

INCOMPLETE DRAFT: NOT FOR
CIRCULATION WITHOUT PERMISSION: The
Effects Of Manipulation on Voting Outcomes
Under the Plurality Rule: A Thought-Randomized
Experiment *

Vicky Barham¹, Herman Demeze², and Roland Pongou³

¹University of Ottawa

²University of Bielefeld

³University of Ottawa

March 15, 2017

Abstract

Whereas the literature on mechanism design typically takes the view that a well-designed social choice mechanism should ideally implement the outcome that is selected when individuals report their preferences truthfully, this paper considers an alternative metric to use in comparing alternative voting mechanisms, namely, the proportion of the population which benefits from manipulation of the mechanism. An important feature of the voting populations that we study is that some proportion of the voters may be highly partisan, and will always vote for their preferred

*We thank Yukio Koriyama, Muhamet Yildiz, Alex Wolitzky and Nicolas Côté for useful discussions and comments. A portion of this paper was completed when Pongou was visiting the MIT Department of Economics; he thanks this institution for its hospitality. Correspondence can be addressed to Pongou at: rpongou@uottawa.ca.

candidate, whereas the remaining voters may find that it is payoff-maximizing to vote strategically. We study two notions of manipulation: Gibbard manipulations, in which any voter who does not vote sincerely must prefer the outcome of the manipulated vote to the sincere voting outcome, and a new concept of manipulation, Nash manipulations, which requires all voters to be choosing best responses to the votes of the other candidate, and any candidate who does not vote for their preferred candidate must weakly prefer the manipulated outcome to the outcome which would prevail if they voted sincerely. We observe that both notions of manipulation lead to outcomes that are Pareto non-comparable with respect to the sincere voting outcome, and we calculate exact minimum and maximum bounds on the number of voters who benefit from strategic voting as a function of the number of voters, the number of candidates, and the number of highly partisan voters. Subsequently, we disaggregate the overall effect and calculate exact bounds for both the sincere and on the strategic voting populations, which provides us with additional insight into how these gains are shared. In some sense, our analysis can be viewed as a cautionary tale against being overly focused on designing collective choice procedures which always select the outcome which prevails when electors vote sincerely, because this may not be an outcome that is particularly worth protecting. In effect, the plurality rule becomes vulnerable to manipulation when it selects an outcome that is the preferred outcome of only a minority of voters, but it is precisely in such circumstances that manipulation leads to an outcome which is Pareto non-comparable to that selected under sincere voting. A voting procedure with respect to which manipulations typically benefit a large proportion of the voting population might therefore be seen as more desirable than those for which manipulations are typically only advantage a narrow minority of citizens.

1 Introduction

Should you vote with your head or your heart? In tightly contested elections, this is a question over which many voters agonize. Voting one's conscience, even if this is for a candidate with no realistic chance of winning, is commonly viewed as a principled approach to the exercise of one of the most fundamental rights of citizens in a democratic society: the right to participate in the election of members of the country's government. But voting one's heart is not without risk: in jurisdictions using the plurality rule, sincere voting often paves the way to the election of an unpopular leader; if more voters were to cast their ballots strategically, this might result in better outcomes. So should strategic voting be encouraged or deplored? This question is the central focus of this paper. We are the first to show that it is possible

to calculate exact upper and lower bounds on the proportion of the population which benefits from manipulation of the voting procedure under the plurality rule; in many settings it is a clear majority of voters who will be made better off. Our results also highlight the impact of those voters who *always* vote sincerely on the distribution of gains from strategic voting. Our analysis can consequently be interpreted as a thought-randomized experiment: if in any particular population Nature randomly assigns voters to either the set of strategic voters, or to the set of invariably sincere voters, then the impact of strategic voting for any of the realized preference profiles will fall between our minimum and maximum bounds.

Why is it important to take account of invariably sincere voters when studying the impact of manipulations of the plurality voting procedure? Empirical investigations of the prevalence of strategic voting suggest that in any given election the vast majority of voters will vote for their preferred candidate (??, ?). What is not clear is whether it is strategic voters who are rare, or whether it is only in rare circumstances that voters who prefer candidates who are certain to lose believe that they can influence the outcome of the election by voting strategically (??, ?). A multitude of factors may in fact explain why some voters *always* vote for their most-preferred candidate (or party). Voting strategically is often portrayed as equivalent to telling a lie — indeed, social choice theorists refer to this as ‘misreporting your preferences’ — and there is considerable experimental evidence (see, for example, ??, ??, ?) which suggests that a significant proportion of the population is lie-averse, and will therefore almost certainly always vote sincerely.¹ Moreover, many voters feel tremendous party loyalty: the psychic costs of voting for any party (or candidate) other than the party they love outweigh the gains that would accrue from blocking the election of a party they intensely dislike (??, ? ?). Regardless of the underlying explanation for the existence of invariably sincere voters, it is critical to take account

¹At some level, this is somewhat surprising. Although there is a moral prohibition on lying, philosophers agree that one cannot lie without asserting (verbally or otherwise) a false proposition and no proposition can be asserted by casting a ballot. In general, on a Kantian view, an act is obligatory if and only if it is *always* desirable that every person perform that action; if voting strategically averts a worse outcome than would prevail if voters report their true preferences then the categorical imperative is not violated by either sincere or strategic voting. In contrast, virtue theorists would take the view that voters, *regardless* of their personal preferences, should always cast their ballots for the candidate who would be preferred by a virtuous (that is, moral) voter. What is less clear, however, is whether or not a virtuous voter would determine their preferred candidate by taking account of the likelihood that a ballot cast for that candidate makes it more or less likely that the candidate wins the election. In contrast, consequentialists (such as ?) argue that citizens have a moral duty to vote strategically, when to fail to do so leads to an outcome that is less preferred - either from the viewpoint of themselves as individuals, or for society as a whole. In effect, it is sincere voting, and not strategic voting, that is the morally reprehensible choice.

of the fact that such voters exist, as this is likely to shape the circumstances in which strategic voters can manipulate outcomes. Below, we separately calculate exact bounds on the gains and losses from manipulation for invariably sincere and strategic voters, and show how the tightness of these bounds is affected by the proportion of the voting population which always votes for its preferred candidate. **ARE WE GOING TO DO THIS???** In addition, we investigate how the existence of sincere voters can resolve the problem of existence of a Nash equilibrium of the voting game.

Whereas the literature on mechanism design typically takes the view that a well-designed social choice mechanism should ideally implement the outcome that is selected when individuals report their preferences truthfully, the approach taken in this paper points to an alternative metric to use in comparing alternative voting mechanisms, namely, the proportion of the population which benefits from manipulation of the mechanism. There is good reason for wanting to design mechanisms which encourage truthful reporting. As noted by ?, the manipulation of social decision procedures is a matter of concern if these procedures select an efficient outcome when voters are truthful, but recommend an inefficient alternative when voters report their preferences insincerely; much effort has consequently been directed to designing strategy-proof mechanisms for collective choice. However, ever since the seminal contributions of ? and ?, it has been well understood that the voting mechanisms used in actual elections (and, in particular, the plurality or first-past-the-post mechanism, which is by far the most widely used mechanism) are vulnerable to manipulation **SHOULD WE ALSO MENTION ARROW?** Given this fact, our analysis suggests that it may be illuminating to compare alternative voting procedures with respect to the likelihood that a manipulation is to the benefit of a majority of the voting population, or will lead to a different outcome on the Pareto frontier rather than to a Pareto inefficient decision. In some sense, our analysis can be viewed as a cautionary tale against being overly focussed on designing collective choice procedures which always select the outcome which prevails when electors vote sincerely, because this may not be an outcome that is particularly worth protecting. In effect, the plurality rule becomes vulnerable to manipulation when it selects an outcome that is the preferred outcome of only a minority of voters, but it is precisely in such circumstances that manipulation leads to an outcome which is Pareto non-comparable to that selected under sincere voting. A voting procedure with respect to which manipulations typically benefit a large proportion of the voting population might therefore be seen as more desirable than those for which manipulations are typically only advantage a narrow minority of citizens.

The structure of this paper is as follows. In section 2, below, we first lay out

our model, which extends the standard model of voting under the plurality rule to include both strategic and invariably sincere voters. In section 3 we study the set of outcomes which satisfy both the standard Gibbard-Satterthwaite definition of manipulation and are Nash equilibria of the voting game, and calculate exact minimum and maximum bounds on the number of voters who benefit from strategic voting as a function of the number of voters, the number of candidates, and the number of invariably sincere voters. Subsequently, we disaggregate the overall effect and calculate exact bounds for both the sincere and on the strategic voting populations, which provides us with additional insight into how these gains are shared. In section 4 we propose an alternative definition of manipulation, and study the bounds associated with this approach. Section 5 concludes.

2 The Model

We consider a set-up that is, in most essentials, identical to classic political economy models: a set of voters ranks a set of candidates, and the winning candidate is selected using the plurality rule with alphabetical tie-break. We denote the set of candidates by $A = \{a_1, \dots, a_m\}$, letting m denote the cardinality of this set; the set of voters is denoted by $N = \{1, \dots, n\}$, and has cardinality n . Unlike more traditional models, however, we distinguish between two types of voters: those who are willing to vote strategically if they anticipate that this will lead to an outcome which they prefer to that which would prevail if they were to vote for their preferred candidate, and those who always vote sincerely. Denote, therefore, the set of (potentially) strategic voters by $S \subseteq N$, and the cardinality of this set by s ; the set of sincere voters is consequently $N \setminus S$, and has cardinality $n - s$. We assume that Nature determines whether or not a voter is tactical or invariably sincere; note that whether or not a tactical voter chooses, in equilibrium, to vote for a candidate who is not their preferred leader depends upon the particular electoral environment.

The true preference profile of the voting population - that is, each voter's ranking of each of candidates in declining order of actual preference - is denoted by R^N ; the true preference ranking of the candidates by any individual voter i is denoted R^i . The set of all possible preference profiles is \mathcal{R}^N . As strategic voters may choose to report a ranking that does not correspond to their true preference profile, we denote the preference profile which is constructed from the actual reports of each voter by Q^N . As noted above, the winner of a given election is determined by application of the plurality rule with alphabetical tie-break. That is, denoting by $F(a_i, Q^N)$ the set of voters who rank $a_i \in A$ as their preferred candidate when the reported ranking is Q^N , then for any $a_i, a_j \in A$, if $F(a_i, Q^N) = F(a_j, Q^N)$ and $F(a_k, Q^N) < F(a_i, Q^N)$

for any $k \neq i, j$, then candidate i is selected as the victor if $i < j$. Let $Pl(Q^N)$ denote the outcome selected by the plurality voting process when the reported preference profile is Q^N .

Below, we study Nash equilibrium outcomes of the voting game. As the set of equilibrium outcomes typically depends on the set of strategic voters, we refer to a Nash equilibrium outcome when the set of strategic voters is S as a S -strong Nash equilibrium. In the tradition of Gibbard (1973) and Satterthwaite (1975), an S -effective manipulation of R^N is a profile Q^N such that (i) all players $i \in S$ who misreport their preferences prefer the outcome under Q^N to the outcome under R^N , that is, $Pl(Q^N) \succ_i Pl(R^N)$ for all i such that $Q^i \neq R^i$, and (ii) all other players report their preferences truthfully.² We denote by $N(S | R^N)$ the set of S -Nash equilibrium profiles that are S -effective manipulations of R^N . Define $N(R^N) = \cup_S N(S | R^N)$, that is, $N(R^N)$ is the set of outcomes which are S -Nash equilibrium outcomes for at least some possible S . Note that although the set of Nash equilibria of the voting game is always non-empty - in particular, any combination of strategies which involves more than the minimum plurality of voters voting for the same candidate is a Nash equilibrium, as no individual voter is pivotal - but the intersection between the set of Nash equilibria and manipulations of the sincere voting outcome which satisfy the Gibbard-Satterthwaite manipulation criterion may be empty, i.e., if it is not a best-response for strategic voters who do not benefit from the manipulation to continue to report their true preferences given that other strategic voters have chosen to misreport. In section ??, below, we study an example of a voting environment in which this problem arises.

It will not typically be the case that all strategic voters will choose to manipulate the sincere voting outcome, and so it is also useful to identify preference profiles which are vulnerable to manipulation. A preference profile $R^N \in \mathcal{L}^N$ is unstable if at least one voter has an incentive to manipulate their report - that is, R^N is unstable if there is at least one player i for whom it is not beneficial to submit a truthful report given that all other players are truthful. Notice that whether or not a given preference profile is stable will typically depend on the set of strategic voters (and, of course, on its complement - the set of invariably sincere voters). Define by $P(S)$ the set of S -unstable profiles of preferences, and let $P = \cup_S P(S)$, that is, the set

²Notice that the Gibbard-Satterthwaite definition of manipulation does not guarantee that all strategic voters are reporting a preference ordering which is a best-response to the preference-orderings reported by the other voters. In particular, strategic voters who are made worse off as a result of a manipulation by other strategic voters are required to report their preferences honestly, even if they would be able to obtain a better outcome by in turn manipulating (in the sense of Gibbard-Satterthwaite) Q^N . The results presented in this paper are for S -Nash equilibria which also satisfy the Gibbard-Satterthwaite definition of manipulation.

of preference profiles that is S -unstable for at least some possible set S . Similarly, let $Z(S) \subset P(S)$ denote the subset of S -unstable profiles of preferences that admit at least one S -effective manipulation that is a S -strong Nash equilibrium, and let $Z = \cup_S Z(S)$, that is, Z is the set of S -unstable preference profiles admitting an S -effective manipulation which is a S -strong Nash equilibrium for at least some possible set S .

Finally, in the event that (some) strategic voters choose to misrepresent their preferences at the strong Nash equilibrium, we need to be able to distinguish the winners and losers as a result of this manipulation. Given $R^N \in N(S)$ and $Q^N \in N(R^N)$, $E(R^N, Q^N)$ is the set of voters - including those who vote sincerely - who benefit from the manipulation from R^N to Q^N .

3 Bounding The Proportion of Winners and Losers In the Population

Our principal objective in this paper is to derive maximum and minimum bounds for the proportion of citizens who benefit - or are adversely affected by - manipulation of the reported preference profile. In this section we derive bounds for the population as a whole; in the next sections, we derive bounds for the populations of strategic and non-strategic voters. Our first step is to relate the set of voters who rank the successful candidate, a_j , first when the reported preference profile is Q^N and the set of voters - including those who always vote sincerely - who benefit from the manipulation Q^N , that is, the set $E(R^N, Q^N)$. This relationship, together with the fact that under the plurality rule the floor on the number of voters who support the winning candidate can be determined as the number of voters divided by the number of candidates, helps us to start to calculate the number of winners and losers as a result of a manipulation.

Lemma 1 *Let $R^N \in P$ and Q^N be an effective manipulation of R^N and suppose that $Pl(R^N) = a_l$ and $Pl(Q^N) = a_j$. Then $F(a_j, Q^N) \subseteq E(R^N, Q^N)$, and therefore $|E(R^N, Q^N)| > \frac{n}{m}$.*

Proof The first inclusion is an immediate consequence of the fact that Q^N is an effective manipulation of R^N ; the inequality follows as a result of the plurality rule. ■

Before stating our first substantive result we need to introduce some additional notation. We denote by $m^*(m, n, s)$ the minimum proportion of the population population - that is, including both strategic and invariably-sincere voters - which benefits

from strategic voting; $M^*(m, n, s)$ denotes the maximum proportion of the population which benefits from this behaviour. It is also useful to decompose the beneficial (or adverse) impact of strategic voting on the populations of strategic and invariably sincere voters. Thus we denote by $m_1^*(m, n, s)$ and $M_1^*(m, n, s)$ the minimum and maximum proportions of the strategic voting population which benefits from manipulation of the sincere voting outcome. Similarly, we denote by $m_2^*(m, n, s)$ and $M_2^*(m, n, s)$ the minimum and maximum proportions of the sincere voting population which is made better off when strategic voters cast their ballots tactically.

Proposition ??, below, establishes an exact upper bound on the proportion of the overall voting population which can benefit from an effective manipulation. Not surprisingly, an effective manipulation cannot make everyone better off. However, in large populations, with large numbers of candidates, the proportion which benefits can be very high.

Proposition 2 *In equilibrium, the maximum proportion of voters who benefit from an effective manipulation, $M^*(m, n, s)$, is equal to $1 - \frac{1}{n} \lceil \frac{n}{m} \rceil$.*

Before providing a proof of this proposition, it is useful to establish some intermediate lemmata. The overall strategy of the proof is to first construct a (unstable) preference profile for which there exists an effective manipulation which is also a strong Nash equilibrium. By calculating the proportion of voters who benefit from this manipulation, it follows that the upper bound on the number of voters benefitting from an effective manipulation can be no less than this number. Then, in the proof of the proposition, we calculate the minimum number of voters that are adversely affected by any effective manipulation. Since the sum of these numbers is equal to one, it follows that the upper bound which is calculated in proving the lemma must be exact.

Lemma 3 *For any number of voters, n , and any number of candidates, m , it is possible to construct a preference profile $R^N \in Z$ and an manipulation Q^N which belongs to the set of Nash equilibria $N(R^N)$ and such that $|E(R^N, Q^N)| = n - \lceil \frac{n}{m} \rceil$ and $Q^{-2} = R^{-2}$, that is, where the effective manipulation coincides with the true preference profile except for the report of player 2.*

Proof For any m, n , let q be the largest non-negative integer such that $qm \leq n$ and then choose $r \in \{0, 1, \dots, m-1\}$ so that $n = qm + r$. We now construct a R^N and Q^N which satisfy the statement of the lemma; note that the R^N, Q^N which are constructed for this purpose depend upon the particular values of q and r .

Case 1 : $r = 0$. Consider a partition $\{N_1, N_2, \dots, N_m\}$ of N in m subsets such that voter 2 belongs to N_2 . Notice that, by construction, each of these subsets is comprised of the same number of voters, that is, for all $k \in \{1, 2, \dots, m\}$, $|N_k| = q$. Let R^N be a profile such that

$$\forall i \in N_3, R^i = a_3 \text{ and } \forall i \in N_k, R^i = a_k a_3 \text{ for } k \neq 3$$

that is, the true preference profile is such that for all voters belonging to N_3 , candidate a_3 is ranked first (and all other candidates can be ranked in any order) whereas for all other players in partition N_k , candidate a_k is ranked first, candidate a_3 is ranked second (and all other candidates can be ranked in any order). We next construct a manipulation, which is also a Nash equilibrium. Thus, consider the (mis)report for player 2, $Q^2 = a_3 a_2 \dots$ so that $Q^N = (Q^2, R^{-2})$. Observe that $Pl(R^N) = a_1$, whereas $Pl(Q^N) = a_3$. Since there are now $q + 1$ voters who rank a_3 as their preferred candidate, $Pl(Q^N) = a_3$. Moreover, since voter 2 (and, indeed, all voters except those in partition N_1) prefers a_3 to a_1 , this misreport satisfies the requirement of a manipulation - that is, any players mis-representing their preferences prefer the outcome under Q^N to the outcome under R^N . Moreover, Q^N is also a Nash equilibrium voting profile - no voter who does not belong to N_3 can change the outcome of the voting process by mis-reporting their true preferences profile given the reports of the other voters, no voter in N_3 wishes to misreport, and there is no alternative strategy which improves the payoff of voter 2.

Case 2 : $q = 0$. Note that this implies that $n < m$. Define the profile R^N as follows:

$$R^1 = a_2 a_1 a_3 a_4 \dots a_N \tag{1}$$

$$\text{and } R^i = a_{i+1} a_1 a_2 \dots a_i a_{i+2} \dots a_N \text{ for all } i \neq 1. \tag{2}$$

Observe that with this preference profile there are n candidates who receive one vote when all voters report their preference profile sincerely, $m - n$ candidates who receive no votes, and $Pl(R^N) = a_2$. Now let $Q^2 = a_1$ and consider $Q^N = (Q^2, R^{-2})$. As when all voters cast their ballots sincerely, no candidate receives more than one vote. However, $Pl(Q^N) = a_1$. Notice that, with the exception of voter 1, candidate a_1 is preferred by all voters to candidate a_2 . As voter 2 prefers a_1 to a_2 , and voter 2 is the only agent who is not voting sincerely, Q^N meets the requirements of a manipulation. Moreover, Q^N is a Nash equilibrium: given that a_1 is the second-most-preferred outcome for all voters - including voter 1 - none of these other voters can obtain an outcome they prefer to the outcome a_1 by misreporting their true preference profile, and therefore R^i is a best-response for player i to (Q^2, R^{-i}) for all $i \neq 2$ and Q^2 is a best response for player 2 to R^{-2} .

Case 3 : $r \geq 1$ and $q \geq 1$. We now allocate voters into blocks of either q or $q + 1$ voters. Consider a partition $\{N_1, N_2, \dots, N_m\}$ of N into m subsets such that voter 2 belongs to N_3 , $|N_k| = q + 1$ if $k \in \{2, 3, \dots, r + 1\}$ and $|N_k| = q$ if $k \in \{1, r + 2, r + 3, \dots, m\}$. Let R^N be the profile defined by

$$\forall i \in N_1, R^i = a_1 a_2 \dots a_m \quad (3)$$

$$\forall i \in N_k, R^i = a_k a_1 a_2 \dots a_{k-1} a_{k+1} a_m \text{ for all } k \neq 1. \quad (4)$$

that is, the true preference profile is such that all voters in N_1 rank a_1 first (and all other candidates in alphabetical order), whereas all other voters in N_k rank candidate a_k first, and candidate a_1 second (and all other candidates in alphabetical order). Observe that there are r partitions with $q + 1$ voters, and that one of these partitions is N_2 , so that $Pl(R^N) = a_2$. Now let $Q^2 = a_1$ and $Q^N = (Q^2, R^{-2})$. Notice that $Pl(Q^N) = a_1$, and that a_1 is preferred by all voters, except for those in N_2 , to the outcome under R^N . Q^N therefore meets the requirements of a manipulation, as only voter 2 misreports, and voter 2 prefers the outcome under Q^N to the outcome under R^N . Moreover, Q^N is a Nash equilibrium: for all voters in partitions $N_k, k \neq 2$, a_1 is preferred to any outcome other than a_k but by misreporting their preference they cannot obtain a_k and so voting a_k is a best response to Q^N ; voters in N_2 cannot misreport and obtain a_2 , and a_1 is voter 2's best response to R^{-2} . This completes Case 3.

To complete the proof of the lemma, it suffices to observe that, in each case, the only voters who do not benefit from the effective manipulation are those q voters (or, in case 2, voter 1) whose preferred candidate is selected under sincere voting. Consequently $|E(R^N, Q^N)| = n - \lceil \frac{n}{m} \rceil$. ■

It is now straightforward to extend this result to a setting in which some proportion of voters are invariably sincere.

Lemma 4 *For any number of voters, n , of candidates, m , and for any $S \subseteq N$, it is possible to construct a preference profile $R^N \in Z(S)$ and an S -effective manipulation Q^N such that $Q^N \in N(S | R^N)$ and such that $|E(R^N, Q^N)| = n - \lceil \frac{n}{m} \rceil$ and $Q^{-2} = R^{-2}$, that is, where the S -effective manipulation coincides with the true preference profile except for the report of player 2.*

Proof Take R^N and Q^N from Lemma ?? and rename the voters so that voter 2 belongs to S . ■

We can now prove our first Proposition, which follows straightforwardly from our preceding lemmata.

Proof of Proposition ??.

Renaming the voters such that $2 \in S$, it follows from Lemma ?? that $M^*(m, n, s) \geq 1 - \frac{1}{n} \lfloor \frac{n}{m} \rfloor$. In addition, if $Pl(R^N) = a_j$, then $|E(a_j, R^N)| \geq \lfloor \frac{n}{m} \rfloor$. This implies that at least $\lfloor \frac{n}{m} \rfloor$ voters suffer from manipulation if it happens. So $n * M^*(m, n, s) \leq n - \lfloor \frac{n}{m} \rfloor$. ■

Proposition ?? provides an exact upper bound on the number of individuals who may, in equilibrium, benefit from an S -effective manipulation; in particular, when the number of candidates is larger than the number of voters, this may be almost the entirety of the population. Such a scenario might arise, for example, if a small group of voters is tasked with selecting a successful applicant from a large pool of job seekers. Strikingly, even when there are only two candidates, this upper bound never falls below one half of the voting population. This result may nonetheless appear to be of limited interest, because in any specific setting there is no reason to expect that the actual preference profile will be similar to the particular preference profile used to construct the proof. Rather, what may seem more pertinent is clearer insight into the minimum proportion of the population which benefits from the manipulation of the sincere voting outcome. This is the focus of our next Proposition.

Proposition 5 *The minimum proportion of the population which benefits from an effective manipulation, $m^*(m, n, s)$, is equal to $\begin{cases} \frac{1}{n} + \frac{1}{n} \lfloor \frac{n}{m} \rfloor & \text{if } s \leq 2 \lfloor \frac{n}{m} \rfloor + 1 \\ \frac{1}{n} \lfloor \frac{s}{2} \rfloor & \text{if } s > 2 \lfloor \frac{n}{m} \rfloor + 1 \end{cases}$*

As a first step towards establishing this result, we first establish a minor lemma.

Lemma 6 *If the number of strategic voters s is greater than $2 \lfloor \frac{n}{m} \rfloor + 1$, then $\frac{s}{2} > \frac{n-s}{m-2}$.*

Proof Assuming that $s \geq 2 \lfloor \frac{n}{m} \rfloor + 2$ we have that $\frac{s}{2} > \frac{n}{m}$. Cross-multiplying and subtracting $2s$ from both sides, we observe that this is equivalent to $\frac{n}{m} > \frac{n-s}{m-2}$. ■

It is now possible to proceed to calculate the lower bound on the number of voters who benefit from an S -effective manipulation.

Proof of Proposition ??. Assume that voter 2 belongs to the set of strategic voters, S . We proceed in 2 cases:

Case 1: $s \leq 2 \lfloor \frac{n}{m} \rfloor + 1$

Let $R^N \in N(S)$, and let Q^N be an S -effective manipulation of R^N . Suppose that $Pl(R^N) = a_i$ and $Pl(Q^N) = a_j$. From Proposition ??, we have that $|E(a_j, Q^N)| > \frac{n}{m}$. From Lemma ??, we have that $|E(R^N, Q^N)| \geq 1 + \lfloor \frac{n}{m} \rfloor$. Consequently, $m^*(m, n, s) \geq \frac{1}{n} + \frac{1}{n} \lfloor \frac{n}{m} \rfloor$. We now show that there exist possible preference profiles for which

$m^*(m, n, s) \leq \frac{1}{n} + \frac{1}{n} \lfloor \frac{n}{m} \rfloor$. As above, let $n = qm + r$ where q is the largest non-negative integer such that $qm \leq n$ and then choose r so that $n = qm + r$.

Case 1.1 Let $r = 0$, and consider a partition $\{N_1, N_2, \dots, N_m\}$ of N into m subsets with equal numbers of voters, i.e., such that $|N_j| = q$ for all $j = 1, \dots, m$. Without loss of generality, assume that voter 2 $\in N_3$ and, moreover, that the set of strategic voters, less voter 2, are all members of N_2 if there are no more than $q + 1$ strategic voters, that is, $S \setminus \{2\} \subset N_2$ if $|S| \leq q + 1$, or alternatively they are all members of the first two subsets, that is, $N_2 \subset S \setminus \{2\}$ and $S \setminus \{2\} \subset N_2 \cup N_1$ if $q + 1 < |S| \leq 2q + 1$. Now consider a possible true preference profile R^N such that voters in N_1 rank a_1 first and all other candidates in any order; voter 2 ranks a_3 first, a_2 second, and all other candidates in any order; and voters in $N_k \setminus \{2\}$, $k \neq 1$, rank a_k first, a_1 second, and all other candidates in any order. Notice that each alternative is top ranked by exactly q voters, so $Pl(R^N) = a_1$. Now suppose that voter 2 were to manipulate the vote by reporting $Q^2 = a_2$ first, a_3 second, and all other candidates in any order. There are now $q + 1$ voters ranking a_2 as the preferred candidate, and so $Pl(Q^2, R^{-2}) = a_2$. Since a_2 is preferred by voter 2 to a_1 , and voter 2 is the only agent who does not vote sincerely, this satisfies the requirement of a manipulation, and we have $R^N \in P(S)$ and $Q^N = (Q^2, R^{-2})$ is, by construction, an S -effective manipulation of R^N . Voters of $N_2 \cup \{2\}$ benefit from the manipulation. In contrast, voters of $S \setminus (N_2 \cup \{2\})$ rank a_1 first and a_2 second in the profile Q^N . However, none of these voters can make themselves strictly better off by changing their reported preference profile. Consequently, Q^N is a Nash equilibrium. We have $|E(R^N, Q^N)| = q + 1$ voters who benefit, and so the minimum proportion of the population which benefits is $m^*(n, m, s) \leq \frac{1}{n} + \frac{1}{n} \lfloor \frac{n}{m} \rfloor$.

Case 1.2 Now consider the case where $q = 0$. By construction, the set of strategic voters is therefore a singleton, that is, $S = \{2\}$ and $n < m$. Now consider a possible true preference profile R^N such that voter 1 ranks a_2 first and all other candidates in any order, voter 2 ranks a_{n+1} first, a_1 second, and all other candidates in any order, and all voters $i \notin \{1, 2\}$ rank a_i first, a_2 second, and all other candidates in any order. If all voters report their true preferences, then $Pl(R^N) = a_2$. Now suppose that the unique strategic voter, voter 2, manipulates the outcome by reporting a_1 as the most-preferred candidate, and a_{n+1} as the second-most preferred candidate. Then $Pl(Q^N) = a_1$. Observe that since voter 2 is the unique strategic voter, then it is trivially true that Q^N is an S -effective manipulation of R^N and a S -Nash equilibrium. Consequently, $|E(R^N, Q^N)| = 1$ and therefore the proportion of the population when benefits from the manipulation is $m^*(m, n, s) \leq \frac{1}{n} = \frac{1}{n} + \frac{1}{n} \lfloor \frac{n}{m} \rfloor$ since $n < m$.

Case 1.3 Finally, suppose that $r \geq 1$ and $q \geq 1$. Consider a partition $\{N_1, N_2, \dots, N_m\}$ of N in m subsets such that voter 2 belongs to N_3 , and such that there are $q + 1$ voters in sets N_2 to N_{r+1} , and q voters in all of the other sets in this partition. Additionally, assume that all of the strategic voters, with the exception of voter 2, are assigned to N_1 if there are no more than $q + 1$ members of S , or are assigned to N_1 and N_2 if $q + 1 < |S| \leq 2q + 1$. Now consider a possible true preference profile R^N such that voters in N_1 rank a_1 first and all other candidates in any order; voters in N_2 rank a_2 first and all other candidates in any order; voter 2 ranks a_3 first and a_1 second and all other candidates in any order, while voters in $N_k \setminus \{2\}, k \notin \{1, 2\}$ rank a_k first, a_2 second, and all other candidates in any order. Then a_1 and a_2 are top ranked by exactly q voters and $q + 1$ voters respectively. Since $|N_k| \leq q + 1$ for all $k \in \{1, 2, \dots, m\}$, $Pl(R^N) = a_2$. Next, notice that the only voters who benefit from a switch to a_1 rather than a_2 when the true preference profile is R^N are the members of $N_1 \cup \{2\}$. Now suppose that voter 2 chooses to strategically misreport, and announces $Q^2 = a_1 a_3 \dots$ so that we have $Q^N = (Q^2, R^{-2})$. The candidate chosen by the plurality rule is now $Pl(Q^N) = a_1$, and Q^N satisfies the definition of an S -effective manipulation of R^N . Moreover, Q^N is also a S -Nash equilibrium since no strategic voter can improve their payoff by changing their reported preference. We have $|E(R^N, Q^N)| = q + 1$ voters who benefit, and so the minimum proportion of the population which benefits is $m^*(n, m, s) \leq \frac{1}{n} + \frac{q}{n} = \frac{1}{n} + \frac{1}{n} \lfloor \frac{n}{m} \rfloor$. This establishes our claim for the case where $s \leq 2 \lfloor \frac{n}{m} \rfloor + 1$, that is, when there are relatively few strategic voters.

Case 2: $s > 2 \lfloor \frac{n}{m} \rfloor + 1$

THIS PARAGRAPH MAKES NO SENSE As above, we first establish that the proportion of the population which benefits from the manipulation of the voting outcome is at least equal to $\frac{1}{n} \lfloor \frac{s}{2} \rfloor$. Without loss of generality, consider an unstable profile $R^N \in Z$ and let Q^N be an S -effective manipulation of R^N that is a S -Nash equilibrium. The proof is by contradiction. Without loss of generality, let $Pl(R^N) = a_1$, $Pl(Q^N) = a_2$ and suppose that the number of voters who prefer outcome a_2 to outcome a_1 under the true preference profile R^N is less than $\lfloor \frac{s}{2} \rfloor$. Denote by $L(a_2, a_1, R^N)$ the set of participants who prefer a_1 to a_2 when the true preference profile is R^N . Then it must be true that $|L(a_2, a_1, R^N) \cap S| \geq \lfloor \frac{s}{2} \rfloor$ and $|L(a_1, a_2, R^N)| = |L(a_1, a_2, Q^N)| \leq |L(a_2, a_1, Q^N)| \leq |L(a_1, a_2, R^N)|$. So $(L(a_2, a_1, R^N) \cap S) \setminus L(a_1, a_2, R^N) \neq \emptyset$. That is there exist at least one voter i of S who prefers candidate a_1 to a_2 and who did not vote for a_1 in R^N and thus in Q^N . Let T^i be a strategic preference of voter i in which he ranks a_1 first and $T^N = (T^i, T^{-i})$. From R^N to T^N , both candidates a_1 and a_2 receive one additional vote while others do not receive any. Since candidate a_1 is elected at R^N , he will still be elected at

T^N . A contradiction arises since Q^N is an S -Nash equilibrium. We conclude that $|L(a_1, b_1, R^N)| \geq \lceil \frac{s}{2} \rceil$. That is $m^*(n, m, s) \geq \frac{1}{n} \lceil \frac{s}{2} \rceil$.

We now construct a true preference profile $R^N \in Z(S)$ and an effective manipulation $Q^N \in N(R^N)$ such that Q^N is a S -Nash equilibrium and $|E(R^N, Q^N)| = \lceil \frac{s}{2} \rceil$. Without loss of generality, choose k, r so that $s = 2k + r$ with $r \in \{0, -1\}$ and $n - s = (m - 2)q + p$ with $p \in \{0, \dots, m - 3\}$. By construction, from Lemma ??, we have $k \geq q + 1$. Now construct a partition $\{N_1, N_2, \dots, N_m\}$ of N such that $|N_1| = k - 1$, $|N_2| = k$, and $|N_j| = q + 1$ for $j = 3, \dots, p + 3 + r$. Moreover, suppose that voter 2 $\in N_3$ and let $S \subset N_1 \cup N_2 \cup \{2\}$.³ Now consider the possible true preference profile R^N such that

$$R^i = a_1 a_2 a_3 \dots \text{ for all } i \in N_1 \tag{5}$$

$$R^i = a_2 a_1 \dots \text{ for all } i \in N_2; \text{ and} \tag{6}$$

$$R^i = a_l a_2 a_1 \dots \text{ for all } i \in N_j \setminus \{2\}, j \geq 3. \tag{7}$$

By construction, we have $Pl(R^N) = a_2$. However, the voters belonging to $N_1 \cup \{2\}$ would prefer that candidate a_1 prevail. Consider, now, a possible misreport by voter 2, such that $Q^2 = a_1 a_3 a_2 \dots$ and let $Q^N = (Q^2, R^{-2})$. Observe that $Pl(Q^N) = a_1$, and since the only voters who misrepresent their preferences under Q^N benefit from the change in the voting outcome, Q^N is an S -effective manipulation of R^N . In addition, Q^N is an S -Nash equilibrium: with the exception of voter 2, all of the strategic voters are voting for their preferred outcome, and cannot improve their payoff by changing their vote. It follows that $|E(R^N, Q^N)| = k = \lceil \frac{s}{2} \rceil$. That is $m^*(n, m, s) \leq \frac{1}{n} \lceil \frac{s}{2} \rceil$. We have therefore shown that $\frac{1}{n} \lceil \frac{s}{2} \rceil \leq m^*(n, m, s) \leq \frac{1}{n} \lceil \frac{s}{2} \rceil$ which completes our claim. . ■

Notice that the bounds derived here are sensitive to three key parameters of the voting problem, namely, the number of candidates, the number of strategic voters, and the total number of voters. In particular, the minimum bound is increasing in the proportion of strategic voters ; if all voters are strategic voters, then the minimum proportion of the voting population that benefits from an effective manipulation is equal to 50%. In contrast, when strategic voters are scarce, and the pool of candidates is large in size relative to the number of voters, then the proportion which actually benefits from a manipulation is approximately equal to the minimal winning coalition size, n/m . In some sense, then, what this result makes clear is that the likelihood that strategic voting is of broad benefit is intrinsically linked to the willingness of

³Note that N_1 or N_2 may contain one voter not in S . This is for the case $s = 2k - 1$.

the general voting population to vote tactically. When the vast majority of electors invariably vote sincerely, then it is much more likely that manipulation of the outcome by strategic voters may benefit only a small minority of the population - and may adversely affect all other agents. In contrast, when most voters cast their ballot with a view to potentially influencing the outcome of the election, then the floor on the proportion of the population that benefits from manipulation rises rapidly. **I AM NOT HAPPY WITH THIS PARAGRAPH.**

Whereas the results derived above provide bounds on overall gains, they do not draw a clear picture of how those gains are shared between strategic and non-strategic voters. This is the focus of the next section.

3.1 Maximal and Minimal Gains For Sincere Voters

The benefits of manipulation of the sincere voting outcome will typically be shared between both sincere and strategic voters, and although there must be at least some proportion of the strategic voting population which benefits - for otherwise a strategic voter has no incentive to manipulate the sincere voting outcome - there is no reason *a priori* to expect that these gains will accrue primarily to either the invariably sincere or to the strategic voters. For this reason, it is useful to derive exact bounds for the maximum and minimum proportion of sincere voters who benefit from a manipulation of the voting outcome. Not surprisingly, in view of Proposition ??, the maximum proportion of sincere voters who benefit from manipulation may attain 100%; what is somewhat less obvious, however, is that the possibility of manipulation being to the benefit of *all* invariably sincere voters arises when the strategic voting population is in fact a relatively small proportion of the overall voting population - approximately equal to n/m - as is shown in Proposition ??. As before, the proof is by construction.

Proposition 7 *Suppose that the proportion of strategic voters is strictly less than 1. Then the maximum proportion of sincere voters who benefit from a manipulation is equal to*

$$M_1^*(m, n, s) = \begin{cases} \frac{1}{n-s}(n-s) & \text{if } s-1 > \lceil \frac{n}{m} \rceil \\ \frac{1}{n-s}(n - \lceil \frac{n}{m} \rceil - 1) & \text{if } s-1 \leq \lceil \frac{n}{m} \rceil \end{cases}$$

Proof We first consider the case when the number of strategic voters is one more than the minimum winning coalition size, i.e., $s-1 > \lceil \frac{n}{m} \rceil$. Consider the true preference profile R^N and the manipulation Q^N used earlier in the proof of Lemma ??. Now rename the voters so that $N \setminus E(R^N, Q^N) \subset S \setminus \{2\}$. We then have that $Q^N \in N(R^N) \subset N(S \mid R^N)$ and $N \setminus S \subset E(R^N, Q^N)$. This means that $M_1^*(m, n, s) \geq \frac{1}{n-s}(n-s)$. That is $M_1^*(m, n, s) = 1$: all sincere voters benefit from the manipulation.

The second case arises when the number of strategic voters is less than or equal to the minimum winning coalition size, i.e., $s - 1 \leq \lfloor \frac{n}{m} \rfloor$. From Proposition ??, we know that at most $n - \lfloor \frac{n}{m} \rfloor$ voters benefit from a manipulation. Since only a voter belonging to S may choose to vote insincerely, at most $n - \lfloor \frac{n}{m} \rfloor - 1$ voters of $N \setminus S$ can benefit from a manipulation. That is $M_1^*(m, n, s) \leq \frac{1}{n-s}(n - \lfloor \frac{n}{m} \rfloor - 1)$. To complete the proof, consider once again the true preference profile R^N and the manipulation Q^N used to establish Lemma ?. Rename the voters so that $S \setminus \{2\} \subset N \setminus E(R^N, Q^N)$. We have $Q^N \in N(R^N) \subset N(S \mid R^N)$ and $E(R^N, Q^N) \subset (N \setminus S) \cup \{2\}$. Since $|E(R^N, Q^N)| = n - \lfloor \frac{n}{m} \rfloor$, we have $M_1^*(m, n, s) \geq \frac{1}{n-s}(n - \lfloor \frac{n}{m} \rfloor - 1)$. ■

Of course, whilst it is interesting to know that - at least in some circumstances - it is the entire population of sincere voters who benefit from a manipulation, there is no reason to believe, *a priori*, that such circumstances arise with any particular frequency. What is of greater importance is to get some sense of the difference between the maximum and minimum proportion of sincere voters who benefit from a manipulation. Proposition ??, below, provides an exact bound on the minimum proportion of strategic voters who benefit from a manipulation. Whereas Proposition ?? takes an optimistic view of the impact of a manipulation on sincere voters, Proposition ?? is pessimistic; it is derived under the assumption that the composition of the group of voters adversely affected by the manipulation is comprised primarily (and possibly exclusively) of sincere voters. What this means in practice is that when strategic voters comprise a large share of the overall voting population, then none of the sincere voters may end up benefiting from the manipulation; as the proportion of strategic voters in the overall population falls, then this makes it more likely that sincere voters end up in the group of voters which are positively impacted by the manipulation.

Proposition 8 $m_1^*(m, n, s) = \begin{cases} 0 & \text{if } s - 1 \geq \lfloor \frac{n}{m} \rfloor \\ \frac{1}{n-s}(1 + \lfloor \frac{n}{m} \rfloor - s) & \text{if } s - 1 < \lfloor \frac{n}{m} \rfloor \end{cases}$

Proof This result follows directly from Proposition ?. ■

This completes our derivation of the maximal and minimal impact of manipulation on sincere voters. What is the take-away message? On the one hand, it is clear that when the proportion of strategic voters in the population is high enough, then the proportion of the sincere voting population which benefits from a manipulation can vary between 0 and 100%, which is not a particularly tight bound; on the other hand, when the proportion of strategic voters falls, then this gap narrows rapidly. Moreover, as the number of candidates falls, then the difference between the maximum and minimum bounds on the proportion of sincere voters who benefit from a

manipulation also narrows. Overall, this suggests than if there is reason to believe that a high proportion of the voting population habitually votes sincerely, and there are relatively few candidates relative to the number of voters, then any manipulation of the voting outcome can be expected to positively impact a significant proportion of the population of invariably sincere voters.

3.2 Maximal and Minimal Gains For Strategic Voters

To complete our analysis, we derive exact bounds for the maximum and minimum proportion of strategic voters who benefit from manipulation of the sincere voting outcome. These results largely echo those established above: we show that in some settings, manipulation may in fact benefit all strategic voters, and that the minimum proportion of the strategic voting population which profits from a manipulation is bounded strictly away from zero.

Proposition 9 $M_2^*(m, n, s) = \begin{cases} 1 & \text{if } s \leq n - \lceil \frac{n}{m} \rceil \\ \frac{1}{s}(n - \lceil \frac{n}{m} \rceil) & \text{if } s > n - \lceil \frac{n}{m} \rceil \end{cases}$

Proof Follows obviously from Proposition ?? . ■

The derivation of the minimum bound, below, is somewhat more interesting. When the proportion of strategic voters in the general population is less than one half of the total population, then it may be the case that only one strategic voter benefits from manipulation, i.e., this is precisely the setting where almost all of the benefits of manipulation accrue to the sincere voters. In contrast, when strategic voters constitute a majority of the voting population, then the floor on the proportion of the strategic voting population which benefits from manipulation grows.

Proposition 10 $m_2^*(m, n, s) = \begin{cases} \frac{1}{s} & \text{if } s - 1 \leq \lfloor \frac{n}{2} \rfloor \\ \frac{1}{s}(s - \lfloor \frac{n}{2} \rfloor) & \text{if } s - 1 > \lfloor \frac{n}{2} \rfloor \end{cases}$

Proof As above, the proof is by construction. Assume that voter 2 $\in S$ and consider a partition $\{N_1, N_2, \{2\}\}$ of N , and a true preference profile $R^N \in L^N$ and a preference $T^2 \in L$ such that:

If $n = 2k$. We have $|N_1| + 1 = |N_2| = k$, $R^i = a_1 a_2 \dots$ for all $i \in N_1$, $R^i = a_2 a_1 \dots$ for all $i \in N_2$, $R^2 = a_3 a_1 a_2 \dots$ and $T^2 = a_1 a_3 a_2 \dots$

If $n = 2k + 1$. We have $|N_1| = |N_2| = k$, $R^i = a_2 a_1 \dots$ for all $i \in N_1$, $R^i = a_1 a_2 \dots$ for all $i \in N_2$, $R^2 = a_3 a_2 a_1 \dots$ and $T^2 = a_2 a_3 a_1 \dots$

Pose $T^N = (T^2, R^{-2})$. It is easy to check that T^N is an effective manipulation of R^N , a Nash equilibrium given R^N and that $E(R^N, T^N) = N_1 \cup \{2\}$. From lemma ??,

we have $T^N \in N(S \mid R^N)$. Now rename voters of $N \setminus \{2\}$ such that $S \setminus \{2\} \subseteq N_2$ if $s-1 \leq k$ and $N_2 \subseteq S \setminus \{2\}$ otherwise. It follows that $|E(R^N, T^N) \cap S| = |N_1 \cap S| + 1$.

Since $|N_1 \cap S| = \begin{cases} 0 & \text{if } s-1 \leq k \\ s - \lfloor \frac{n}{2} \rfloor - 1 & \text{otherwise} \end{cases}$, we have $|E(R^N, T^N) \cap S| = \begin{cases} 1 & \text{if } s-1 \leq k \\ s - \lfloor \frac{n}{2} \rfloor & \text{otherwise} \end{cases}$.

We deduce that $m_2^*(m, n, s) \leq \begin{cases} \frac{1}{s} & \text{if } s-1 \leq \lfloor \frac{n}{2} \rfloor \\ \frac{1}{s}(s - \lfloor \frac{n}{2} \rfloor) & \text{if } s-1 > \lfloor \frac{n}{2} \rfloor \end{cases}$.

It is left to show that $m_2^*(m, n, s) \geq \begin{cases} \frac{1}{s} & \text{if } s-1 \leq \lfloor \frac{n}{2} \rfloor \\ \frac{1}{s}(s - \lfloor \frac{n}{2} \rfloor) & \text{if } s-1 > \lfloor \frac{n}{2} \rfloor \end{cases}$.

Obviously, since only members of S can manipulate, we have $m_2^*(m, n, s) \geq \frac{1}{s}$.

Now assume that $s-1 > \lfloor \frac{n}{2} \rfloor$ and that there exist $R^N \in Z(S)$ and $Q^N \in N(S \mid R^N)$ such that $|E(R^N, Q^N) \cap S| < s - \lfloor \frac{n}{2} \rfloor$. Let $S_1 = S \setminus E(R^N, Q^N)$, $Pl(R^N) = a$ and $Pl(Q^N) = b$. We have $S_1 = S \cap \bar{E}(b, a, R^N)$ and $|S_1| > \lfloor \frac{n}{2} \rfloor$. On the other hand, $|E(a, R^N)| = |E(a, Q^N)| \leq |E(b, Q^N)| \leq |E(a, b, R^N)| < s - \lfloor \frac{n}{2} \rfloor \leq \lceil \frac{n}{2} \rceil$. Consequently, $S_1 \setminus E(a, R^N) \neq \emptyset$. Let $i \in S_1 \setminus E(a, R^N)$, $T^i = ab\dots$ be a strategic preference of voter i and $T^N = (T^i, Q^{-i})$. From R^N to T^N , both candidates a and b receive one additional vote while others do not receive any. Since candidate a wins election at R^N , he will still be elected at T^N . A contradiction arises since Q^N is an S -equilibrium. We then conclude that $|E(R^N, Q^N) \cap S| \geq s - \lfloor \frac{n}{2} \rfloor$. That is $m_2^*(n, m, s) \geq \frac{1}{s}(s - \lfloor \frac{n}{2} \rfloor)$. ■

4 When Honesty Is Not The Best Policy: Nash Manipulation

The Gibbard-Satterthwaite definition of manipulation requires that all voters report their preferences truthfully unless they are made better off as a result of misreporting their true preferences than at the sincere voting outcome. A weakness in this definition, however, is that it may be a best-response for strategic voters to misreport their preferences in equilibrium, even if this does not lead to a higher payoff than at the sincere voting outcome, as long as such a mis-report prevents an outcome which they dislike even more intensely. This possibility is illustrated by the following example. Consider a population of seven voters, and suppose that $R^1 = R^2 = abcd$, $R^3 = R^4 = badc$, $R^5 = R^6 = cdab$ and $R^7 = dcba$. Observe that $Pl(R^N) = a$ and that there is a unique manipulation which satisfies the Gibbard-Satterthwaite manipulation criterion, and in which player 7 chooses the strategy $Q^7 = c$ and that $Pl(Q^7, R^{-7}) = c$; this would make three players better off, and four players worse off than at a . However, the reports (Q^7, R^{-7}) are not Nash equilibrium strategies, which means that the intersection of the set of Nash equilibria of the

voting game and the set of manipulations which satisfy the Gibbard-Satterthwaite criterion union R^N is in fact empty. In particular, (Q^7, R^{-7}) is not a Nash equilibrium because it is a best response of either player 3 or player 4 to change their reported preference ordering to $abdc$ — in which case, the selected outcome would be a — or for player 1 or 2 to change their reported preference ordering to $bacd$ — in which case the selected outcome would be b . Note that if outcome b is selected then there is (at least) one voter who is mis-reporting their preference ordering who is worse off than at the sincere voting equilibrium. However, the player who is mis-reporting is better off than would be the case if c were selected. In contrast, if outcome a is selected, then there is (at least) one voter who is mis-reporting their preference ordering who is no better off than at the sincere voting equilibrium, but this player is strictly better off than at the outcome selected by the manipulation which satisfies the Gibbard-Satterthwaite criterion. If outcome b is selected, then strategic voting makes three players better off, and four players worse off as compared to the outcome selected under sincere voting; if outcome a is selected, then strategic voting neither harms nor benefits any voter.

It is then straightforward to consider all possible partitions of the set of voters into blocs of strategic and invariably sincere voters. For any partition of N , if player 7 belongs to S but players 1, 2, 3 and 4 belong to $N \setminus S$, then the unique S -Nash equilibrium of the voting game is c — the outcome which is a Gibbard-Satterthwaite manipulation, but which is not a Nash equilibrium when $N = S$. If player 7 belongs to $N \setminus S$ then the sincere voting outcome is always a S -Nash equilibrium of the voting game. If player 1 (and/or player 2) belongs to S , as well as player 7, but players 3 and 4 belong to $N \setminus S$, then b is the unique S -Nash equilibrium of the voting game, whereas if player 3 (and/or player 4) belongs to S as well as player 7, but players 1 and 2 belong to $N \setminus S$ then a is the unique S -Nash equilibrium of the voting game. If S consists of player 7, along with one of voters 1 or 2 and one of voters 3 or 4 then there are two S -Nash equilibria of the voting game: one in which a is selected, and the other in which b is selected.

This suggests that the standard Gibbard-Satterthwaite definition of manipulation could be usefully modified to require all (strategic) voters to choose voting strategies which are best responses to the voting strategies of the other players. Consider, therefore a true preference profile R^N , and let the reported preference profile \tilde{Q}^N be a Nash manipulation if $\forall i \in S$ such that $\tilde{Q}^i \neq R^i$, (i) $Pl(\tilde{Q}^N) \succeq_i Pl(R^i, \tilde{Q}^{-i})$ and (ii) if $Pl(\tilde{Q}^N) \neq Pl(R^N)$ then there exists at least some $i \in S, \tilde{Q}^i \neq R^i$ such that $Pl(\tilde{Q}^N) \succ_i Pl(R^N)$, or if $Pl(\tilde{Q}^N) = Pl(R^N)$ then there exists at least some $i \in S, \tilde{Q}^i \neq R^i$ such that $Pl(\tilde{Q}^N) \succ_i Pl(R^i, \tilde{Q}^{-i})$. As with a Gibbard-Satterthwaite manipulation, our definition of a Nash manipulation requires all strate-

gic voters to report their preferences sincerely, unless there is some benefit to doing otherwise. Here, however, the benefit may be to block an outcome which would be even worse (for them) than the outcome which prevails when other voters misreport their preferences, and it is possible that they nonetheless end up worse off than would be the case at the sincere voting outcome. Additionally, as with the Gibbard-Satterthwaite definition of manipulation, *someone* must benefit from any departure from the sincere voting outcome, ensuring that strategic voting leads to a movement along the Pareto frontier, rather than a loss in efficiency.

STILL TO BE ADDED: PROOF THAT INTERSECTION OF NASH MANIPULATION + NASH EQUILIBRIA IS NON-EMPTY (COMPLETE.); CALCULATION OF UPPER BOUND (COMPLETE); CALCULATE OF LOWER BOUND (IN PROGRESS).

5 Conclusions

Whereas the social choice literature traditionally grades collective choice mechanisms with respect to their capacity to deliver the outcome that would be selected if all participants report their preferences truthfully, the analysis conducted in this paper explores an alternative metric for judging voting mechanisms, namely, the maximum and minimum bounds on the proportion of the population which benefits from manipulation when voting proceeds under the plurality rule. This is the first paper to calculate such bounds. The salience of this alternative metric is undeniable given that the mechanisms the most widely used in actual elections are all known to be vulnerable to manipulation by strategic voters. Voting procedures which are likely to select outcomes that are preferred by a broad swathe of the population when manipulated to the outcome that is selected when all voters report their true preferences are arguably better mechanisms than ones which are more likely to generate a large proportion of losers. One of the lessons to be drawn from this analysis is that - at least under the plurality rule - strategic voting may often be a virtue, rather than a vice, and citizens should indeed be encouraged to vote with their heads, rather than their hearts.

The importance of this alternative metric is well illustrated by the most recent US election cycle. Arguably, the key factor underlying Donald Trump's success in the Republican primaries was that the anti-Trump vote was split for too long amongst too many candidates, and that many of the voters who supported Rubio, Cruz and Kasich persisted in voting sincerely for their preferred candidate, rather than coalescing around a single opponent, thereby clearing a path for Trump to victory. Moreover, once Trump was confirmed as the Republican Presidential candidate, it was arguably

crystal clear to those anti-Trump voters who preferred a third-party candidate to Hillary Clinton that it was risky to vote sincerely. Whether or not these warnings were heeded is of course a matter of speculation, but it is worth noting that in many of the so-called swing states which ended up favouring Trump (including Michigan, Wisconsin, and Pennsylvania), the number of votes separating Trump and Clinton was smaller than the vote totals which accrued to Jill Stein (of the Green Party). In this particular election, the plurality voting mechanism arguably delivered the outcome that would have been selected had all voters cast their ballots sincerely. However, had more voters been persuaded that it was acceptable to vote strategically (either at the initial primary stage, or later in the Presidential election) it is certainly possible that a President who was ultimately selected who would have commanded broader support.

References

- Alvarez, R Michael, and Jonathan Nagler (2000) ‘A new approach for modelling strategic voting in multiparty elections.’ *British Journal of Political Science* 30(01), 57–75
- Barbera, Salvador (2010) ‘Strategy-proof social choice.’ *Handbook of Social Choice and Welfare*, K. Arrow and K. Suzumura (eds)
- Blais, André (2002) ‘Why is there so little strategic voting in canadian plurality rule elections?’ *Political Studies* 50, 445–454
- Blais, André, Richard Nadeau, Elisabeth Gidengil, and Neil Nevitte (2001) ‘Measuring strategic voting in multiparty plurality elections.’ *Electoral Studies* 20, 343–352
- Brennan, Jason (2011) *The Ethics of Voting* (Princeton University Press)
- Degan, Arianna, and Antonio Merlo (2009) ‘Do voters vote ideologically?’ *Journal of Economic Theory* 144(5), 1868–1894
- Erat, Sanjiv, and Uri Gneezy (2012) ‘White lies.’ *Management Science* 58, 723–733
- Gibbard, Allan (1973) ‘Manipulation of voting schemes: a general result.’ *Econometrica* pp. 587–601
- Gneezy, Uri (2005) ‘Deception: The role of consequences.’ *The American Economic Review* 95, 384–394

- Kawai, Kei, and Yasutora Watanabe (2013) ‘Inferring strategic voting.’ *The American Economic Review* 103(2), 624–662
- Li, Christopher M, and Ricardo Santiago Pique Cebrecos (2016) ‘Ethical motives for strategic voting.’ Available at SSRN: <https://ssrn.com/abstract=2318096> or <http://dx.doi.org/10.2139/ssrn.2318096>
- Muller, Daniel, and Lionel Page (2015) ‘A new approach to measure tactical voting: evidence from the british elections.’ *Applied Economics* 47(36), 3839–3858
- Satterthwaite, Mark Allen (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
- Shachar, Ron (2003) ‘Party loyalty as habit formation.’ *Journal of Applied Econometrics* 18(3), 251–269
- Vanberg, Christoph (2008) ‘Why do people keep their promises? an experimental test of two explanations.’ *Econometrica* 76, 1467–1480