In Praise of Manipulation

Keith Dowding (LSE)

Martin van Hees (University of Groningen)

THE LONDON SCHOOL
OF ECONOMICS AND
POLITICAL SCIENCE

In Praise of Manipulation

Keith Dowding
Government Department, London School of Economics and Political Science
Houghton Street, London, WC2A 2AE
United Kingdom
K.M.Dowding@lse.ac.uk
fax: +44(0)20 7955 6352

Martin van Hees*
Faculty of Philosophy, University of Groningen
Oude Boteringestraat 52, 9712 GL Groningen
Netherlands
Martin.van.Hees@rug.nl

September 20, 2006

Abstract

Many theorists believe that the manipulation of voting procedures is a serious problem. Accordingly, much of social choice theory examines the conditions under which strategy proofness can be ensured, and what kind of procedures do a better job of preventing manipulation. In this paper we argue that democrats should not be worried about manipulation. Two arguments against manipulation are examined. First we discuss the 'sincerity argument', according to which manipulation should be rejected because it displays a form of insincere behaviour. We argue however that a distinction can be made between sincere and non-sincere manipulation and we show that a familiar class of social choice functions is immune to insincere manipulation. We then discuss the 'transparency' argument against manipulation. We argue that (sincere or insincere) manipulation may indeed lead to non-transparency of the decision making process, but that, from a democratic perspective, such non-transparency is often a virtue rather than a vice.

1 Introduction

Results from social choice still drive many of the ideas in modern voting theory. We learned from Arrow that there was no general will for a voting mechanism to track – at least if the will is interest based.[1] We have learned from subsequent developments in social choice theory that the decisions of a voting body depend on features of the decision mechanism itself as

---

*We would like to thank Nick Baigent, Anne Gelling, Aki Lehtinen and Hugh Ward for their comments on this paper.

[1] If the will is epistemic then voting might track the truth, see Bernard Grofman and Scott L. Feld 'Rousseau's General Will: A Condorcetian Perspective' *American Political Science Review* (1988) 82: 567-576 for an early perspective on this.

much as the preferences of the members of the voting body; and depend as much on the preferences expressed as any preferences that might underlie that expression. In other words we have learned that decision mechanisms are manipulable.[2] That is, they all display the quality that gives individual voters an incentive to misrepresent their correct orderings for personal advantage. Voting mechanisms are games and not simply procedures for tracking some truths about the underlying preferences or interests of the electorate.

The manipulability of decision mechanisms has generally been viewed as unfortunate. It allows some voters to manage to get their way behind the backs, so to speak, of other voters. However, not all news about manipulation is bad news. Miller has shown that under successive procedures – where proposals are voted up or down in sequence – strategic voting by all can lead to results that are Pareto-preferred to the ones that result if all vote naively. Similarly he shows that under amendment procedures – where proposals are explicitly paired for votes – strategic voting reduces the set of possible outcomes in a manner preferred by the majority.[3] Riker, who popularized the unfortunate effects of Arrow's theorem and manipulation in political science, also sees some advantages in manipulation for heresthestic politicians to transform the political landscape in ways that might be socially advantageous.[4] Others view manipulation with equanimity, not because they see no harm in manipulation itself, but because they believe its practical importance has been overplayed and that manipulation rarely occurs in the political world.[5] Other social choice writers view manipulation as deeply problematic and accordingly examine under what conditions strategy-proofness might be ensured, and what procedures do a better job of preventing manipulation from ever occurring.

Perhaps the first debate in social choice concerning manipulation occurred when Condorcet and others found fault with Borda's rank-order count, proposed to the French Academy of Science in 1770, for being open to strategic misrepresentation of preferences.[6] Even today the debate between defenders of the Borda count and those who privilege the Condorcet winner is essentially concerned with issues about manipulation and how likely it is to occur under different voting schemes and populations. In order to manipulate successfully agents must know (a) their own preferences, (b) the preferences of all other voters, and (c) the technical operation of the voting mechanism.[7] These are quite stringent conditions in the real world, and in large populations successful manipulation is difficult to achieve. Surprisingly, however, few writers have ever queried the normative status of manipulation itself. With the exceptions we have noted, writers simply assume that the manipulation of voting schemes is abhorrent and then consider the possibilities for avoiding strategic voting. As a result, debate concerns

---

[2] Alan Gibbard, 'Manipulation of Voting Schemes: A General Result' *Econometrica* 41 (1973): 587-601; See also Mark Satterthwaite, 'Strategy Proofness and Arrow's Conditions' *Journal of Economic Theory* 10 (1975): 187-217.

[3] Nicholas R. Miller, 'Graph-Theoretical Approaches to the Theory of Voting' *American Journal of Political Science* 21(1977): 769-803.

[4] William H. Riker, *Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice* (San Francisco: W. H. Freeman and Co, 1982); William H. Riker, *The Art of Political Manipulation* (New Haven: Yale University Press, 1986).

[5] Gerry Mackie, *Democracy Defended* (Cambridge: Cambridge University Press, 2003).

[6] Iain McLean, 'The First Golden Age of Social Choice, 1784-1803' in W. A. Barnett, H. Moulin, M. Salles and N. J. Schofield (eds) *Social Choice, Welfare, and Ethics: Proceedings of the Eighth International Symposium in Economic Theory and Econometrics* (Cambridge Cambridge University Press, 1995).

[7] The second demand, knowledge of the preferences of the others, is *not* needed if an individual has a dominant strategy. However, the main result on manipulation – the Gibbard-Satterthwaite theorem – derives from a more general result that there is no non-dictatorial rule in which all individuals always have a dominant strategy. Gibbard, 'Manipulation of Voting Scheme'; Satterthwaite, 'Strategy Proofness'.

how likely manipulation is,[8] or how it might be avoided.[9]

There are, however, two important exceptions to the lack of attention to the normative underpinning for the quest of strategy-proofness, and they constitute two different arguments in favour of strategy-proofness. The first argument is what we call the *sincerity argument* against manipulation. On this view, manipulation of a social choice function is an instance of a form of insincerity or dishonesty that is undesirable or even immoral. The insincerity consists of the discrepancy between the true preferences and the submitted preferences; the things that an individual says she prefers (her submitted preferences) are not her real ones. An example of the use of the sincerity argument is Borda's famous negative reaction to those who pointed out the manipulability of his rule: 'My scheme is only intended for honest men'.[10] That the idea that manipulation is a form of dishonesty is held by many is indicated by the opening line of a recently published (excellent) introduction to the literature on manipulation: 'Were honesty always the best policy, this indeed might be a better world'.[11]

There is an interesting parallel between the sincerity argument and the famous argument put forward by Riker against populism.[12] Riker does not so much deplore the possibility of insincerity as argue that the results from social choice theory show that a conception of democracy as a representation of the honest views of the population should be abandoned. He uses two arguments against the idea that a voting mechanism truly or sincerely represents the preferences of the electorate (or 'the will of the people'). The first is directed against a strong notion of populism and follows directly from Arrow's theorem – which is taken to imply that there is nothing outside of the way of measuring the people's will by which to judge the people's will. Hence there is no way of telling which of two voting mechanisms which differently aggregate the same preferences is the correct one. The second argument is directed against a weaker form of populism that recognizes that there might not be anything that is non-controversially 'the will of the people' but once a voting mechanism has been accepted as the process of decision-making within a community whatever it produces can be considered the people's will. Riker argues from Gibbard-Satterthwaite that as all voting mechanisms are manipulable we can never be sure that the result of any ballot has not been manipulated. Hence we can never claim that the results of any ballot represents the will of the people in this weak form. Riker's conclusion is that the results of all ballots are 'arbitrary' and cannot be trusted. Although Riker does not argue that manipulation is undesirable or even immoral because it is a form of insincerity (as the sincerity argument is formulated by us), he does argue that manipulation entails that the collective outcome need not reflect the true preferences of the electorate. Hence he also invokes a notion of 'true' or 'sincere' preferences that may differ from the preferences that individuals actually express in a decision process.[13]

---

[8]Gary W. Cox, *Making Votes Count* (Cambridge: Cambridge University Press, 1997); Mackie, *Democracy Defended*.

[9]For example, John S. Dryzek and Christian List, 'Social Choice and Deliberative Democracy: A Reconciliation' *British Journal of Political Science* 33(2003): 1-28.

[10]Recounted in Duncan Black, *The Theory of Committees and Elections* (Cambridge: Cambridge University Press, 1958), p. 182. Contemporary defenders of the Borda count seem to take a sanguine stance towards the possibility of manipulation. For recent defenders of Borda see Michael Dummett, *Voting Procedures* (Oxford: Clarendon Press, 1984); P. J. Emerson, *Consensus Voting Systems* (Dublin: Hazelwood Integrated College, 1991); Donald G. Saari, 'The Borda Dictionary', *Social Choice and Welfare* 7 (1990): 279-317; Donald G. Saari, *Geometry of Voting* (Berlin: Springer, 1994); Mackie, *Democracy Defended*.

[11]Alan D. Taylor, *Social Choice and the Mathematics of Manipulation* (Cambridge: Cambridge University Press, 2005), p. ix.

[12]Riker, *Liberalism Against Populism*.

[13]See Keith Dowding, 'Can Populism be Defended? William Riker, Gerry Mackie and the Interpretation of

The only writer that we have been able to discover who considers what might be morally problematic about manipulation without invoking the sincerity argument is Mark Satterthwaite, who discusses the issue in his doctorate.[14] Satterthwaite uses five arguments against strategic voting – which may be summarized as inequality, inefficiency, non-transparency of voters' preferences, non-transparency of representatives' preferences, and randomness. In our discussion of Satterthwaite's five arguments we see that the first four refer to possible asymmetries of information, whilst the fifth also refers to the opaqueness or lack of transparency in the result. Thus all five arguments receive their normative force from a certain *non-transparency* of the decision process and we shall therefore refer to them as constituting the *transparency argument* against manipulation. If it would always be in one's best interest to vote for one's most preferred outcome, the answer to the question of how to vote would be completely transparent, and the need to collect further information would be absent. Furthermore, an outside observer who knows the preferences of the voters would be able to easily predict the outcome of the decision process. Hence, none of the five phenomena would emerge.

We thus distinguish a *sincerity argument* and a *transparency argument* against manipulation.[15] Rather than further working out these arguments, we shall first argue that not all forms of manipulation represent insincerity or dishonesty. In fact, we show that there is a familiar class of voting mechanisms – plurality rules – which either induce straightforward sincere voting or allow only for manipulation that is sincere.[16] Familiar voting mechanisms are immune to insincere manipulation and the sincerity argument therefore loses much of its relevance. Stated differently, if one is worried about insincere manipulation, our result might be taken as a strong argument for the plurality rule.[17]

After having thus questioned the relevance of the sincerity argument, we turn to the transparency argument. We do not argue that Satterthwaite is wrong in claiming that the possibility of manipulation can ensue in forms of non-transparency, but we argue that these non-transparencies are not as problematic as is believed. It is not manipulation itself that might be problematic, but inequality of information and understanding. Indeed, we argue that there may well be virtues in manipulation overall. Rather than being concerned about the strategic or gaming nature of voting systems, we should celebrate those aspects as providing incentives for people to understand democratic procedures, learn about each other and engage

---

Democracy' *Government and Opposition* 41 (2006): 327-346 for more detail on Riker's position.

[14]Mark Satterthwaite, *Existence of a Strategy Proof Procedure: A Topic in Social Choice Theory* (University of Wisconsin: Ann Arbor, Michigan, 1973), pp. 5-16.

[15]A possible third argument against manipulation would be based on the idea that we value voting because it enables us to express ourselves – voting has an *expressive value* (Geoffrey Brennan and Loren Lamasky, *Democracy and Decision: The Pure Theory of Electoral Preference* (Cambridge: Cambridge University Press, 1993); Alexander A. Schuessler, *A Logic of Expressive Choice* (Princeton: Princeton University Press, 2000). Our discussion in Section 3 about the relation between feasibility and sincerity suggests, however, that this expressive value need not be undermined by strategic voting. Clearly, if one manipulates sincerely then one hopes that the alternative voted for will indeed be the outcome and in that sense one's choice still expresses one's will. However, insincere manipulation, that is – voting for an alternative in order to improve the chances that another alternative wins – may seem to be less expressive of one's will.

[16]Lest one thinks that 'sincere manipulation' is an oxymoron because the term 'manipulation' carries negative (dishonest) normative connotations, we point out that the lexical root of manipulation is to dextrously handle or manage.

[17]Closely related to our argument is Brian Barry's brief exposition of the idea that there is no proper distinction between sophisticated (strategic) and naive (sincere) voting since the underlying preferences which lead one to vote naively or strategically do not lend themselves to those terms. We simply have preferences and vote accordingly given the situation we are in. See Brian Barry 'Wasted Votes and Mare's Nests: A View of Electoral Reform' – an undated, unpublished book manuscript.

in politics: the possibility of manipulation is often a virtue rather than a vice.[18]

2 Defining Manipulation

There are two elements of manipulation: first, strategic or sophisticated voting; secondly, agenda setting. This paper's focus is on the first aspect of manipulation. The fundamental result about strategic voting is the Gibbard-Satterthwaite theorem.[19] This result is about a specific kind of decision mechanism, to wit *social choice functions*. A social choice function is then a function that specifies for each possible combination of preference orderings (one for each individual) over a set of alternatives exactly one alternative as the outcome. The Gibbard-Satterthwaite result states that any non-dictatorial and non-imposed social choice function is vulnerable to manipulation – that is, expressing one's true preferences is not always a dominant strategy: there will always be a preference profile in which at least one of the individuals has an incentive not to express her 'true' preferences.[20] Alternatively, we can say that the result shows that the only (non-imposed) social choice function that is not vulnerable to manipulation, i.e. that is *strategy-proof*, is a dictatorship.[21]

To make this more precise, let $N = \{1, \ldots, n\}$ ($n \geq 2$) denote the set of all individuals and X the set of alternatives (containing at least three alternatives). We define a *preference game* as any $n + 2$-tuple $G = (R(X), \ldots, R(X), s, \pi)$ in which $R(X)$ is the set of all possible preference orderings over X, s the profile of preference orderings that the individuals actually have (their 'true' or 'real' preferences) and π a mapping from the set of all possible preference profiles into X.[22] The set of all possible preference profiles consists of the n-fold Cartesian product of $R(X)$. This product also forms the set of all strategy profiles, or *plays*, of the game. The strategies that the individuals have thus consist of preference orderings – the mapping π is a social choice function that yields a decision on the basis of the preferences that the individuals submit. The submitted preferences may coincide with the 'true' or 'sincere' preferences of the individuals, as specified by s, but they need not do so. The Gibbard-Satterthwaite results show that it is not always a dominant strategy to express one's true preferences or, equivalently, for any non-dictatorial (and non-imposed) social choice

---

[18]From the start it should be noted that we limit the discussion to strategic behaviour in the context of *democratic* decision making. Hence we do not consider the literature on implementation. For an overview see Eric Maskin and T. Sjöström, 'Implementation Theory' in Kenneth J. Arrow, Amartya K. Sen, Katoro Suzumura, (eds) *Handbook of Social Choice and Welfare* (Amsterdam: Elsevier Science, 2002). Nor do we discuss the issue of the implications that the results of manipulation of voting procedures have for the manipulability of decision procedures in general.

[19]Gibbard 'Manipulation of Voting Schemes'; Satterthwaite, 'Strategy Proofness'.

[20]Non-dictatorship means that there is no individual whose top alternative is always chosen. The condition of non-imposition says that there are at least three distinct alternatives such that for each of them there is at least one situation in which it will be the outcome.

[21]Social choice functions do not exhaust the set of possible social decision mechanisms. A social choice correspondence, for instance, is a function that assigns a (non-empty) set of outcomes to preference profiles, rather than a unique alternative. Although several impossibility results have been formulated in the context of social choice correspondences, for such mechanisms there is no general result comparable to the Gibbard-Satterthwaite theorem. A different kind of social decision mechanism is one that assigns lotteries over the outcomes. Although possibility results here are possible – that is, there are non-dictatorial decision mechanisms that assign lotteries and that are strategy-proof – they also have some other undesirable properties. See Alan Gibbard, 'Straightforwardness of Game Forms with Lotteries as Outcomes', *Econometrica*, 46 (1978): 595-614.

[22]An ordering $R \in R(X)$ is a transitive and complete binary relation. The symmetric and anti-symmetric parts of it are denoted by I and P, respectively.

function π there is at least one possible profile of true preferences s such that s is not a Nash-equilibrium of the preference game $G = (R(X), \dots, R(X), s, \pi)$: there is an individual i and a profile s' which is the same as s except that i submits a different ordering, and that leads to an outcome $\pi(s')$ that i strictly prefers – in terms of her real preferences – to the outcome $\pi(s)$ that would have resulted otherwise. In other words, for any such social choice function there is at least one situation in which one of the individuals has an incentive not to submit her true preferences. A *strategy-proof* social choice function is then defined as any social choice function π such that for all possible preferences profiles s, s is a Nash-equilibrium of the game $G = (R(X), \dots, R(X), s, \pi)$ – nobody has an incentive to submit an ordering different from his or her real one.

3 Sincere Manipulation

Consider the 2000 US elections. Some of the supporters of Ralph Nader voted for Al Gore rather than for Nader, their most preferred alternative. Their rationale for doing so was clear: these voters (although not enough as it turned out) voted for Gore since they did not believe that Nader stood a real chance of gaining the presidency. In fact, a vote for Nader would increase the probability that Bush would win – an outcome they found inferior to Gore winning the election. Galbraith and Rae define such voting as *tactical voting*: 'voting for other than one's preferred party of candidate in order to increase the probability of a satisfactory overall outcome'.[23] The 'satisfactory overall outcome' is then described as the outcome that results if an unwanted alternative – usually the incumbent party – is replaced by a party of the opposition, although that party need not be their most preferred party.

Clearly, such a tactical vote is not quite an example of manipulation as it is defined in the Gibbard-Satterthwaite theorem. The main reason for this is that such manipulation is not defined in terms of probabilities, but in terms of the absence of equilibriums. In the 2000 US election it was almost certainly the case that each voter submitting a vote for their most preferred candidate formed a Nash-equilibrium: a single Nader-supporter could not have affected the outcome by switching his vote to Gore when all others voted for their most preferred outcome. But suppose, for the sake of the argument, that such an unlikely scenario could have obtained: if everybody voted for their most preferred outcome Bush would win, but if some Nader supporter switched his vote to his second-best outcome – that is, to Gore – Bush would lose the election and Gore would win. Suppose furthermore that the Nader supporter who could bring about such a swing was aware of this fact and therefore decided to vote for Gore. In this case, the vote for Gore would constitute an act of manipulation – by not voting for her most preferred alternative (Nader) an outcome results (Gore) which the voter prefers to the outcome which would have resulted otherwise (Bush).

Though a form of manipulation, we call this *sincere manipulation* because the individual votes for the outcome (Gore) that she prefers most from among the feasible outcomes. She acts sincerely in the sense that she votes for an outcome in order to bring *that* outcome about. To forgo pursuing our most preferred outcomes for reasons of feasibility is not insincerity. We do not think that someone has revealed a preference for beer over champagne when they buy beer rather than champagne, when we know their finances will not stretch to a bottle of the bubbly. Rather they have revealed a preference for beer, given their budget constraints.

---

[23] John W. Galbraith and Nicol C. Rae, 'A Test of the Importance of Tactical Voting: Great Britain, 1987' *British Journal of Political Science* 19(1989): 126-136, p. 126.

Feasibility of options may also be determined by the actions of others.[24] Even though I would very much like to buy the latest Harry Potter as soon as it is out, I may decide not to go to the bookshop, since I do not want to queue and know that the book will almost certainly be sold out on the first day. Surely we do not want to maintain that staying at home is acting insincerely. Similarly we can think that some forms of manipulation simply reveal one's preferences given the constraints of what others think or do. We spend our votes wisely on the feasible alternatives, just as we spend our money wisely given our budget constraints.[25] Furthermore, it might be thought that voting for Gore reveals a preference for how much one dislikes Bush. One would sooner vote for one's second preferred candidate because one wants to stop one's least preferred candidate.

It can only be claimed that manipulation is always insincere if the notion of insincerity is stretched very far, and in fact too far. To say that a person voted insincerely may say something about *how* she takes account of feasibility constraints, but we cannot say that it is insincere as such to take account of such constraints. We not only label the neglect of feasibility considerations as naive or stubborn rather than sincere, but to define sincerity in terms of such neglect seems to be at odds with itself. After all, someone is hardly insincere if they try to accomplish their objectives as well as they able. In fact, precisely for this reason many Nader supporters were criticized for voting for Nader in the 2000 US elections. If they had voted for Gore the result would have been more in accord with their overall preferences.

Thus, not all forms of manipulation are insincere. It is a form of sophisticated voting, and acting in a 'sophisticated manner' may in fact be more sincere than acting naively since it implies that one is better able to reach one's objectives.[26] But *which* forms of manipulation can indeed be said to be sincere? To give an exact definition, we introduce some extra notation. For any ordering R and any $x \in X$, let R^x denote the ordering that results from bringing x to the top of R while leaving the ranking of all other alternatives unchanged. Formally, for all $y, z \in X - \{x\}$ we have yR^xz if and only if yRz, and for all for all $y \in X - \{x\}$ we have xP^xy.

Definition 1 *Given a preference game $H = (R(X), \ldots, R(X), s, \pi)$, we say that a group of individuals $S \subseteq N$ can manipulate sincerely if, and only if, there is some $\hat{s} = (\hat{R}_1, \ldots, \hat{R}_n)$ and some $y \in X$ such that*

1. *$\hat{R}_i = R_i$ for all $i \notin S$;*

2. *$\hat{R}_i = R_i^y$ for all $i \in S$;*

3. *$\pi(\hat{s}) = y$ and $yP_i\pi(s)$ for all $i \in S$;*

4. *for all profiles $\tilde{s}$ such that $R_j = \tilde{R}_j$ for all $j \notin S$: if $\pi(\tilde{s}) \neq y$, then $yP_i\pi(\tilde{s})$ for some $i \in S$.*

Thus a group of individuals S is said to be able to manipulate sincerely if, assuming that everybody else expresses their real preferences, (Clause 1), they can make some alternative

---

[24]In fact, budget constraints are often also the result of actions by others.

[25]One interpretation of 'authenticity' in existentialist writings might include taking into account feasibility constraints when choosing actions as 'inauthentic' or 'insincere'. However, outside of such writing we would not normally think so.

[26]The term 'sophisticated voting' is taken from Robin Farquharson, *Theory of Voting* (New Haven: Yale University Press, 1969) although he uses the term to designate a specific kind of strategic voting.

y the outcome by bringing it to the top of their ranking (Clause 2), and if they all prefer y to the outcome that would have resulted otherwise (Clause 3). Furthermore, there is no other outcome that the group can achieve by a unilateral switch and which they all prefer to y (Clause 4).

Before examining the possibility of having decision procedures that only allow for sincere manipulation, it may be helpful to give an example of what would count as *insincere manipulation*. Let $X = \{x, y, z\}$, and let the decision procedure stipulate that an alternative x (x may for instance be the status quo) is chosen unless there is some other alternative that is more often the top alternative than any of the other alternatives. Next assume that there are four individuals with the following preferences:

R_i	R_j	R_k	R_l
y	x	z	z
z	y	x	y
x	z	y	$x.$

If each individual submits her actual preferences, the outcome will be z. That outcome is however individual j's least preferred alternative. Suppose that j therefore submits an ordering in which y rather than x is the top outcome. Assuming that the other individuals continue to express their real preferences, we see that the two alternatives y and z both have the same number of top positions. The social choice function will therefore choose x as the social choice (note that the mechanism chooses x unless there is an alternative which occupies more top positions than any other alternative). Clearly, j manipulates the social choice function. Moreover, the manipulation can be labelled insincere in the sense that j tries to realize an objective which differs from y by an action (i.e., saying that y is her best alternative) that suggests differently.

Now one may want to argue that such behaviour may in fact also be a form of sincere manipulation – say on the grounds that insincerity entails a conscious attempt to withhold information to the others whereas that need not be the case here. We could indeed try to formulate a stronger notion of sincere manipulation than the one used here. However, our primary aim is to examine the possible existence of social choice functions that are immune to insincere manipulation. Clearly, the weaker the notion of sincere manipulation that we use, the stronger the corresponding notion of immunity to insincere manipulation, and thus also the stronger any positive result about the possibility of such immunity.

Definition 2 *A social choice function π is immune to insincere manipulation if for each $S \subseteq N$ and each preference game $H = (R(X), \ldots, R(X), s, \pi)$: if S has a manipulation possibility at s, then S can manipulate sincerely at s.*

Our main result in this section shows that the class of social choice functions that are immune to insincere manipulation can be characterized in terms of the following property:

Definition 3 *A social choice function π satisfies top-monotonicity iff for all profiles s, s' and all i: if $\pi(s)P_i'x$ for all $x \in X - \{\pi(s)\}$ whenever $R_i \neq R_i'$, then $\pi(s') = \pi(s)$.*

Top-monotonicity states that for any profile s, if some of the individual change their preferences by putting the outcome corresponding to s, $\pi(s)$, at the top of their preference ordering, then the outcome in the new profile should also be $\pi(s)$. Note that top-monotonicity is a stronger condition than conditions of monotonicity that also stipulate that the individual

rankings of the alternatives different from $\pi(s)$ stay the same. This we do not assume here - the move of $\pi(s)$ to the top of a person's preference ordering need be the only change in that person's preferences. On the other hand, top-monotonicity is a weaker requirement than the monotonicity condition (labelled 'strong positive association') introduced by Muller and Satterthwaite in the context of the analysis of manipulation.[27] Strong positive association says that an outcome should remain the outcome *whenever* it has gained extra support. With top-monotonicity the extra support always consists of having been moved to the top of a person's preference ordering, whereas strong positive association only demands that the outcome has been moved up some places but not necessarily to the top. Strong positive association was shown by Muller and Satterthwaite to be equivalent to strategy-proofness. Our main result of this section shows that the weaker requirement of top monotonicity is equivalent to demanding immunity from insincere manipulation.

Theorem 1 *A social choice function π is immune to insincere manipulation if, and only if, it satisfies top-monotonicity.*

This result already takes away much of the relevance of the existing results, most notably the Gibbard-Satterthwaite theorem, about the impossibility of having voting rules that always elicit sincere behaviour. The notion of manipulation as used in the Gibbard-Satterthwaite theorem does not render the idea of insincere voting correctly. Our result shows that if we are only worried about insincerity as such we need not be worried about the theorem. There is a class of voting mechanisms that only allow for sincere manipulation. The class consists of all those social choice functions that are top-monotonic – and the social choice function that is most commonly used, plurality rule, satisfies this requirement.

In other words, if sincerity is all that one cares about then this constitutes a strong argument for the plurality rule. Note, however, that the plurality rule is not the only social choice function that satisfies top-monotonicity and is thus immune to insincere manipulation. Take for instance a quota-rule according to which some alternative x is chosen (where x may stand for the status quo or the chairman's most preferred alternative) unless some other alternative occupies a certain minimal number (the quotum) of the top positions. It satisfies top-monotonicity and yet may assign an outcome which does not occupy the highest number of top positions.

4   The Transparency Argument

If insincerity does not form the problem of manipulation, what *is* the problem? We now turn to the transparency argument against manipulation, and do so on the basis of a discussion of Sattherwaite's five transparency problems with manipulation. The arguments against strategic voting may be summarized as (1) inequality of skills, (2) inefficiency, (3) non-transparency of voters' preferences, (4) non-transparency of representatives' preferences, (5) randomness:[28]

1. *Inequality of skills.* The possibility of manipulation gives some voters an unfair advantage over others. Voting is supposed to be equal – one person one vote – but if some

---

[27]Eitan Muller and Mark Satterthwaite, 'The Equivalence of Strong Positive Association and Strategy-proofness', *Journal of Economic Theory*, 14 (1977): 412-418. See also David-Austen Smith and Jeffrey Banks, 'Monotonicity in Electoral Systems', *American Political Science Review*, 85 (1991): 531-537.

[28]We list the five arguments in a different order to Satterthwaite. See also J. S. Kelly, *Social Choice Theory: An Introduction* (Berlin: Springer Verlag, 1988), p. 103.

voters understand the strategic nature of the voting game and others do not then those with that knowledge gain an advantage. They are more likely to vote strategically. The ignorant will not understand the strategic possibilities and so be disadvantaged.

2. *Inefficiency.* Strategic voting is wasteful. Finding out the preferences of others is a waste of resources. If there are strategic possibilities inherent in the process, then in order to discover this, the voters need to know not only their own preferences, but the preferences of all others. Discovering these preferences is taken to be a wasteful and inefficient activity.

3. *Non-Transparency of Voters' Preferences.* The possibility of manipulation gives people incentives to hide their preferences. If someone can hide their preferences they make it harder for others to manipulate and may make their own strategic possibilities easier.

4. *Non-Transparency of Representatives' Preferences.* Manipulation blurs the voting records of politicians. Politicians may have voted strategically on some occasions. If in assessing the performance of a politician one relies on their voting records only, a distorted picture may emerge: the records do not reveal the reasons why politicians voted as they did. As a result, the voters may have difficulty in ascertaining whether their representative did indeed represent their interests.

5. *Randomness.* Manipulation introduces an element of randomness into the voting process. If there is no dominant strategy, then there may be different ways of trying to secure one's interests as well as possible. In those cases it may be very difficult to predict the outcome of the decision process.

Each of these arguments is about non-transparency: the last three focus on specific types of non-transparency whereas the first two arguments point out possible negative consequences of non-transparency. We shall not dispute the possibility that non-transparency results when individuals need to take account of strategic considerations, but we shall argue that this is not as problematic as these arguments seem to suggest.

Consider first the possibility that the knowledgeable will gain an advantage over the ignorant. This is surely true. But why should we take this as a criticism of the inherent strategic possibilities in some voting games? The criticism entirely relies upon informational asymmetry and this asymmetry is the problem rather than the strategic possibilities themselves. Indeed we might turn this criticism upon its head. In order to understand the strategic possibilities voters must have an understanding of the mechanics of the voting mechanism, of their own preferences – that is, they must not only know what their most-preferred alternative is but also their overall ranking of all alternatives – and they must have a general knowledge of the likely preference profile of the rest of the electorate. They do not need to know this latter in detail – they do not need to know the exact ranking of everyone in the electorate – but they must have a general idea of the underlying structure of the electorate as a whole. Armed with that knowledge they can then judge the strategic possibilities, see what feasible outcomes there are and vote accordingly. Of course, if we want a good working democracy we want voters to have this kind of knowledge. Do we really want an electorate that does not understand how their voting mechanism works? Do we want an electorate that does not know their own preference ordering? Do we want an electorate that does not care what everyone else's ordering is? Not only deliberative democrats want people to debate and understand others' preferences; liberals may also see such knowledge as a desideratum. The inherent

possibilities of strategic voting encourage voters to learn more about their democracy and the views of their fellows. And that, surely, is good for democracy.

Satterthwaite's second problem is the inefficiency that results from strategic voting. One way of interpreting this criticism is that it suggests that all a voting mechanism should do is consult people and tot up their preferences. Encouraging them to do more than give their preferences is wasteful. The objection reduces the democratic process to a counting device and no more. This view is a rather mean and meagre vision of democracy. It is not only in modern accounts of deliberative democracy that the process of discussion and deliberation is thought to lead to a better populace as well as superior decisions; similar thoughts can be found in the classic defences of democracy.[29] Not only does such a skimpy account of democracy deny the Aristotelian idea that we are a political animal, it also denigrates the whole process of forging coalitions and working with others. Rather than seeing the process of discovering others' preferences and forging coalitions as a waste of our precious resources, we might see it as a positive part of being human. Knowledge of others' preferences enables people to realize that their first preference may not be possible. They may begin to compromise, to forge coalitions, to make allies. Again it is not only deliberative democrats that see advantages in this process, as a process; liberals can too. By deliberating and discussing people improve their negotiating and bargaining skills, and they can begin to understand and empathize with others' viewpoints. In other words, the fact that aggregative systems are manipulable encourages the very deliberation that makes politics what it is. Satterthwaite argues that such behaviour will concentrate the efforts of people on working out sophisticated voting strategies rather than concentrating on the content of the committees' decisions. It cannot be simply assumed that the resources expended on discovering others' preferences is a dead-weight loss, as Satterthwaite's argument seems to assume. Working out what one's views are on the issue and what others' views are and why we all take those views is part-and-parcel of working out the content of a committee's decision. The possibility of manipulation is not a vice here but a virtue for democracy since it enables due consideration.[30]

Surely, however, the critic responds, the fact that one person may manipulate a result when others cannot must be wrong. That person gains an advantage over everyone else that is inequitable and against the very idea of democracy that each person's views (at least as expressed in their votes) should count equally. The fact that the underlying structure of preferences means that some – the Nader – supporters can vote for their next most preferred candidate but Gore and Bush supporters cannot gives Nader supporters a strategic advantage. It is almost as though they have two candidates, their most favoured candidate and, if he is not feasible, their second most favoured candidate. Since the underlying structure of preferences make only Bush and Gore feasible candidates, this means their supporters do not have any viable response to the Nader supporters' strategic advantage. This objection only works if one believes that the equality condition of voting means not 'one person one vote' but 'one

[29]A. de Tocqueville, *Democracy in America* vol 1. edited by Phillips Bradley (New York: Vintage Books [1835] 1945), pp. 246-63; John Stuart Mill, *Considerations on Representative Government* in John Gray (ed) *John Stuart Mill: On Liberty and Other Essays* (Oxford: Oxford University Press, [1861] 1991), pp. 225-6, 255; Hannah Arendt, *The Human Condition* (Chicago: Chicago University Press, 1958); Benjamin Barber, *Strong Democracy* (Berkeley: California University Press, 1984), p. 155; Robert A. Dahl, *Democracy and Its Critics* (New Haven: Yale University Press, 1989), pp. 97-105.

[30]A different argument against the charge that manipulation is inefficient is developed by A. Lehtinen, 'Welfarist Evaluation of Voting Schemes', *mimeo*, 2004. Taking a dynamic perspective and making use of simulations, he shows that manipulation of some well-known voting rules lead to a sequence of outcomes that yields a higher average utility than the sequence of non-strategic (sincere) voting.

person one vote for your top alternative'. Nader supporters only get one vote, and if they choose to use that to support Gore that cannot be considered unfair. Indeed we cannot put conditions on what people choose to do with their one vote. If they wish to vote for their least preferred candidate, abstain or spoil the ballot paper, that is up to them.

Satterthwaite thinks, thirdly, that people will have an incentive to hide their preferences. Bush supporters, for instance, have no third alternative to go to should Nader supporters switch to Gore. This gives Bush supporters the incentive to pretend they are not going to vote for Bush. They might claim that they are Gore supporters and so lead Nader supporters to believe that their votes for Gore are not required thus allowing them to express their top preference in a vote for Nader. Or they may pretend they support Nader to lead Nader (and perhaps some Gore supporters) to think that Nader is a feasible candidate. Bush supporters do not have a strategic vote response to the strategic possibilities open to Nader supporters, but they do have a strategic response. They could lie and pretend to be other than they truly are.[31]

We might begin by noting that to the extent that such deception is a bad feature of voting mechanisms that are manipulable, it is also a feature of ones which are not. In a two-candidate majority election there is no possibility of strategic voting. But there is still the possibility of the lying strategic manipulation just described. Imagine that in a two candidate race where everyone votes for their top alternative, Gore is expected to beat Bush by a whisker. One response for Bush supporters is to pretend they are Gore supporters, make his victory seem a shoo-in and encourage abstention from genuine Gore supporters. Even if abstention is not allowed, such tactics might work. Imagine someone in 2000 who found themselves in agreement with Bush on many issues but feared his competence. She might decide to vote for Gore as the safe option. However, if she believed that Gore was a shoo-in she might indulge her expressive preference and vote for Bush. The non-transparency argument is too strong as directed against the manipulability of voting games, since it applies to non-manipulable ones too.

In fact the possibility of such deceitful behaviour is not really a problem of democracy. It is a problem of social life itself. People are capable of deceit in their everyday lives, which is why evolution has also fitted us with ways of recognizing deceit in others. The fact that Bush supporters may have aided Nader in 2000 probably gave greater emphasis to Nader's natural supporters to vote Kerry in 2004. But again, the possibility of deceit gives people incentives to find out more information about the preferences and activities of others engaging in the political processes around them. Again, we submit, this is a virtuous not an inherently vicious feature of democratic practice. This objection is again based on the asymmetry of information that exists between people. I tend to know my own preferences better than others do. And I may have an incentive to hide those preferences. But that may give you an incentive to find out what my preferences are. That is not inherently bad, and it is not inherently wasteful. It is part-and-parcel of human life.

Next consider the fourth argument, the non-transparency of representatives' preferences. Why is this a bad thing? Of course, it makes the job of political scientists and historians harder, but that is hardly worth taking into consideration.[32] Rather it makes it harder for

---

[31]There is no evidence that Bush supporters pretended they were really Gore or Nader supporters but there is evidence that in both the 2000 and 2004 elections Bush supporters aided Nader to get on the ballot in some states and provide some facilities for his campaign.

[32]It is certainly possible to reconstruct the preference profiles of representatives, and their views, both from their voting records and their recorded views in speeches, letters, writings and contemporary reports, even

constituents to know what the 'true' preferences of their representatives are. But surely, all constituents really want to know is the voting record of their representatives. Constituents, like political scientists and historians, can make judgements about whether representatives voted their top alternative or not. All they really need to know is how their representatives vote. It is in terms of how their representatives vote that constituents will decide whether or not to support them next time. Indeed, to some extent, we might not care what our representatives 'really' think. What we are interested in is how they vote, and how they defend our interests. Again, what is important here is the information flow between the actions, including the speech-acts, of representatives and the public at large. Again it is not the manipulation that is problematic, but asymmetric information.

Satterthwaite's fifth argument is that manipulation leans to the possibility of randomness of the voting process. Strictly speaking 'randomness' means haphazard, or occurring heedlessly without aim or principle, or statistically speaking each potential outcome is unpredictable. Does manipulation induce an element of randomness into collective decisions? Satterthwaite discusses this in the context of people casting strategic votes through miscalculating the preferences of others and so the result depends not only on voters' preferences but also on their unreliable subjective estimates of others' preferences. However, another way of thinking about randomness occurring in collective decisions is to follow what Riker means by 'arbitrariness'.[33] In the absence of a clear equilibrium outcome (there may be several or there be none at all as is the case in the voting paradox), it may be difficult to predict the outcome: there is no obvious way of playing the game.

This randomness objection is perhaps the strongest of all the transparency objections. However, it can be argued that the advantages that we ascribe to the possibility of manipulation, and which all relate to the ensuing importance of obtaining information about one's political environment, are even greater if there are situations like those of the voting paradox in which it is very difficult to predict the outcome. Indeed, as far as any claim might be made that the results of a voting mechanism are hard to interpret because of the possibility of strategic voting, manipulation requires that the manipulator has a good enough idea of what the underlying structure of individual preferences looks like. The greater the incentives to manipulate, the greater the incentive to obtain information about the voting situation. Assuming that it is important from a democratic perspective to have such incentives, we again see that non-transparency is an advantage rather than a disadvantage.

5   Conclusion

In this paper we have examined two arguments against manipulation in voting systems. The 'sincerity argument' suggests that manipulation displays insincerity since it entails not expressing one's 'real' preferences. Instead, for their advantage voters may decide not to reveal their true or sincere preference. We have demonstrated a distinction between sincere and insincere manipulation and have subsequently shown that a familiar class of social choice functions is immune to insincere manipulation. Translated to a voting scheme, the plurality rule is immune to insincere manipulation. The 'transparency argument' against manipulation

---

in highly strategic contexts as Mackie *Democracy Defended*, chs. 9-15 has amply demonstrated. See also technical fixes to judge strategic misrepresentation in roll call and other data, Thomas König, Daniel Fink and Stephanie Daimer 'Ignoring the Non-Ignorables? Missingness and Missing Positions' *European Union Politics* 6 (2005): 269-290.

[33]Riker, *Liberalism Against Populism*.

suggests that manipulation leads to opaqueness in the decision process. We discussed the 'transparency argument' against manipulation on the basis of Satterthwaite's five arguments against manipulation. Many of his arguments are really about asymmetries in information rather than about the properties of voting schemes themselves, and so the basis of any harmful effects are these power imbalances. From this fact that informational asymmetries form the basis of Satterthwaite's argument we argued that whilst (sincere or insincere) manipulation may indeed lead to non-transparency of the decision making process, from a democratic perspective such non-transparency of voting schemes is less problematic than it is taken to be. In fact, we suggest that the conception of democracy that leads to the negative appraisal of non-transparency is rather unattractive. Democracy is more than simply collecting votes. It is also about knowing one's political environment, about understanding the preferences of others, and about knowing the nature of the process by which collective outcomes come about. Of course, this does not mean that manipulation may not have negative consequences – not any degree of non-transparency is desirable. Whether particular instances of manipulation are indeed troublesome will depend on the specifics of the voting situation, on how likely it is that individuals will indeed manipulate, on the specific ways in which they can do so (e.g. sincerely or non-sincerely), on the kind of equilibriums that may or may not arise, and so forth. It may well turn out to be that the price – inequality, inefficiency or randomness – will be very high in some cases.

One may, however, assess specific instances of manipulation negatively without thereby committing oneself to a negative appraisal of the possibility of manipulation as such. Indeed, the lessons learned about manipulation in general can be used to learn how voting situations can be manipulated, and about the powers people have for strategic voting and the good results that these can accomplish.[34] They should not be used to denigrate democracy as though democracy were simply the aggregation of pre-set preferences. The very fact that politics is a game, where the strategies include argument, persuasion, the understanding of the position of oneself and others, entails strategic interaction. In this regard strategic manipulation can be recognized as a virtue rather than a vice. The possibility of manipulation gives incentives for citizens to learn more about those around them, the political situation and the democratic process itself. In that sense we think the possibility of manipulation is to be welcomed.

---

[34]This is the stance taken by Riker, *Political Manipulation.*

Appendix

Proof of Theorem 1. Let π be top-monotonic. Assume there is a preference profile $s = (R_1, \ldots, R_n)$, a profile $s' = (R'_1, \ldots, R'_n)$ and a group of individuals S such that for all $i \notin S$, $R_i = R'_i$, and for all $i \in S$, $\pi(s')P_i\pi(s)$. In fact, assume $\pi(s')$ is the best that S can do: there is no profile s^* which also differs from s only because the members of S have a different preference ordering but which leads to an outcome $\pi(s^*)$ that all $i \in S$ strictly prefer to $\pi(s')$. Consider the profile $\tilde{s}$ which is the same as s except that all $i \in S$ have $\pi(s')$ as their unique top outcome. Comparing $\tilde{s}$ and s', we see that top-monotonicity entails that $\pi(\tilde{s}) = \pi(s')$. This means that the individuals in S can manipulate sincerely at s (to wit, by submitting $\tilde{s}$).

We prove the converse direction by contradiction: assume that π is immune to non-sincere manipulation and that π is not top-monotonic. We then have some preference profiles s, s' with $\pi(s) \neq \pi(s')$ such that for all i either $R'_i = R_i$ or the unique top outcome in R'_i is $\pi(s)$. In particular, let $(R_1, \ldots, R_n)$ be a 'smallest' profile for which this is true: if one of the individuals, say i, who changed their preferences would still have expressed R_i the outcome would have remained $\pi(s)$. Suppose that s' is the real preference profile. If i submits the preference ordering R_i while the rest of the voters continue to express the same preference ordering as in s', the outcome is $\pi(s)$. Since $\pi(s)$ is the unique top outcome of R'_i, the change from R'_i to R_i is a form of insincere manipulation, which is a contradiction. □