Manipulation & Single-peakedness: A General Result *

Elizabeth Maggie Penn†  Sean Gailmard‡  John W. Patty§

August 12, 2010

Abstract

This paper considers environments in which individual preferences are single-peaked with respect to an unspecified, but unidimensional, ordering of the alternative space. We show that in these environments, any institution that is coalitionally strategy-proof must be dictatorial. Thus, any non-dictatorial institutional environment that does not explicitly utilize an a priori ordering over alternatives in order to render a collective decision is necessarily prone to the strategic misrepresentation of preferences by an individual or group. Moreover, we prove in this environment that for any nondictatorial institution, the truthful revelation of preferences can never be a dominant strategy equilibrium. Accordingly, an incentive to behave insincerely is inherent to the vast majority of real-world lawmaking systems, even when the policy space is unidimensional and the core is nonempty.

*We are grateful for helpful suggestions from Tom Schwartz, Ken Shepsle, Rick Wilson, two anonymous reviewers, and seminar participants at Princeton University, the University of Wisconsin—Madison, the Harris School of Public Policy, and the California Institute of Technology. Of course, all remaining errors are the fault of the authors.

†Associate Professor, Department of Political Science, Washington University in St. Louis. Email: penn@wustl.edu.

‡Associate Professor, Department of Political Science, University of California, Berkeley. Email: gailmard@berkeley.edu.

§Associate Professor, Department of Political Science, Washington University in St. Louis. Email: jpatty@wustl.edu.
The question of manipulation – the misrepresentation of one’s true preferences in order to achieve a more favorable collective decision – has alternately intrigued and frustrated social scientists for over two centuries.\(^1\) Manipulation is clearly a potential problem for those interested in drawing inferences from individuals’ observed behaviors. For example, how does a legislator’s roll call vote relate to his or her policy preferences? How does an individual voter’s vote choice reflect his or her preferences over the parties and/or candidates? More subtly, does the composition of a legislative committee reflect the preferences of the members of the legislature? Does a juror’s vote to convict a defendant truthfully reflect the juror’s beliefs about the defendant’s guilt or innocence? Is the President’s choice of political appointees indicative of the policies that he or she would pursue in their positions? In short, manipulation is simultaneously a methodological and normative conundrum of broad relevance to those who study politics.

Furthermore, the conundrum is – in a very precise sense – inescapable: Allan Gibbard and Mark Satterthwaite famously and independently demonstrated that, when there are at least 3 alternatives to choose from, any collective choice procedure that eliminates the potential incentive for manipulation must be dictatorial.\(^2\) The Gibbard-Satterthwaite theorem closely mirrors Kenneth Arrow’s impossibility result, which states that any democratic institution that can aggregate individual preferences so as to always produce a transitive social ranking of the alternatives without violating either Pareto efficiency or independence from irrelevant alternatives must also be dictatorial.\(^3\) The theorems of Arrow, Gibbard, and Satterthwaite all have strongly negative implications about the predictability and performance of democratic group choice. Perhaps the most famous arguments along these lines was offered by Riker (1982).\(^4\) It is indeed accurate to summarize the perception of democratic governance following these seminal results as “bedevilled by impossibil-

---

\(^1\) Illustrating the topic’s long lineage is Condorcet’s 1788 treatise “On the Constitution and the Functions of Provincial Assemblies,” in which argued that Borda’s scoring method was highly prone to strategic manipulation by voters (c.f., Young (1995)).

\(^2\) Gibbard (1973); Satterthwaite (1975).

\(^3\) Arrow (1963).

\(^4\) Of course, Riker’s arguments have – quite rightfully – not gone unchallenged. While they adopt very different starting points, both Mackie (2003) and McGann (2006) have recently offered counterarguments to Riker’s criticisms of democratic institutions.
ity, instability and arbitrariness.”

On the other hand, Black’s median voter theorem implies that Arrow’s negative conclusions regarding preference aggregation disappear if one presumes that the policies can be ordered along a left-right spectrum so that every individual’s preference over the alternatives is single-peaked. The assumption that individuals have single-peaked preferences guarantees the existence of at least one alternative that can not be defeated by any other under majority rule (i.e., the majority rule core is nonempty). Furthermore, the importance of this result for the development of political theory over the past half century is hard to overstate: contrasted with the results of McKelvey (1976), Schofield (1978), Rubinstein (1979), and Riker (1980), the theoretical power of undimensional spatial model underlying the institutional theories of the past 30 years has proven undeniably productive in advancing our understanding of the interactions between individual preferences, strategic behavior, and political institutions. Indeed, the central insight of the notion of structure-induced-equilibrium (Shepsle (1979)) is that institutions might be structured so as to reduce a single, difficult-to-predict multidimensional policy decision into a well-ordered set of unidimensional, partial policy decisions. The following quotes demonstrate that the distinction between the unidimensional and multidimensional settings has been repeatedly cast as dispositive with respect to, variously, the simplicity, stability, predictability, and even coherence of political decision making.

“When all individuals have single-peaked preference orderings the process of collective decision making is dramatically simplified.” (Feld and Grofman, 1988, p. 776)

“Where preferences are single peaked . . ., one option must be the Condorcet winner and it would be possible to find this by repeated binary votes. (Miller, 1992, p. 63)

“In a unidimensional world, there is no incentive for ‘sophisticated voting.’” (Mouw and Mackuen, 1992, p. 102 (fn. 4))

---

6 Black (1948).
“[When preferences are single-peaked] voters know the political world is coherently organized, the possibility of cycles is zero and the method of majority rule is wholly consistent and never tyrannical.” (Riker, 1992, p. 107)

“[R]elaxation of [unrestricted domain] provides acceptable escape-routes from Arrow’s theorem and from the Gibbard-Satterthwaite theorem, compatible with all other conditions of these theorems.” (Dryzek and List, 2003, pp. 27-28)

“We do know for sure that if the distribution of preference orders is such that they are single-peaked, the Gibbard-Satterthwaite Theorem does not apply, there is no chance for strategic voting to succeed.” (Mackie, 2003, p. 161)

In this paper, we distinguish between the implications of single-peakedness for Arrow’s result and the implications of this assumption for the Gibbard-Satterthwaite theorem. Specifically, while single-peakedness is sufficient to escape from the negative conclusions of Arrow, we demonstrate in this paper that the assumption of single-peaked preferences is not sufficient to eliminate the possibility of profitable manipulation of democratic collective choice. That is, when there are at least 3 alternatives to choose from, the potential for strategic behavior is endemic to all democratic political institutions, even when preferences are single-peaked.8

Our results rest upon the fact that, in order for single-peaked preferences to provide an escape-route from the conclusions of Gibbard-Satterthwaite, one must require that individuals may vote for alternatives only in a way that is consistent with the underlying ordering of alternatives. In other words, if the underlying ordering of alternatives on the left-right spectrum is $X < Y < Z$, then ruling out profitable manipulation requires that one disallow any individual, for example, from casting a vote for $x$ over $z$ and a vote for $z$ over $y$. In other words, Black’s theorem implies that

---

8It is almost obligatory to note at this point that – in spite of its potentially unsettling mendaciousness – “manipulation” need not be a bad thing. For example, Miller (1977) shows that when all individuals vote strategically, outcomes may be obtained that are Pareto superior to outcomes obtained if all individuals vote truthfully. Furthermore, “manipulation” is a large tent, encompassing a wide array of behaviors. For example, Dowding and Van Hees (2007) distinguish between “sincere” and “non-sincere” forms of manipulation.
majority rule is strategy-proof provided that individuals aren’t too strategic. In this sense, our results illustrate that, even with a strong restriction on the structure of the underlying preferences, “institutions matter” insofar as every democratic procedure necessarily generates the incentive for strategic misrepresentation by one or more individuals.

Thus, the arguments in this paper demonstrate that Riker’s famous – and arguably unwarranted – dismissal of single-peakedness as empirically irrelevant was unnecessary for his conclusion that political outcomes “are the consequence not only of institutions and tastes, but also of the political skill and artistry of those who manipulate agenda, formulate and reformulate questions, generate ‘false’ issues, etc., in order to exploit the disequilibrium of tastes for their own advantage.” (Riker, 1980, p. 445) Riker’s discussion of political strategy – particularly the introduction of ‘false issues’ is directly spoken to by our results. This is because one of the keys to our result is that knowing with certainty that the alternatives can be ordered so as to induce a single-peaked profile of preferences is not equivalent to knowing how the alternatives will be ordered. As we demonstrate in both the proof of the result and two motivating examples following it, the assumption of single-peaked preferences is not sufficient to rule out some situation existing in which a group of individuals agree that they would prefer to misrepresent (or, perhaps, more charitably, “dispute”) the true left-right ordering of alternatives. Thus, referring to the passage from Riker (1992) quoted earlier, our results demonstrate while the presumption of single-peaked preferences does imply that collective preference is “coherently organized,” and the possibility of sincere majority preference cycles is indeed zero, one should still not neglect political institutions, as any democratic choice procedure necessarily holds the possibility of a revealed collective preference cycle.


discuss further.

9This point was – to our knowledge – first noted by Blin and Satterthwaite (1976), who examined one specific collective choice rule (majority rule with Borda completion) and showed that it is strategy-proof when preferences and ballots are required to be single-peaked with respect to a common ordering, but is manipulable when ballots are no longer required to be single-peaked with respect to the common ordering. Our results are stronger in one sense – we consider all collective choice rules – and somewhat qualified in another – we focus on coalitional strategy-proofness, which is a stronger condition for a collective choice rule to satisfy. In the canonical special cases of exactly three voters, our results are strict generalization of the insights of Blin and Satterthwaite.

1 Notation and Definitions

There is a finite collection of $K$ alternatives (or policies), $X$, and a finite collection of $n$ individuals (or voters) $N$. We assume that $K \geq 3$ and $n \geq 2$. Individual $i$'s preferences are represented by a strict, transitive and complete binary relation $P_i$. The notation $xP_i y$ implies that $i$ strictly prefers $x$ to $y$. If $x^*_i P_i y$ for all $y \neq x^*_i$, then $x^*_i$ is referred to as $i$'s most-preferred policy or ideal point.

Throughout, $\rho = (P_1, ..., P_n)$ denotes an $n$-dimensional preference profile describing the preferences of all individuals, and we let $x^*_i(\rho)$ denote $i$'s ideal point at profile $\rho$. The notation $\mathcal{P}^n$ represents the collection of all $n$-dimensional profiles of strict orders on $X$. Any nonempty set $D \subseteq \mathcal{P}^n$ is referred to as a preference domain, and when $D$ is a strict subset of $\mathcal{P}^n$ then $D$ is referred to as a restricted domain. We discuss restricted domains in more detail in Section 1.2. For any preference profile $\rho \in \mathcal{P}^n$, $\rho|_S$ denotes the restriction of $\rho$ to the set of alternatives $S \subseteq X$. Similarly, for any individual preference $P_i \in \mathcal{P}$, $P_i|_S$ denotes the restriction of $i$'s preference relation to the set $S$. For any preference profile $\rho \in \mathcal{P}^n$ and pair of alternatives $(x, y) \in X^2$, the notation $P(x, y; \rho) \equiv \{i \in N : xP_i y\}$ denotes the set of individuals who strictly prefer $x$ to $y$ under $\rho$.

1.1 Collective Choice Functions

A collective choice function, or choice function, is any function, $\phi : \mathcal{P}^n \rightarrow X$ that maps any strict profile of orderings over alternatives into the policy space $X$. Throughout, we will assume that $\phi$ has full range: for any $x \in X$, there exists a $\rho \in \mathcal{P}^n$ such that $\phi(\rho) = x$. The notation $\phi(\rho) = x$ says that choice function $\phi$ produces outcome $x$ at profile $\rho$.

While we require choice functions to map any strict profile into a social outcome, we do not require individuals’ true preference orderings to be drawn from the full set $\mathcal{P}^n$. This is because we are interested in the, possibly insincere, behavior induced by a choice function when true individual preferences are drawn from a restricted domain. We call the preference domain of a choice function $\mathcal{D}$, while the ballot domain of all choice functions is assumed to be $\mathcal{P}^n$. In other
words, while true individual preferences may come from a restricted set of orderings, individual behavior is only required to appear individually rational, in the sense of being rationalizable by a transitive binary relation. Throughout, we will use the notation \((P'_i, \rho_{-i})\) to denote a ballot profile in which \(i\) submits ballot \(P'_i\), and all others submit ballots as in profile \(\rho\). More generally, the notation \((\rho'_L, \rho_{-L})\) denotes a ballot profile in which all members \(i \in L \subseteq N\) submit ballots as under \(\rho'\), and all individuals not in \(L\) submit ballots as under \(\rho\).

The following definitions characterize several properties of collective choice functions.

**Definition 1 (Weakly Paretian).** A collective choice function \(\phi\) is weakly Paretian on \(D\) if for all \(\rho \in D\) and all \((x, y) \in X^2\),

\[ P(x, y; \rho) = N \Rightarrow \phi(\rho) \neq y. \]

**Definition 2 (Monotonic).** A collective choice function \(\phi\) is monotonic on \(D\) if, for all \((x, y) \in X^2\) and all \(\rho, \rho' \in D\),

\[ P(x, y; \rho) \subseteq P(x, y; \rho') \text{ and } \phi(\rho) = x \Rightarrow \phi(\rho') \neq y. \]

**Definition 3 (Dictatorial).** A collective choice function \(\phi\) is dictatorial if for some \(i \in N\) and for all \(\rho \in \mathcal{P}^n\),

\[ \phi(\rho) = x^*_i(\rho), \]

where \(x^*_i(\rho)\) is \(i\)'s reported ideal point under ballot profile \(\rho\).\textsuperscript{11}

**Definition 4 (Strategy-Proof (SP)).** A collective choice function \(\phi\) is manipulable if, for some \(\rho = (P_1, ..., P_n) \in D\) and \(i \in N\) there exists a \(P'_i \in \mathcal{P}\) such that

\[ \phi(P'_i, \rho_{-i}) P_i \phi(\rho). \]

\textsuperscript{11}Note that the previous definitions are defined on the preference domain \(D\) while dictator is defined on ballot domain \(\mathcal{P}^n\). We make these definitional choices conservatively; we are ultimately going to prove that coalitional strategy-proofness when preferences are single-peaked will imply dictator on the full ballot domain. Our intermediary steps in this proof will only utilize the weak Pareto and monotonicity properties on the restricted preference domain. However, by implying dictator on the full ballot domain, our proof also guarantees that the collective choice functions we consider are both weakly Pareto and monotonic on the full ballot domain.
A choice function is strategy-proof if it is not manipulable.

**Definition 5** (Coalitionally Strategy-Proof (CSP)). A collective choice function \( \phi \) is coalitionally manipulable if, for some \( \rho = (P_1, ..., P_n) \in D \) and \( L \subseteq N \) there exists a \( \rho' \in \mathcal{P}^n \) such that

\[
\phi(\rho'_L, \rho_{-L}) \succ_i \phi(\rho) \quad \text{for all} \quad i \in L.
\]

A choice function is coalitionally strategy-proof if it is not coalitionally manipulable.

Note that individual manipulability of social choice functions implies coalitional manipulability (for a coalition of one); equivalently, coalitional strategy proofness implies strategy proofness. The converse is not generally true; there may be instances in which a social choice function may be manipulable by a sufficiently large coalition, but not by an individual. However, when \( D = \mathcal{P}^n \), then individual strategy proofness implies coalitional strategy proofness, because it implies dictatorship. Thus, in a setting with unrestricted domain, individual and coalitional manipulability and non-manipulability are equivalent.

### 1.2 Single-Peaked Preferences

In this section we define the domain of single-peaked preferences. This domain has attracted the interest of many scholars because it has been shown to lead to the existence of non-dictatorial Arrovian preference aggregation rules and, when ballots are required to be single-peaked, to non-manipulable, non-dictatorial collective choice functions. Our interest is less about existence, and more about the characterization of choice functions on this restricted domain.

**Single-Peaked Preferences.** A preference profile is single-peaked if there exists a way of ordering the collection of alternatives along a left-right scale so that each individual’s ranking of the alternatives decreases as one moves away from his ideal point. (Austen-Smith and Banks, 1999, p. 93) We denote the single-peaked preference domain by \( S \subseteq \mathcal{P}^n \). We will at times denote by \( Q_\rho \)
a particular ordering of alternatives with respect to which profile $\rho \in S$ is single-peaked. When referring to an ordering $Q$, if alternative $x$ is above $y$ with respect to $Q$ we write $xQy$.

While the single-peaked preference restriction is widely utilized and intuitively quite simple, Ballester and Haeringer (2007) prove that the set $S$ is completely characterized by two conditions, worst-restriction\textsuperscript{12} and $\alpha$-restriction, both defined below.

**Definition 6** (Worst-restriction). A profile $\rho$ is worst-restricted if, for every triple of alternatives, $(x, y, z) \in X^3$, there exists an $a \in \{x, y, z\}$ such that for all $i \in N$, $aPib$ for some $b \in \{x, y, z\} \setminus \{a\}$.

In words, a profile is worst-restricted if for every triple $(x, y, z) \in X^3$, there is some element of that triple that no individual ranks last relative to the other two elements of the triple.

**Definition 7** ($\alpha$-Restriction). A preference profile $\rho$ is $\alpha$-restricted if there do not exist two agents, $i, j \in N$, and four alternatives $w, x, y, z$ such that

1. The preferences over $w, x, \text{ and } z$ are opposite: $wP_i xP_iz$ and $zP_jxP_jw$.

2. The players both rank $y$ higher than $x$: $yP_ix$ and $yP_jx$.

**Theorem 1** (Ballester and Haeringer (2007)). A preference profile is single-peaked if and only if it satisfies worst-restriction and $\alpha$-restriction.

It is important to note at this point that the domain $S$ is the set of all single peaked preference profiles. In other words, in a priori terms, any ordering of the alternatives is possible.\textsuperscript{13}

\textsuperscript{12}See Sen (1966) and Sen and Pattanaik (1969) for a more thorough discussion of worst-restriction.

\textsuperscript{13}This point is a technical one, but important for broader considerations of the results in this paper. In particular, for any given linear ordering of the alternatives, $Q \in \mathcal{P}$, one can identify the set of preferences that are single-peaked with respect to $Q$, this set is denoted by $S_Q$, and the set of all profiles of such preferences is denoted by $S^d_Q$, a rectangular (product) space. Conversely, the set $S$ that we consider is not rectangular. Unlike $S$, the space $S^d_Q$ is widely discussed in the political economy literature. In particular, it has been shown that in domain $S^d_Q$ the collection of strategy-proof collective choice functions can be completely characterized by the set of augmented median voter rules. See Moulin (1980), Border and Jordan (1983), Austen-Smith and Banks (2004, Theorem 2.4, p. 41).
Ubeda (2003) has recently used a different domain restriction, the 2-free triple domain \( T_2^n \),\(^\text{14}\) to demonstrate that on any domain satisfying this restriction, any weakly Paretian and IIA aggregation rule must be neutral, where neutrality characterizes those aggregation rules that disregard the labeling of alternatives when making pairwise comparisons. Because our paper is concerned with collective choice rules and not preference aggregation rules \emph{per se}, we omit the technical details of Ubeda’s result and our previously published Theorem 2 below. The key distinction between Ubeda’s result and ours is that the 2-free triple domain and the single-peaked domain are not nested. In other words, satisfaction of either the 2-free triple restriction or single-peakedness does not imply satisfaction of the other. We provide Theorem 2 below because it demonstrates the implications of Arrow’s axioms on the single-peaked domain, and thus is substantively linked to our results that follow concerning manipulation on this domain. In Section 3.1 we will more explicitly discuss how the implications of single-peakedness for Arrow’s result differ from the implications for the Gibbard-Satterthwaite theorem.

**Theorem 2** (Gailmard \textit{et al.} (2008)). \textit{Let } \( F \) \textit{be a weakly Paretian and IIA preference aggregation rule. If } \( D = S \), \textit{then } \( F \) \textit{is neutral.}

\section{Manipulation on Single-Peaked Domains}

With these preliminaries in hand, we are now in a position to state and prove our main results. First we will prove that single-peakedness is not sufficient to ensure that groups of individuals cannot be made better off by submitting insincere ballots; in other words, any nondictatorial collective choice function is necessarily manipulable by a coalition at a single-peaked profile of preferences. Then we will show that a collective choice function being manipulable by a coalition implies that it cannot be a dominant strategy equilibrium for all individuals to truthfully reveal their preferences.

\footnote{This domain restriction says that for any triple of alternatives, only two orderings of the triple are possible across all individuals. While profiles on this domain will satisfy worst-restriction, they may fail \( \alpha \)-restriction, with a clear example being the case with two individuals with preferences: \( wP_1yP_1xP_1z \) and \( zP_2yP_2xP_2w \). Similarly, the following 3-player profile is single-peaked but is not an element of the 2-free triple domain: \( xP_1yP_1z, yP_2xP_2z \), and \( zP_3xP_3y \).}
Thus, when a non-dictatorial institution is required to consider all possible collections of ballots as potential inputs, truth telling can never be a dominant strategy for all individuals, even when preferences are single-peaked.

2.1 The impossibility of coalitional strategy-proofness

To prove that coalitional strategy-proofness implies dictatorship on preference domain $D = S$, we utilize the following lemmas and theorem.

**Lemma 1.** Let $\phi$ be a coalitionally strategy-proof collective choice function with full range. If $D = S$, then $\phi$ is weakly Paretian.

*Proof:* Consider a $\rho \in S$ such that for all $i \in N$, $x P_i y$, but $\phi(\rho) = y$. By full range, $\exists \rho' \in P^n$ with $\phi(\rho') = x$. Then $\phi(\rho') P_i \phi(\rho)$ for all $i \in N$, and $\phi$ is manipulable by coalition $N$. It follows that $\phi(\rho) \neq y$ if $\phi$ CSP. □

**Lemma 2.** Let $\phi$ be a coalitionally strategy-proof collective choice function with full range. If $D = S$, then $\phi$ is monotonic.

*Proof:* Suppose that $\rho$ and $\rho'$ are single-peaked, but violate monotonicity, with $\phi(\rho) = x, \phi(\rho') = y$, and $P(x, y; \rho) \subseteq P(x, y; \rho')$. We will show that this implies $\phi$ is not CSP on $S$. Throughout the proof, let $P(x, y; \rho) = A$ and $P(y, x; \rho') = B$. We know that $A \cap B = \emptyset$.

First, to simplify notation, change $\rho$ so that for all $i \in A$, $P_i$ is replaced by a new preference ordering $P_A$. We construct $P_A$ so that $x$ is top-ranked and $y$ is as high in the $P_A$ ranking as possible while maintaining $P_A$ single-peaked with respect to $Q_\rho$. It is easy to verify that such an ordering exists.

Similarly, change $\rho'$ so that for all $j \in B$, $P'_j$ is replaced by a new preference ordering $P'_B$ that is constructed so that $y$ is top-ranked and $x$ is as high as possible in the $P'_B$ ordering while maintaining that $P'_B$ is single peaked with respect to $Q_{\rho'}$. CSP implies that for each of these new profiles (which in an abuse of notation we will still call $\rho$ and $\rho'$) $\phi(\rho) = x$ and $\phi(\rho') = y$. 11
Consider the triple, \(x, y, z \in X\), with \(z\) arbitrary. \(\rho, \rho' \in S\) imply that for each profile, one of these elements cannot be lowest-ranked by any individual. If at profile \(\rho\) the two remaining elements are \(a, b\), we say that \((a, b)\) is lowest-ranked for \(\rho|_{\{a, b, z\}}\).

Construct a new profile \(\rho^* \in \mathcal{P}^n\) in which \(P^*_i = P_A\) for \(i \in A\), and \(P^*_j = P'_B\) for \(j \in B\). Note that rankings are unspecified for \(k \notin A \cup B\), if \(N \setminus \{A \cup B\} \neq \emptyset\). These individuals’ rankings can be arbitrarily assigned. Also note that \(\rho^*\) may not be single-peaked; our choice function is still required to produce an outcome at this profile, however. Then it must be the case that \(\phi(\rho^*) \notin \{x, y\}\), else either coalition \(N \setminus A\) could manipulate \(\rho\) with \(\rho^*\), or \(N \setminus B\) could manipulate \(\rho'\) with \(\rho^*\). Thus, \(\phi(\rho^*) = z\).

Note that in constructing \(\rho\) (and \(\rho^*\)) above, we have ensured that the only instances in which \(zP_A y\) are those in which \(z\) lies between \(x\) and \(y\) under the ordering \(Q_\rho\) (i.e. \(xQ_\rho zQ_\rho y\), or the reverse). This is because, while maintaining both single-peakedness and \(x\) as top-ranked for \(P_A\), \(y\) is ranked as high as possible for \(P_A\). Similarly, \(zP'_B x\) implies that \(z\) lies between \(x\) and \(y\) according to \(Q_{\rho'}\).

Case 1: \((x, y)\) is lowest-ranked for either \(\rho|_{\{x, y, z\}}\) or for \(\rho'|_{\{x, y, z\}}\). Without loss of generality, assume that \((x, y)\) is lowest-ranked for \(\rho|_{\{x, y, z\}}\). This implies that for all individuals \(k \notin A\), \(yP_k x \Rightarrow zP_k x\). Thus, \(\phi\) is manipulable at \(\rho\) by coalition \(N \setminus A\) submitting ballots as in \(\rho^*\); these individuals can guarantee themselves the outcome \(z\), which they all prefer to \(x\). It follows that \(\phi\) is not CSP.

Case 2: \(z\) is a lowest ranked element of both \(\rho|_{\{x, y, z\}}\) and \(\rho'|_{\{x, y, z\}}\). Note that by the construction of \(\rho\) and \(\rho'\) above, \(yP_A z\) and \(xP'_B z\).\(^{15}\)

To recap, we have that all members of \(A\) (resp. \(B\)) have the same preference ordering over all alternatives, and that this ordering ranks \(xP_A yP_A z\) (resp. \(yP'_B xP'_B z\)). Note also that \(x, y, z\) need not appear consecutively in these individuals’ rankings. Furthermore, \(\rho^*\) as constructed above still

\(^{15}\text{This immediately implies that } \phi \text{ does not satisfy Pareto efficiency on the full domain } \mathcal{P}^n, \text{ because } \rho^* \text{ can be constructed so as to have all individuals not in } A \cup B \text{ place } x \text{ and } y \text{ at the top of their ballots, and so } xP'_k z \text{ and } yP'_k z \text{ for all } k, \text{ and yet } \phi(\rho^*) = z.\)
yields $z$ as its choice. However, now members of coalition $N \setminus A$ (resp. $N \setminus B$) may rank $z$ last relative to $x$ and $y$, and may not have an incentive to manipulate $\phi$ by submitting ballots as in $\rho^*$. We will now construct a new $\rho^o \in S$ and show that $\phi$ CSP for $\rho$ implies that $\phi$ not CSP for $\rho^o$, thus establishing a contradiction. This will take several intermediate steps. First, consider the ordering over alternatives induced by the preferences of individuals in $A$ under $\rho$; we will call this ordering $Q_A$, and we know that under this ordering $xQ_AyQ_Az$. Construct a new profile $\hat{\rho}$ where $\hat{P}_i = P_A$ for all $i \in A$. Clearly this $\hat{P}_i$ is single-peaked with respect to the ordering $Q_A$, as it is this ordering. For all $j \not\in A$, assign each member of $j$ an identical preference ordering $\hat{P}_j$ that ranks $y$ at the top of the ballot, ranks $z\hat{P}_jx$, and is single-peaked with respect to $Q_A$. Such an ordering exists because $y$ lies between $z$ and $x$ according to $Q_A$. Thus, $\hat{\rho} \in S$.

We know that coalition $N \setminus A$ can submit ballots as in $\rho^*$ and receive $z$ as an outcome, which they prefer to $x$. Thus, $\phi(\hat{\rho}) \neq x$. Furthermore, $\phi(\hat{\rho}) \neq z$, by weak Pareto, because for all $i \in N$, $y\hat{P}_i z$. And by CSP of $\rho$, $\phi(\hat{\rho}) \neq y$, because then $\rho$ would be manipulable by coalition $N \setminus A$ submitting ballots as in $\hat{\rho}$. Thus, if $|X| = 3$ the proof is complete. If not, it follows that $\phi(\hat{\rho}) = w$, with all members $j \in N \setminus A$ ranking $y\hat{P}_j w\hat{P}_j z\hat{P}_jx$. However, this implies that for all $i \in A$, $wP_{Ay}$, otherwise $\hat{\rho}$ would be manipulable by coalition $A$ submitting ballots identical to those for $N \setminus A$, and ensuring an outcome of $y$ by weak Pareto. Thus $Q_A$ ranks the alternatives $xQ_A wQ_A yQ_A z$.

Now consider a new profile $\hat{\rho}^1$ in which everyone in $A$ ranks the alternatives as in $\hat{\rho}$. For all $j \in N \setminus A$, assign each $j$ an identical preference ordering that ranks $y$ at the top, ranks $z\hat{P}_j^1 w\hat{P}_j^1 x$, and is single-peaked with respect to $Q_A$. Again, such an ordering exists given that $Q_A$ ranks the alternatives $xQ_A wQ_A yQ_A z$. Thus, $\hat{\rho}^1 \in S$. Using the same logic as above, we get that $\phi(\hat{\rho}^1) \neq x$ or $w$ (else $N \setminus A$ manipulates with $\rho^*$ to get $z$), that $\phi(\hat{\rho}^1) \neq z$ by Pareto efficiency, and that $\phi(\hat{\rho}^1) \neq y$ by CSP of our original $\rho$. If $|X| = 4$, the proof is complete. Otherwise, $\phi(\hat{\rho}^1) = w^1$, with individuals $j \in N \setminus A$ ranking $y\hat{P}_j^1 w^1 \hat{P}_j^1 z\hat{P}_j^1 w^1 \hat{P}_j^1 x$. Again, it follows that for all $i \in A$, $\hat{P}_i^1 = P_A$ ranks $w^1 \hat{P}_i^1 y$, otherwise this coalition would manipulate $\hat{\rho}^1$ in order to get $y$ as the outcome. Thus, $Q_A$ ranks the alternatives $xQ_A \{w^1, w\} Q_A y Q_A z$ (the $w^1, w$ ranking need not be specified in the proof, although by coalition $N \setminus A$’s preferences, the ranking must respect
Repeat the above steps for \( k = 2, \ldots, |X| - 4 \) by choosing a new \( \hat{\rho}^k \) with \( y\hat{P}_j^k \hat{z}_j \) for all \( j \in N \setminus A \), and \( \hat{P}_i^k = P_A \) for \( i \in A \). Such a \( \hat{\rho}^k \) can be constructed so as to always remain single-peaked with respect to \( Q_A \) because it must always be the case that the social choice at each stage, \( w^k \), is ranked between \( x \) and \( y \) for all \( i \in A \). We ultimately get that \( Q_A \) ranks the alternatives \( xQ_A \{w, w^1, w^2, \ldots, w^{|X| - 4}\} \) \( Q_A yQ_A z \). At this point, construct \( \rho^o \) so that \( \phi^o \) is such that \( y\phi^o_j \phi^o_j \{w, w^1, \ldots, w^{|X| - 4}\} \phi^o_j x \) for all \( j \in N \setminus A \) and \( \phi^o_i = P_A \). Again, \( \rho^o \) is single-peaked with respect to \( Q_A \). CSP requires \( \phi(\rho^o) \phi^o_j \phi^o_j \) for all \( j \in n \setminus A \), otherwise this coalition would manipulate with ballots as in \( \rho^* \). CSP of \( \phi \) at \( \rho \) requires \( \phi(\rho^o) \neq y \). And these two statements imply a contradiction. Thus, \( \phi \) is not CSP. It follows that \( \phi \) CSP implies \( \phi \) monotonic when \( \mathcal{D} = S \). \( \square \)

For the next lemma we will need the following two definitions:

**Definition 8** (Blocking coalition for \( (x, y) \)). A coalition \( L \subseteq N \) is a blocking coalition for \( (x, y) \) if for all \( \rho = (P_1, \ldots, P_n) \in \mathcal{D} \) such that \( xP_i y \) for all \( i \in L \) and \( yP_j x \) for all \( j \notin L \), then \( \phi(\rho) \neq y \).

**Definition 9** (Blocking coalition). A coalition \( L \subseteq N \) is a blocking coalition if for all \( \rho = (P_1, \ldots, P_n) \in \mathcal{D} \) and all pairs \( (a, b) \in X^2 \), \( aP_i b \) for all \( i \in L \) implies \( \phi(\rho) \neq b \).

**Lemma 3.** \( \phi \) coalitionally strategy-proof with full range implies that if there exists one \( \rho \in S \) with \( xP_i y \) for all \( i \in L \) and \( yP_j x \) for all \( j \notin L \) and \( \phi(\rho) = x \), then \( L \) is a blocking coalition.

**Proof:** Let \( \rho \) be such that \( xP_i y \) for all \( i \in L \) and \( yP_j x \) for all \( j \notin L \) and \( \phi(\rho) = x \). We will first show that \( L \) is blocking for \( (x, y) \), and then that \( L \) is a blocking coalition.

Suppose that \( L \) is not blocking for \( (x, y) \), so that there exists a \( \rho' \in S \) with \( xP_i y \) for all \( i \in L \) and \( yP_j x \) for all \( j \notin L \) and \( \phi(\rho') = y \). By monotonicity, this implies that \( \phi(\rho) \neq x \), a contradiction. Thus, \( L \) is blocking for \( (x, y) \).

We will now show that \( L \) blocking for \( (x, y) \) implies that for any \( a \notin \{x, y\} \), \( L \) is blocking for \( (x, a) \) and for \( (a, y) \). Consider any \( \rho \in S \) where \( i \in L \) rank the alternatives \( xP_i yP_i a \), \( j \notin L \) rank \( wQ_A w^1 \).
them \( yP_j aP_j x \) and all \( k \in N \) have \( cP_k d \) when \( c \in \{a, x, y\} \) and \( d \not\in \{a, x, y\} \). Then \( \phi(\rho) = x \), by \( L \) blocking for \((x, y)\) and by weak Pareto. Thus, \( L \) is blocking for \((x, a)\).

Now construct a \( \rho' \in \mathcal{S} \) with \( P'_j = P_j \) for \( j \not\in L \) and with \( i \in L \) having \( aP_i xP_i y \) and \( cP_i d \) when \( c \in \{a, x, y\} \) and \( d \not\in \{a, x, y\} \). In this case \( \phi(\rho') = a \), again by \( L \) blocking for \((x, y)\) and by weak Pareto. Thus, \( L \) is blocking for \((a, y)\).

Because \( a \) was chosen at random, the above argument proves that \( L \) is also blocking for any distinct pair \((c, d)\): \( L \) blocking for \((x, y)\) implies \( L \) blocking for \((c, y)\), and this implies \( L \) blocking for \((c, d)\), for any \( d \neq y \).

Last, by monotonicity, we will show that \( L \) blocking for all \((a, b)\) implies that at any profile \( \rho \in \mathcal{S} \) in which \( aP_i b \) for all \( i \in L \), then \( \phi(\rho) \neq b \). Suppose not; suppose that \( \phi(\rho) = b \). Let \( Q \) be the ordering that \( \rho \) is single-peaked with respect to. Now consider a \( \rho' \) that is also single-peaked with respect to \( Q \), in which \( P'_i = P_i \) for all \( i \in L \), and for each \( j \not\in L \), \( P_j \) is replaced by a \( P'_j \) in which \( b \) is top-ranked for \( P'_j \). By monotonicity, \( \phi(\rho') = b \). However, under \( \rho' \) we have \( aP_i b \) for all \( i \in L \) and \( bP_j a \) for all \( j \not\in L \). \( \phi(\rho') = b \) contradicts \( L \) blocking for \((a, b)\). Thus, \( L \) is a blocking coalition. \( \square \)

**Theorem 3.** Let \( \phi \) be a coalitionally strategy-proof collective choice function with full range. If \( D = \mathcal{S} \), then \( \phi \) is dictatorial.

**Proof:** Consider three profiles \( \rho_1, \rho_2, \rho_3 \in \mathcal{S} \) in which the alternatives \( x, y, z \) are at the top of each person’s preference ordering, and all other alternatives are ordered according to a fixed ordering \( Q_{-\{xyz\}} \). Thus, save for alternatives \( \{x, y, z\} \), rankings over all other alternatives are identical across all individuals and all three profiles.

Let \( L \subseteq N \) be a “minimal” blocking coalition. Thus, for any \( i \in L \), the set \( L \setminus \{i\} \) is not a blocking coalition. Such a coalition exists because, by Pareto, we know that the collection of blocking coalitions is nonempty. Define \( \{x, y, z\} \) rankings under \( \rho_1, \rho_2, \rho_3 \) as follows:
We know the following: $\phi(\rho_2) \neq z$ because all in $L$ prefer $y$ to $z$; $\phi(\rho_3) \neq z$ because all in $L$ prefer $y$ to $z$; $\phi(\rho_1) \neq y$ because everyone not in $L \setminus \{i\}$ prefers $z$ to $y$, and we have assumed that $L \setminus \{i\}$ is not a blocking coalition; $\phi(\rho_2) \neq y$ because everyone not in $L \setminus \{i\}$ prefers $x$ to $y$, and we have assumed that $L \setminus \{i\}$ is not a blocking coalition. These last two statements are consequence of Lemma 3.

Condensing the above paragraph, we now know that $\phi(\rho_1) = x$ or $z$, that $\phi(\rho_2) = x$, and that $\phi(\rho_3) = x$ or $y$.

Case 1: First, suppose that $\phi(\rho_1) = z$. This implies that $\phi(\rho_3) = y$, because $(x, z)$ preferences are identical across $\rho_1$ and $\rho_3$. Thus, $\phi(\rho_3) = x$ would violate monotonicity.

Now consider an insincere ballot $\hat{\rho}$ that is identical to $\rho_1, \rho_2, \rho_3$ for all $w \notin \{x, y, z\}$ (i.e. these alternatives are, for every individual, ordered according to $Q_{-\{xy,z\}}$), and with a Condorcet cycle over $x, y, z$ at the top:

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\rho}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$i$</td>
<td>$xPzPz$</td>
</tr>
<tr>
<td>$L \setminus {i}$</td>
<td>$yPzPx$</td>
</tr>
<tr>
<td>$N \setminus L$</td>
<td>$zPzPx$</td>
</tr>
</tbody>
</table>

We know that $\phi(\hat{\rho}) = d \notin \{x, y, z\}$, otherwise $\rho_1, \rho_2$ or $\rho_3$ would be manipulable by an individual or coalition submitting ballots as in $\hat{\rho}$: $\phi(\hat{\rho}) = x \Rightarrow i$ manipulates $\rho_1$, $\phi(\hat{\rho}) = y \Rightarrow L \setminus \{i\}$ manipulates $\rho_2$, and $\phi(\hat{\rho}) = z \Rightarrow N \setminus L$ manipulates $\rho_3$.

Now, construct a new profile $\hat{\rho}_1 \in S$ with the preferences of all $j \neq i$ identical to those given by $\hat{\rho}$. Player $i$’s new preferences rank $d$ highest, with $i$’s rankings over all $w \notin \{d, x, y, z\}$ unchanged. This profile is single-peaked according to the ordering specified by $i$’s preferences, which can be
verified by considering $\alpha -$ and worst-restriction. Since all players have identical orderings over alternatives not in $\{d, x, y, z\}$, moving $d$ to the top of $i$'s ranking cannot break worst-restriction (as all other players have the same ranking of $d$ and any other alternative $a \neq d$), and cannot break $\alpha -$restriction, as only player $i$ has preferences over a triple that are the reverse of another player's preferences, and $i$ is the unique player with $d$ at the top of his ballot. $\phi(\hat{\rho}_1) = d$, otherwise $\phi(\hat{\rho}_1)$ would be manipulable by $i$ submitting a ballot as in $\hat{\rho}$. But, by Lemma 3, this implies that $i$ is a blocking coalition, because all other players prefer $x, y$ and $z$ to $d$. Thus, $i$ is a dictator on $S$.

Case 2: It must now be the case that $\phi(\rho_1) = x$. Then immediately, by Lemma 3, this implies that $i$ is a blocking coalition, because $i$ is the unique person who prefers $x$ to $z$ at profile $\rho_1$. Thus, $i$ is a dictator on $S$.

We have shown that $i$ is a blocking coalition on preference domain $S$. To finish, we must show that $i$ is a dictator on the full ballot domain. Suppose not, so that there exists $\rho \in P^n$ with $\phi(\rho) = y$ and with $y \neq x^*_i(\rho)$, with $x^*_i(\rho)$ being $i$'s reported ideal point under ballot profile $\rho$. We also know that there exists a $\rho' \in S$ with $P_i = P'_i$ and with $y = x^*_j(\rho')$ for all $j \neq i$. Voter $i$ being a blocking coalition on $S$ implies $\phi(\rho') = x$. However this contradicts $\phi$ coalitionally strategy-proof, because $\phi(\rho')$ would be manipulable by $N \setminus \{i\}$ submitting ballots as in $\rho$ and attaining their ideal point, $y$. It follows that $i$ is a dictator on the full ballot domain. □

When $n = 3$, Theorem 3 can be strengthened to say that if $\phi$ is a strategy-proof collective choice function and $D = S$, then $\phi$ is dictatorial. However, we cannot weaken coalitional strategy-proofness to strategy-proofness when $n \geq 5$. This point, along with what strategy-proofness and coalitional strategy-proofness imply about dominant strategy and Nash implementation, is explored in the next section.

---

16Such a $\rho' \in S$ would exist and be single-peaked, for example, for an ordering of the policy space $Q$ equal to $i$'s preference ordering, so that $xQy$ if and only if $xP_iy$. 17
2.2 Strategy-proofness and truth telling as a dominant strategy equilibrium

It is important to note that on a restricted preference domain and unrestricted ballot domain, strategy-proofness of a choice function is no longer equivalent to that choice function being truthfully implementable in dominant strategies by a direct mechanism. Strategy-proofness simply tells us that at any truthful profile of ballots, no individual can submit an insincere ballot and strictly benefit from that insincerity. It says nothing about whether individuals may benefit from insincerity at profiles of ballots in which other individuals are also being insincere. When the ballot domain and preference domain are equivalent, the definition of strategy-proofness implies dominant strategy implementability, because every possible strategy (or ballot) profile is also a preference profile. However, in the setting we consider, the unrestricted domain of ballot profiles (or strategy profiles) is larger than the collection of single-peaked preference profiles.

In the next theorem we will leverage our previous result on coalitional strategy-proofness to show that, when considering an unrestricted ballot domain, there exists no nondictatorial choice function that is implementable in dominant strategies when true preferences are single-peaked. When preferences are single-peaked there do, however, exist strategy-proof collective choice functions, or choice functions truthfully implementable in Nash equilibrium. These results highlight the fact that on the single-peaked domain we consider, there is not the typical equivalence between coalitional strategy-proofness and strategy-proofness. The only coalitionally strategy-proof choice function on $S$ is dictator; however in Theorem 5 that follows we provide an example of a non-dictatorial choice function that is strategy-proof on $S$. It is not, however, dominant strategy implementable. First we must introduce some notation.

For any individual $i \in N$, $M_i = \mathcal{P}$ is the collection of actions, or messages, available to individual $i$, and $\mathcal{M} = \prod_{i \in N} M_i$ is the set of all profiles of messages. Recall that $\mathcal{P}$ is the collection of all strict orders on $X$. Thus, an individual’s message is an ordering of the policy space. An outcome function $g$ maps a profile of messages $m = (m_1, ..., m_n) \in \mathcal{M}$ into the policy space, $X$. The pair $(\mathcal{M}, g)$ is a mechanism, or game form. If the collection of messages available to each individual is equal to the set of preferences that the individual may have, the mechanism is
called direct.

Given a mechanism \((\mathcal{M}, g)\), a strategy for person \(i\) is a function \(\sigma_i\) mapping preference profiles \(\rho \in \mathcal{D}\) into messages, so that \(\sigma_i(\rho) = m_i\). A game, \(G\), consists of a mechanism and a preference profile, so that \(G = ((\mathcal{M}, g), \rho)\). For any game \(G\), a strategy profile \(\sigma^*(\rho) \in \mathcal{M}\) is

- a **dominant strategy equilibrium for** \(G\) if and only if for all \(i \in N\), all messages \(m_i \in \mathcal{M}_i\) and all \(m_{-i} \in \prod_{j \neq i} \mathcal{M}_j\),

\[
g(\sigma^*_i(\rho), m_{-i}) R_i g(m_i, m_{-i}),
\]

- a **Nash equilibrium for** \(G\) if and only if for all \(i \in N\) and all \(m_i \in \mathcal{M}_i\),

\[
g(\sigma^*_i(\rho), \sigma^*_{-i}(\rho)) R_i g(m_i, \sigma^*_{-i}(\rho)).
\]

In words, then, a dominant strategy equilibrium is one in which the best choice for any individual \(i\), \(\sigma^*_i(\rho)\) is independent of the strategies played by other players: it is an unconditionally optimal choice. In a Nash equilibrium, on the other hand, the strategy of any given player is optimal, **given the Nash equilibrium strategies being played by other players**.

A choice function \(\phi\) with preference domain \(\mathcal{D}\) and ballot domain \(\mathcal{P}^n\) is

- **truthfully implementable in dominant strategy equilibrium** if and only if there exists a direct mechanism \((\mathcal{M}, g)\) so that for all \(\rho \in \mathcal{D}\), \(\sigma^*(\rho) = \rho\) is a dominant strategy equilibrium for the game \(G = ((\mathcal{M}, g), \rho)\), and \(g(\rho) = \phi(\rho)\),

- **truthfully implementable in Nash equilibrium** if and only if there exists a direct mechanism \((\mathcal{M}, g)\) so that for all \(\rho \in \mathcal{D}\), \(\sigma^*(\rho) = \rho\) is a Nash equilibrium for the game \(G = ((\mathcal{M}, g), \rho)\), and \(g(\rho) = \phi(\rho)\).

With these definitions in hand, the next theorem formally links our social choice result, Theorem 3, with implementation theory.

**Theorem 4.** If \(\phi\) is coalitionally manipulable, then \(\phi\) is not truthfully implementable in dominant strategy equilibrium.
Proof: Let $\phi$ be coalitionally manipulable. Then there exists a coalition $L \subseteq N$, a $\rho \in D$ and a $\rho' \in \mathcal{P}^n$ such that $\phi(\rho'_L, \rho_{-L}) \mathcal{P}_i \phi(\rho)$ for all $i \in L$. Number the individuals in $N$ so that the individuals in $L$ are numbered from 1 through $|L|$. $\phi$ coalitionally manipulable implies that there exists an $i \in L$ such that $\phi(\rho_{\{1,\ldots,i-1\}}, P'_i, \rho_{N \setminus \{1,\ldots,i\}}) \mathcal{P}_i \phi(\rho_{\{1,\ldots,i-1\}}, P_i, \rho_{N \setminus \{1,\ldots,i\}})$. Thus, $\sigma_i(\rho) = P_i$ is not a dominant strategy for person $i$. It follows that $\phi$ is not truthfully implementable in dominant strategy equilibrium. □

The following Corollary follows immediately from Theorems 3 and 4 and summarizes an important implication of our results that we return to in Section 3.2. Specifically, the corollary implies that designing deliberative institutions so as to implement the majority will will require accounting for obfuscation so long as individuals are strategic: sincere revelation can not be made a weakly dominant strategy.

Corollary 1. When $D = S$, there exists no non-dictatorial collective choice function that is truthfully implementable in dominant strategies. In particular, a collective choice function that always returns a Condorcet winner when one exists is not truthfully implementable in dominant strategies.

While Theorem 4 proved that coalitional manipulability precludes the possibility of truthful dominant strategy implementation, it does not rule out the possibility of truthful Nash implementation. However, as noted earlier, when $n = 3$, Theorem 3 can be strengthened to say that if $\phi$ is a strategy-proof collective choice function and $D = S$, then $\phi$ is dictatorial. Thus, when $n = 3$ and $D = S$, no non-dictatorial choice function is truthfully implementable in Nash equilibrium.

We cannot weaken coalitional strategy-proofness to strategy-proofness when $n \geq 5$, and therefore, in this situation we can design a mechanism in which the truthful revelation of preferences is a Nash equilibrium. This is because, with four or more voters and a single person submitting an insincere ballot, the set of potential manipulators can be narrowed to two or fewer individuals in any situation in which the submitted profile of ballots is not single-peaked.\(^\text{17}\) In the absence of a core alternative in the submitted profile of ballots, removing these individuals would yield a profile

\(^{17}\)This is to say that we may not be able to uniquely identify an insincere ballot, but we can identify a unique pair of individuals, one of whom has submitted an insincere ballot.
admitting a non-empty core, and even if the resulting core is multi-valued, the individual can never profit from having his ballot dropped. We run into a minor problem when the core of a sincere ballot profile is potentially multi-valued. In particular, without knowledge of the underlying ordering of alternatives, \( Q \), we cannot break ties within the core in a way that is not manipulable by an individual. However, in this case we can specify an individual whose ballot is always discarded, thus bringing us back to an effectively odd number of players. When \( n = 4 \), however, this technique does not work because it reduces us to a three-player world.

**Theorem 5.** For any \( n \geq 5 \), there exists a choice function that is

- **strategy-proof on the single-peaked preference domain and unrestricted ballot domain and**

- **truthfully implementable in Nash equilibrium.**

**Proof:** Fix \( n \geq 5 \) and let \( Q \) represent any ordering of the \( k \) alternatives in \( X \), with \( Q \) ordering the \( k \) alternatives \( x_1, \ldots, x_k \). Let \( \rho_{-L} \) be the ballot profile generated by dropping the ballots of all individuals in \( L \subset N \) and retaining all other ballots. Let \( C(\rho) \in 2^X \) be the collection of majority core alternatives at profile \( \rho \), so that \( C(\rho) = \{ x \in X : \text{ for all } y \in X, |i : x P_i y| \geq \frac{n}{2} \} \). Consider the following “Drop 2” choice function on \( \rho \in \mathcal{P}^n \):

1. If \( n \) is odd and \( \rho \) satisfies both worst-restriction and \( \alpha \)-restriction, then \( \phi(\rho) = C(\rho) \).

2. If \( n \) is even and \( \rho \) satisfies both worst-restriction and \( \alpha \)-restriction, then \( \phi(\rho) = C(\rho_{-\{1\}}) \).

3. If \( \rho \) violates one of these conditions, then for each \( i = 1, \ldots, n \) test whether \( \rho_{-\{i\}} \) satisfies both worst-restriction and \( \alpha \)-restriction. If so, let \( i \in W \).

   (a) If \( |W| < n \) and \( n \) is odd, then \( \phi(\rho) = \arg\min_{x_j \in C(\rho_{-W})} \{ j \} \).

   (b) If \( |W| < n \) and \( n \) is even, then \( \phi(\rho) = \arg\min_{x_j \in C(\rho_{-W\cup\{1\}})} \{ j \} \).

   (c) If \( |W| = n \) then \( \phi(\rho) = x_1 \).
We will consider two cases: the case in which a ballot profile is single-peaked, and the case in which it is not. Let \( \rho^* = (P_1, \ldots, P_n) \in S \) be a sincere preference profile, with \( \phi(\rho^*) = x^* \), and throughout, let \( \succ \) denote the strict majority preference relation induced by a profile \( \rho \).

First, suppose that \( i \) can profitably manipulate \( \phi \) at profile \( \rho^* \) with ballot \( P'_i \), and that \( (P'_i, \rho^*_{-i}) \) is single-peaked. Assuming that \( n \) is odd, this implies that \( x = C(P'_i, \rho^*_{-i}) \neq C(\rho^*) \), and that \( xP_ix^* \). However, we know that \( x^* \succ(\rho^*_{-i})x \), contradicting the fact that \( x \) is the core of \( (P'_i, \rho^*_{-i}) \). If \( n \) is even, this same logic holds, with \( p^* \) replaced by \( p^*_{-\{1\}} \).

Second, suppose that \( (P'_i, \rho^*_{-i}) \) is not single-peaked. Since \( i \) is the sole person submitting an insincere ballot, it follows that \( i \in W \). It may also be the case that there exists one other \( j \in W \), however if \( \rho^* \in S \) then \( |W| \leq 2 \). First, suppose \( n \) is odd. Then for this manipulation to be profitable for \( i \), we know that \( x = \phi(P'_i, \rho^*_{-i}) \in C(\rho_{-W}) \), and that \( xP_ix^* \). Let \( W = \{i\} \) or \( W = \{i, j\} \). Because \( x^* \succ x \), then \( x^* \succ x \) also, and regardless of whether \( W \) contains one or two individuals. Either one or two supporters of \( x \) are removed and so \( x^* \) is still majority-preferred to \( x \), or a supporter of \( x \) and a supporter of \( x^* \) are removed, thus canceling each other out. In either case, \( x^* \succ x \), contradicting \( x \in C(\rho_{-W}) \). For \( n \) even this same argument holds, replacing \( \rho^* \) with \( \rho^*_{-\{1\}} \) as appropriate.

Finally, having shown that this choice function is strategy-proof is equivalent to having shown that everyone submitting a sincere ballot, or \( \sigma_i^* = P_i \) for all \( i \), is a Nash equilibrium of the game \( (\mathcal{P}^n, \phi, \rho) \), for all \( \rho \in S \). Thus, \( \phi \) is truthfully implementable in Nash equilibrium. \( \square \)

### 3 Examples of Manipulation in One Dimension

In this section, we consider two applications that illustrate the broader implications of our findings. The first example, discussed in Section 3.1, demonstrates a feature of our results that may not be apparent from the theoretical sections of this paper, and that is important to note. As discussed in

---

\[\text{18} \text{This is because with } n \geq 5, \text{ at least two people will rank any element of a triple last; thus worst-restriction could be violated by at most two ballots. Furthermore, } \alpha-\text{restriction can only be violated by a pair of individuals, and the true manipulator will be in any such pair. Thus, } \alpha-\text{restriction can be violated by at most two individuals.}\]
the introduction, domain restrictions pose different challenges when considering Arrow’s theorem versus G-S. While coalitional strategy-proofness implies dictatorship of choice functions on $S$, IIA and weak Pareto only imply neutrality of preference aggregation rules on $S$. Section 3.1 illustrates how this distinction plays out in binary voting procedures by providing an example of a generally non-neutral institution that happens to be neutral on $S$, that satisfies IIA and weak Pareto on $S$, and that is also manipulable on $S$.

Then, in Section 3.2, we present a simplification of the argument we use in our main theorem, Theorem 3. It demonstrates that the incentive to manipulate will manifest itself as an individual or coalition either attempting to make the policy space appear multidimensional, or attempting to manipulate the true dimension of conflict. We then discuss the implications of the argument for recent work on deliberative democracy.

### 3.1 Amendment Agendas

In this section we briefly consider an institution that has received much attention in formal models of politics: the amendment agenda. Under an amendment agenda, alternatives are voted upon in an ordered sequence of pairwise votes. It is well-known that in the absence of a Condorcet winner, these agenda procedures are highly manipulable, with any alternative in the top cycle being attainable as a policy outcome depending on the sequence of votes taken. Thus, amendment agendas are, in general, not neutral, because they privilege alternatives appearing later on in the agenda.\(^{19}\) It is also well-known that in the presence of a Condorcet winner, any sequence of voting will yield the Condorcet winner as an outcome, regardless of whether all individuals vote sincerely or all vote sophisticatedly.\(^{20}\) Thus, over the domain $S$, amendment agendas are neutral: if the collection of ballots an amendment agenda is given is single-peaked, then so is any relabeling of that collection, and the outcome of voting will be the Condorcet winner (and the relabeled

\(^{19}\)The last alternative considered in a pairwise vote need only defeat the winning alternative that preceded it in order to become a policy outcome. However, an alternative considered first must defeat every other policy in order to be chosen as a policy outcome.

\(^{20}\)Sophisticated voting in this context refers to individuals playing subgame-perfect Nash equilibrium strategies.
Condorcet winner) of each ballot profile.

What is perhaps less well-known is that amendment agendas are highly manipulable at sincere profiles of ballots, even in the presence of a Condorcet winner. In this section we consider amendment agendas to be choice functions, $\phi_A$, in which an individual’s ballot dictates how that individual will vote on any pair of alternatives.

In the presence of a Condorcet winner, assuming that all individuals vote either sincerely or sophisticatedly yields the same outcome. However, the sequence of votes that individuals cast will differ, and at a sincere profile of ballots, or sequence of votes, an amendment agenda is manipulable. To see this, consider the amendment agenda pictured in Figure 1, in which alternatives $x$ and $y$ are first put to a vote via majority rule, and the winner is then pitted against $z$ in order to determine the final outcome. Suppose that there are three individuals with the following preferences: $xP_1yP_1z$, $yP_2zP_2x$, and $zP_3yP_3x$. This preference profile is single peaked, and is pictured graphically in Figure 2; it yields $y$ as a Condorcet winner.

![Figure 1: A two-stage amendment agenda](image)

Under a truthful collection of ballots, the amendment agenda pictured in Figure 1 yields $y$ as an outcome: $y$ defeats $x$ at the first stage of voting by the votes of Players 2 and 3, and $y$ defeats $z$ at the second stage by Players 1 and 2. Now consider a collection of ballots in which Players 1 and 2 truthfully reveal their preferences over alternatives, but Player 3 claims to have the preference ordering $zP'_3xP'_3y$. Under this ballot profile our amendment agenda now yields $z$ as the winner:

---

21 Others have noted that when the behavioral assumption of sincerity or sophistication is not uniformly made across all voters, amendment agendas may no longer be Condorcet consistent. See Denzau et al. (1985) and Austen-Smith (1987), among others.
Figure 2: A single-peaked preference profile

$x$ defeats $y$ at Stage 1 by the (sincere) vote of Player 1 and the (insincere) vote of Player 3, and $z$ defeats $x$ at Stage 2 by the sincere votes of both Players 2 and 3. Furthermore, this is a beneficial manipulation by Player 3, as it enables him to attain his ideal point as the policy outcome.

Clearly the insincere ballot of Player 3 is not single-peaked with respect to the underlying ordering of alternatives. However, without \textit{a priori} restricting how people can cast votes, manipulation is endemic to this form of agenda, even when the majority will is clearly well-defined. And we know of no real-world institution that restricts how pairwise votes may be cast. At the same time, when handed a truthful profile of ballots, the amendment agenda produces outcomes and sequences of votes that are consistent with pairwise majority voting. Thus, satisfying Arrow’s conditions does not, “by easy implication” imply satisfaction of the conditions of Gibbard-Satterthwaite on single-peaked domains, as claimed by Dryzek and List (2003). Pairwise majority voting is transitive, weakly Paretian and IIA (and thus, neutral) when a collection of preferences is single-peaked, and produces outcomes consistent with those produced by an amendment agenda “choice function” $\phi_A$. Amendment agendas, and in general any choice function representing the outcome of a binary voting process, are not strategy-proof.

3.2 Deliberative Democracy

Deliberation within democratic governance has attracted a great deal of attention in the past two decades. This attention has focused primarily on the potential impact of deliberation on the quality
of democratic decision-making, broadly construed. For example, deliberative democracy has been forwarded as a means of divining the “best” choice (Estlund et al. (1989), Estlund (1997)), as a legitimating device (e.g., Manin (1987), Cohen (1989)), and as a means by which individuals’ preferences may be brought more in line with each other (e.g., Miller (1992)). In addition, some have argued that deliberation is in and of itself a good thing (e.g., Shapiro (2002)).

Dryzek and List (2003) have recently extended such arguments and explored the linkages between deliberative democracy and social choice theory. In many respects, their arguments provide hope for deliberative democrats in spite of the generally negative conclusions about the coexistence of collective rationality and democratic collective choice. The heart of their argument is that deliberative decision-making may allow individuals within a group to leverage the underlying common structure of their individual preferences to choose an outcome that satisfies undesirable normative (e.g., democratic) properties. The principle example of such a structure is single-peakedness. Dryzek and List link single-peakedness with agreement at a meta-level, a notion loosely describing the agreement by participants “on a common dimension in terms of which the alternatives are to be conceptualized.”

A key conclusion for Dryzek and List’s purposes is that preference structuration can eliminate the incentive to for a deliberator to misrepresent his or her preferences at the point at which the collective is faced with making a final decision. Claims that structuration may be produced through deliberation have been forwarded by many scholars and some empirical evidence supports this claim. Our arguments in this paper provide some insight into the conditions that one must impose on the structuration process to guarantee that it offers no benefit from misrepresentation, and show that these conditions are very restrictive. Indeed, the conditions are far more restrictive than has been claimed elsewhere. Specifically, we have shown that a nondictatorial decision-making

---

22Knight and Johnson (1994) present a critique of the parallels drawn (particularly as argued by Miller (1992)) between preference aggregation and deliberation.
24Theories discussing the production of structuration through deliberation include Mansbridge (1983), Goodin (1986), and Miller (1992). Empirical studies of the emergence of structuration have been offered by Radcliff (1993) and Farrar et al. (2006).
25In addition to Dryzek and List (2003), our results call into question some of the relevant claims of Grofman and Feld (1988), Miller (1992), and Mackie (2003), among others.
institution can be ensured to offer no benefit from misrepresentation by a person or coalition only if the details of the structure of the individuals’ preferences are written into the rules of the institution itself (i.e. the institution can utilize the underlying ordering of alternatives). This is akin to saying that if we wish to design a deliberative institution that will give individuals the incentive to participate in a process of collective preference structuration, the institution must have a priori knowledge of the true structure of preferences; it must know the true left-right ordering of the alternative space before the deliberative process begins.

The following table illustrates exactly why individuals with single-peaked preferences may not want the true “dimension of conflict” revealed. It depicts three different preference profiles that are all single-peaked. Because each profile is single-peaked and there are three individuals, each profile yields a unique Condorcet winner, or policy that is the median voter’s ideal point. Consider a choice function that yields a Condorcet winner whenever a Condorcet winner exists (presumably, the benefit of preference structuration is the existence of such a policy). Thus, for \( \rho_1 \) it yields \( z \) as the outcome, and so on.

<table>
<thead>
<tr>
<th>Person 1</th>
<th>( \rho_1 )</th>
<th>Person 2</th>
<th>( \rho_2 )</th>
<th>Person 3</th>
<th>( \rho_3 )</th>
<th>Cycle</th>
</tr>
</thead>
<tbody>
<tr>
<td>x ( \succ ) z ( \succ ) y</td>
<td>x ( \succ ) y ( \succ ) z</td>
<td>x ( \succ ) y ( \succ ) z</td>
<td>y ( \succ ) z ( \succ ) x</td>
<td>y ( \succ ) z ( \succ ) x</td>
<td>z ( \succ ) x ( \succ ) y</td>
<td></td>
</tr>
<tr>
<td>x ( \succ ) y ( \succ ) z</td>
<td>y ( \succ ) x ( \succ ) z</td>
<td>y ( \succ ) z ( \succ ) x</td>
<td>z ( \succ ) y ( \succ ) x</td>
<td>z ( \succ ) x ( \succ ) y</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What does this choice function yield when it receives a collection of ballots that does not admit a Condorcet winner? The table shows that regardless of which outcome the choice function chooses, the function is always manipulable by someone at a single-peaked profile of preferences. For example, if the choice function chooses \( x \) as the outcome when it receives a cyclic profile of ballots, then the function is manipulable by Person 1 at the (sincere) profile \( \rho_1 \). In particular, Person

\[
\begin{array}{|c|c|c|c|}
\hline
\text{Outcome} & \text{Median} = z & \text{Median} = x & \text{Median} = y \\
\hline
\text{x: 1 manipulates } \rho_1 & \text{y: 2 manipulates } \rho_2 & \text{z: 3 manipulates } \rho_3 \\
\hline
\end{array}
\]
1 has an incentive to report that his preferences are $x \succ y \succ z$, when they are truly $x \succ z \succ y$. By misrepresenting his preferences this way, he switches the outcome from $z$ (the median voter’s ideal point) to $x$ (his own ideal point).

In each case, for manipulation to occur an individual must submit a ballot that is not single-peaked with respect to the true ordering of alternatives. In this sense, if individuals are not constrained to only submitting ballots that are single-peaked with respect to precisely the ordering that the deliberative process was intending to uncover, we cannot guarantee that an individual will have no incentive to lie about what he believes the true dimension of conflict to be, or to claim that the policy space is two dimensional, or more complex than it truly is. In other words, even if we begin with a situation in which preferences are already structured in a desirable way, we cannot design an institution that is guaranteed to truthfully elicit those preferences.

4 Conclusions

Theorems 2 and 3 imply that one must be careful in interpreting collective will in any real-world policymaking institution even when preferences are presumed to be single-peaked. This point is highly relevant for those scholars who insist that majority rule cycles are infrequent or untroubling (e.g., Mackie (2003)). We show that even if preferences do not admit cycles, any non-dictatorial institution provides groups with the ability to profit by claiming that preferences are cyclic. Thus, appeals to aggregate outcomes as indicators of collective will are not necessarily well-founded even when the majority will is assumed to exist. Accordingly, the normative and descriptive issues raised by Arrow’s theorem and the Gibbard-Satterthwaite theorem are more than “logical exercises,” “abstract limit cases” and “mathematical curiosities” dreamed up for the purpose of scholarly debate, as claimed by some.26 Single-peakedness does not solve the problems raised by Arrow’s Theorem in the real-world because policymaking institutions are generally not neutral. And single-peakedness does not eliminate the possibility of gains through strategic manipulation

---

26See Mackie (2003), pp. 156, 192.
within real-world institutions, because few (if any) policymaking institutions are dictatorial, and few (if any) policymaking institutions limit the preferences that individuals and groups can *claim* to have.

**References**


