Strategic Manipulation in Representative Institutions

(Incomplete draft) [*]

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

November 15, 2009

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. Accordingly, insincere behavior is inherent to the vast majority of real-world lawmaking systems, even when the policy space is unidimensional and the core is nonempty. We conclude with a discussion of the implications for claims about the political representation of interests within real-world political institutions.

[*]We are grateful for the helpful suggestions of Tom Schwartz, Ken Shepsle, and seminar participants at Princeton University and the University of Wisconsin—Madison.

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

[‡]Assistant 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: jwpatty@gmail.com.

1 Collective Rationality and Neutrality

The following theorem — the Gibbard-Satterthwaite theorem (henceforth, *G-S*) — is central to the study of both collective choice institutions themselves and individual behavior within such institutions.

Theorem 1. *(Gibbard, 1973), (Satterthwaite, 1975). If there are three or more feasible policy alternatives and each individual may have any preference over the alternatives, then the only collective choice institution in which no individual ever has an incentive to misrepresent his or her true preferences is a dictatorship.*

Like Arrow's Possibility Theorem, G-S depends heavily on the assumption of *unrestricted domain*, which states that all individuals may rank all alternatives however they want. In particular, both theorems leverage the fact that, in the presence of a cyclic profile of preferences, any responsive institution that produces a transitive ranking over alternatives must on occasion be called upon to break a tie in favor of a minority. Many scholars have called into question the empirical relevance of both G-S and Arrow's theorem precisely because of the unrestricted domain assumption. Mackie provides perhaps the most comprehensive case for the non-existence of majority-preference cycles in real politics, arguing that virtually every published empirical claim of a majority-preference cycle has been made in error (Mackie, 2003). Mackie claims that voter preferences in general are single-peaked over any domain of issues under consideration at a given time. The validity of this claim is not an issue we tackle in this paper. Suffice it to say that many issue spaces in politics and economics can be naturally interpreted as being one-dimensional, and it is not a stretch to assume that preferences are single-peaked over these spaces. Furthermore, it is well-known that when voter preferences *are* single-peaked with respect to a fixed ordering over alternatives, there exists a non-dictatorial institution in which every individual always has an incentive to truthfully represent his or her preferences – namely, pairwise majority rule.

Theorem 2. *(Black, 1948).*[1] *If individual preferences and ballots are single-peaked with respect to a fixed ordering, then pairwise majority rule is a non-dictatorial collective choice institution with full range in which no individual ever has an incentive to misrepresent his or her preferences.*

And so it would appear that if an issue space admits single-peaked preferences, then we have solved the question of whether preferences have been represented accurately. Or have we? What, exactly, does the existence of a well-defined majority will and nonempty core get us? Does it, for example, imply that reasonably democratic institutions will produce the core as a policy outcome? Does it imply that that individuals have an incentive to behave sincerely? As the following statements demonstrate, there is a large literature in political science that argues that single-peakedness implies a great deal:

> "When all individuals have single-peaked preference orderings the process of collective decision making is dramatically simplified." (Feld and Grofman, 1988, p. 776)
>
> "[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, 2002a)
>
> "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

[1] Black's Median Voter Theorem specifically states that in the presence of single-peaked preferences, a Condorcet winner exists. That a Condorcet winner arises as the outcome of pairwise majority rule over single-peaked ballots, and that majority rule is strategy-proof in this case, was proved early on by Dummett and Farquharson (1961) and Pattanaik (1976).

voting to succeed." (Mackie, 2003, p. 161)

However, there is an important assumption implicit in our statement of Black's Theorem that is necessary in order for single-peakedness to be the "escape route" from G-S that Dryzek and List claim it to be. It requires that individuals may only vote over issues 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 a strategic individual may not cast a vote for x over z and a vote for z over y. Black's Theorem then tells us that majority rule is strategy-proof provided that individuals are not allowed to lie in particular ways.[2]

In this paper we consider incentives for the strategic manipulation of institutions when individuals have preferences that are single-peaked, but are allowed to *claim* to have any preference ordering that they wish. In other words, if the true ordering of alternatives on the left-right spectrum is $x < y < z$, an individual or collection of individuals may claim to prefer z to x to y. We show that in such situations, inducing individuals to truthfully reveal their preferences is not automatic *even when preferences are single-peaked and there exists a clear majority rule core.* In particular, inducing truthful revelation at every single-peaked profile of preferences requires the use of a dictatorial collective choice procedure – a procedure that always grants decision-making authority to a unique individual. Formally, we show that any institution that is not dictatorial necessarily is not *coalitionally strategy-proof*: there exists some situation in which a person or a collection of people have an incentive to misrepresent their preferences in pursuit of a different collective choice. More importantly, this conclusion holds even in the absence of majority-preference cycles, and in

[2]This point was noted by Blin and Satterthwaite (1976), who show that majority rule with Borda completion 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.

the presence of a Condorcet winner.

1.1 Why Manipulation is Important

The question of manipulation – the misrepresentation of one's true preferences for individual gain – has been examined by social scientists for over two centuries.[3] Ignoring the possibly unsettling mendaciousness inherent in manipulative behavior, manipulation is generally problematic for any problem relating to inference. 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 committee reflect the preferences of a group's members? Does a juror's vote to convict a defendant truthfully reflect the juror's beliefs about the defendant's guilt or innocence? Does an executive's best choice of political appointees necessarily pursue the policies that the executive would individually pursue in those positions? The inference problem is also highly relevant for mechanism design issues, such as how a social planner might go about choosing the most efficient form of regulation, when the regulated firms have private information about their cost structures.

In a nutshell, we argue that the possibility of manipulation is endemic to virtually all political institutions, even in settings where it has been considered unimportant, and regardless of the attractiveness of certain policies under consideration. This is not to say that manipulation is a bad thing; in fact, some have convincingly argued that the opposite is actually true.[4] For example, Miller (1977) shows that when all individuals vote strategically, outcomes may be obtained that are Pareto superior to outcomes obtained if

[3] In his 1788 treatise "On the Constitution and the Functions of Provincial Assemblies," Condorcet argued that Borda's scoring method was highly prone to strategic manipulation by voters (Young, 1995).

[4] See also Dowding and Van Hees (2007), in which the authors distinguish between "sincere" and "non-sincere" forms of manipulation.

all individuals vote truthfully. Our argument is simply that manipulation may be problematic for those of us wishing to study the relationship between individual goals and behaviors, even in settings in which a "majority will" is well-defined.

In this respect, and before continuing on to the theory and results, it is useful to compare problems of manipulation to the "Arrovian problem" of the conditions under which one can define an unambiguous notion of "social will." This contrast is relevant because, as mentioned above, real-world institutions are subject to manipulation even if one assumes that there exists a well-defined notion of the social will. We state and discuss Arrow's result below.

Theorem 3. *(Arrow, 1951). If there are three or more alternatives and at least two individuals, each of whom may have any preference over the alternatives, then the only Pareto efficient and transitive preference aggregation rule that satisfies independence of irrelevant alternatives is dictatorial.*

G-S does not preclude any profile of individual preferences, a direct analogue of Arrow's requirement of *unrestricted domain*. Arrow's result, which establishes that the notion of a well- (and universally-) defined "social will" is incompatible with very minimal democratic principles, is closely related to G-S and it has been demonstrated that both theorems can be obtained from what is essentially a single proof (Reny, 2000). However, the conclusions of G-S differ from those of Arrow's possibility theorem in several important ways. As opposed to the acceptance of social indifference within the Arrovian framework, G-S requires that a *unique* collective choice be generated (*i.e.*, G-S deals with "revealed collective preference"). In practical terms, this requirement is equivalent to assuming that any tie for the most-preferred choice must be broken. This requirement implies that G-S is more "institutionally constructive" than Arrow: while Arrow's result maps preference profiles into social preferences, G-S requires that the social preference generate a non-manipulable choice. In other words, G-S not only requires that ties be broken – it requires that they be broken in a way that does not strictly reward insincere behavior by the voters.

There is another important difference between Arrow's theorem and G-S that is particularly relevant when considering preference domain restrictions. In Arrow's theorem, institutions consider only sincere profiles of preferences; in the G-S theorem, they do not. Thus, a preference domain restriction (e.g. single-peakedness) is a stronger restriction in Arrow's theorem than under the G-S theorem, unless an institution under consideration specifically requires *ballots* to be drawn from the same domain as *preferences*. The differential effect of preference domain restrictions in the two theorems plays out in our results. Our proof that coalitional non-manipulability requires dictatorship on single-peaked preference domains leans heavily on a related, but normatively weaker, result that we have proved for Arrow's theorem (Gailmard et al., 2008): Any weakly Paretian preference aggregation rule that is independent of irrelevant alternatives must be *neutral*, even when preferences are known to be single-peaked. In other words, even in instances in which there is a well-defined, transitive majority preference relation, neutrality is required for collective choice to be simultaneously transitive, weakly Paretian and independent of irrelevant alternatives. In our extension of this result to a weakened version of G-S we can replace neutrality with dictatorship precisely because of the fact that the single-peaked preference domain restriction is weakened by assuming that institutions are required to take *all* ballot profiles as an input.

The key to both results 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. Thus, single-peakedness in and of itself is not "enough" information to enable a non-neutral institution to produce an IIA and weakly Paretian social ranking over alternatives, and it is not enough information to enable a non-dictatorial institution to be coalitionally strategy-proof. Given the generality and power of Arrow's Theorem and the Gibbard-Satterthwaite Theorem, the novelty of our results lies in their application to a canonical setting for models of political institutions: the unidimensional spatial model.[5] Furthermore, the unidimensional spatial model is often invoked precisely because it is considered to be immune from the conclusions of Arrow and G-S. In this paper we argue that these conclusions can be interpreted as institutional problems, rather than problems stemming from underlying majority preference cycles. In other words, violating conditions such as Pareto efficiency, independence of irrelevant alternatives, transitivity of collective choice, and strategy-proofness is an irresolvable consequence of the vast majority of collective decision-making procedures in use in the world, and poses problems that are separate from the underlying structure of preferences.

The following section defines the theoretical framework that we utilize in both results. Section 3 proves our main results: that coalitional strategy-proofness on the domain of single-peaked preferences first implies neutrality, and then dictatorship. Section 4 presents several examples of the applicability of our results to real-world institutions, including binary voting procedures. This section also briefly discusses several normative implications of our results, including questions of optimal delegation and the desirability of Pareto efficiency in certain decision-making procedures. In Section 5, we conclude and offer a brief discussion of the connections between our results and the analysis of representative institutions in general.

2 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_iy implies that i strictly prefers x to y. If $x_i^*P_iy$

[5] A smattering of examples to justify the term "canonical" might include Downs (1957), Davis et al. (1970), McCubbins et al. (1994), Poole and Rosenthal (1997), and Krehbiel (1998) among, of course, many others.

for all $y \neq x_i^*$, then x_i^* is referred to as i's *most-preferred policy* or *ideal point*. Let $x_i^*(\rho)$ be i's ideal point under profile ρ.

Throughout, $\rho = (P_1, ..., P_n)$ denotes an n-dimensional preference profile describing the preferences of all individuals: the notation $\mathcal{P}^n$ represents the collection of all n-dimensional profiles of strict orders on X. Any nonempty set $\mathcal{D} \subseteq \mathcal{P}^n$ is referred to as a *preference domain*, and with strict inclusion, $\mathcal{D}$ is referred to as a *restricted domain*. We will come back to restricted domains in more detail in Section 2.3. For any preference profile $\rho \in \mathcal{P}^n$, $\rho|_S$ denotes the restriction of ρ to the set of alternatives $S \subseteq X$. Similarly, for any individual preference $P \in \mathcal{P}$, $P_i|_S$ 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 $P(x, y; \rho) \equiv \{i \in N : xP_iy\}$ denotes the set of individuals who strictly prefer x to y under ρ.

2.1 Preference Aggregation Rules

Letting $\mathcal{R}$ be the collection of weak orders over X, a *preference aggregation rule* is any function, $F : \mathcal{D} \rightarrow \mathcal{R}$, that maps a strict preference profile from domain $\mathcal{D}$ into a weak order over X. The notation $xR_F(\rho)y$ denotes weak social preference under F at profile $\rho \in \mathcal{P}^n$ and $xP_F(\rho)y$ denotes strict social preference. The following definitions characterize several properties of preference aggregation rules.

Definition 1 (Weakly Paretian)**.** *A preference aggregation rule F is* weakly Paretian *if for all $\rho \in \mathcal{D}$ and all $(x, y) \in X^2$,*

$$P(x, y; \rho) = N \Rightarrow xP_F(\rho)y.$$

Definition 2 (Neutral)**.** *A preference aggregation rule F is* neutral *if for every $x, y, a, b \in X$, and every $\rho, \rho' \in \mathcal{D}$,*

$$P(x, y; \rho) = P(a, b; \rho') \textit{ implies } xR_F(\rho)y \Leftrightarrow aR_F(\rho')b.$$

Definition 3 (Independent of Irrelevant Alternatives (IIA))**.** *A preference aggregation rule F is* independent of irrelevant alternatives *(IIA) if, for all $(x, y) \in X^2$ and all $\rho, \rho' \in \mathcal{D}$,*

$$\rho|_{\{x,y\}} = \rho'|_{\{x,y\}} \Rightarrow F(\rho)|_{\{x,y\}} = F(\rho')|_{\{x,y\}}.$$

2.2 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 X. Throughout, we will assume that ϕ has *full range*: for any $x \in X$, there exists a $\rho \in \mathcal{P}^n$ such that $\phi(\rho) = x$. Thus, a preference aggregation rule produces an ordering over the elements of X while a choice function simply produces a single outcome. The notation $\phi(\rho) = x$ says that choice function ϕ produces outcome x at profile ρ.

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 $\mathcal{P}^n$.[6] In other words, while true individual preferences may come from a restricted set of orderings, individual behavior is only required to be individually rational, in the sense of being rationalizable by a transitive binary relation. Throughout, we will use the notation (P_i', ρ_{-i}) to denote a ballot profile in which i submits ballot P_i', and all others submit ballots as in profile ρ. More generally, the notation (ρ_L', ρ_{-L}) denotes a ballot profile in which all members $i \in L \subseteq N$ submit ballots as under ρ', and all individuals not in L submit ballots as under ρ.

[6] In Section **??** we discuss the implications of restricting the ballot domain for a specific class of institutions with an agenda setter.

The following definitions characterize several properties of collective choice functions.

Definition 1′ (Weakly Paretian)**.** *A collective choice function ϕ is* weakly Paretian *if for all $\rho \in \mathcal{D}$ and all $(x, y) \in X^2$,*

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

Definition 4 (Monotonic)**.** *A collective choice function ϕ is* monotonic *if, for all $(x, y) \in X^2$ and all $\rho, \rho' \in \mathcal{D}$,*

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

Definition 5 (Dictatorial)**.** *A collective choice function ϕ is* dictatorial *if for some $i \in N$ and for all $\rho \in \mathcal{D}$,*

$$\phi(\rho) = x_i^*(\rho),$$

where $x_i^*(\rho)$ is i's ideal point under profile ρ.

Definition 6 (Strategy-Proof (SP))**.** *A collective choice function ϕ is* manipulable *if, for some $\rho = (P_1, ..., P_n) \in \mathcal{D}$ and $i \in N$ there exists a $P_i' \in \mathcal{P}$ such that*

$$\phi(P_i', \rho_{-i}) P_i \phi(\rho).$$

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

Definition 7 (Coalitionally Strategy-Proof (CSP))**.** *A collective choice function ϕ is* coalitionally manipulable *if, for some $\rho = (P_1, ..., P_n) \in \mathcal{D}$ and $L \subseteq N$ there exists a $\rho' \in \mathcal{P}^n$ such that*

$$\phi(\rho_L', \rho_{-L}) P_i \phi(\rho) \text{ for all } i \in L.$$

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

Note that individually manipulable social choice functions are coalitionally manipulable (by a coalition of one). However, the converse need not be true; there may be instances in which a social choice function may only be manipulable by a sufficiently large coalition. When $\mathcal{D} = \mathcal{P}^n$, then individual non-manipulability implies coalitional non-manipulability, because it implies dictatorship. Thus, in a setting with unrestricted domain, individual and coalitional manipulability and non-manipulability are equivalent.

2.3 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 the existence, and more about the characterization, of such preference aggregation rules and choice functions on this restricted domain.

Single-Peaked Preferences. The domain of *single-peaked preferences* is the set of all profiles of preferences such that there exists a function $Q : X \to \{1, 2, \ldots, K\}$ such that Q is a bijection and every individual's preferences are consistent with a quasi-concave utility function of $\{Q(x) : x \in X\}$. We denote the single-peaked preference domain by $\mathcal{S}^n \subset \mathcal{P}^n$. We will at times refer to the ordering that profile $\rho \in \mathcal{S}^n$ is single-peaked with respect to as Q_ρ. When referring to this ordering, if alternative x is above y with respect to Q we write $x >_Q y$.

While this preference restriction is widely utilized and intuitively quite simple, Ballester and Haeringer (2007) prove that the set $\mathcal{S}^n$ is completely characterized by two conditions, worst-restriction[7] and α-restriction, both defined below.

[7] See Sen (1966) and Sen and Pattanaik (1969) for a more thorough discussion of worst-restriction.

Definition 8 (Worst-restriction)**.** *A profile* ρ *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* $i \in N$*,* $a \succ_i b$ *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 9 (α-Restriction)**.** *A preference profile* ρ *is* α-restricted *if there do not exist two agents,* $i, j \in N$*, and four alternatives* w, x, y*, and* z *such that*

1. *The preferences over* w, x*, and* z *are opposite:* wP_ixP_iz *and* zP_jxP_jw*.*

2. *The players agree about the ranking of* y *and* x*:* yP_ix *and* yP_jx*.*

Definition 10 (Single-Peakedness)**.** *A preference profile is* single-peaked *if and only if it satisfies worst-restriction and* $\alpha-$*restriction (Ballester and Haeringer, 2007).*

It is important to note at this point that the domain $\mathcal{S}^n$ is the set of *all* single peaked preference profiles. In other words, in *a priori* terms, any ordering of the alternatives is possible.[8]

Ubeda (2003) has recently used a different domain restriction, the 2-free triple domain ($\mathcal{T}_2^n$),[9] to demonstrate that on any domain satisfying this restriction, weak Pareto and IIA imply neutrality, a conclusion that

[8] 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 $\mathcal{Q}$, this set is denoted by $\mathcal{S}_Q$, and the set of all profiles of such preferences is denoted by $\mathcal{S}_Q^n$. This space is widely discussed in the political economy literature. For a succinct and lucid overview of the power of the assumption that Q is known *a priori*, see Chapter 2.4 of Austen-Smith and Banks (2004), in particular Theorem 2.4.

[9] 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 α-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.

mirrors our own (Theorem 4, below). The key distinction between Ubeda's result and Theorem 4 is that the 2-free triple domain and the single-peaked domain are not nested. Specifically, for all $n \geq 2$, $\mathcal{T}_2^n \nsubseteq \mathcal{S}^n$ and $\mathcal{S}^n \nsubseteq \mathcal{T}_2^n$. In other words, satisfaction of either the 2-free triple restriction or single-peakedness does not imply satisfaction of the other. With these preliminaries in hand, we are now in a position to state and prove our main results.

3 Stability and Coalitional Strategy-Proofness on Single-Peaked Domains

To prove that coalitional strategy-proofness implies dictatorship on domain $\mathcal{D} = \mathcal{S}^n$, we utilize the following lemmas and theorem.

Lemma 1. *Let ϕ be a coalitionally strategy-proof collective choice function. If $\mathcal{D} = \mathcal{S}^n$, then ϕ is weakly Paretian.*

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

Lemma 2. *Let ϕ be a coalitionally strategy-proof collective choice function. If $\mathcal{D} = \mathcal{S}^n$, then ϕ is monotonic.*

Proof: Suppose that ρ and ρ' 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 ϕ is not CSP on $\mathcal{S}^n$. 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 ρ so that for all i such that $x\overline{P}_iy$ under ρ, every $\overline{P}_i$ is replaced by an identical new ordering P_i with x top-ranked and y as high in i's ranking as possible while maintaining P_i

single-peaked with respect to Q_ρ. It is easy to verify that such an ordering exists. Similarly, change ρ' so that for all j with $y\overline{P}'_j x$ under ρ', y is top-ranked with respect to the new P'_j, and x is as high as possible while maintaining that P'_j be 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 ρ and ρ') $\phi(\rho) = x$ and $\phi(\rho') = y$.

Consider the triple, $x, y, z \in X$, with z arbitrary. $\rho, \rho' \in \mathcal{S}^n$ imply that for each profile, one of only two elements of this triple may be lowest-ranked by any individual. If at profile ρ these two elements are a, b, we say that (a, b) is lowest-ranked for $\rho|_{\{a,b,z\}}$. Note that in constructing ρ above, we have ensured that the only instances in which zP_iy for $i \in A$ are those in which z lies between x and y under the ordering Q_ρ (i.e. $x >_{Q_\rho} z >_{Q_\rho} > y$, or the reverse). Similarly, zP'_jx for $j \in B$ implies that z lies between x and y according to $Q_{\rho'}$.

Construct a new profile $\rho^* \in \mathcal{P}^n$ in which $P^*_i = P_i$ for $i \in A$, and $P^*_j = P'_j$ for $j \in B$. Note that rankings are unspecified for $k \notin A \cup B$, if such individuals exist. These individuals' transitive rankings can be arbitrarily assigned. Also note that ρ^* 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 ρ with ρ^*, or $N \setminus B$ could manipulate ρ' with ρ^*. Thus, $\phi(\rho^*) = z$.

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 $k \notin A$, $yP_kx \Rightarrow zP_kx$. Thus, ϕ is manipulable at ρ by coalition $N \setminus A$ submitting ballots as in ρ^*; these individuals can guarantee themselves the outcome z, which they all prefer to x. It follows that ϕ 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 ρ and

ρ' above, yP_iz for all $i \in A$ and xP'_jz for all $j \in B$.[10]

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_iyP_iz (resp. yP_jxP_jz). Note also that x, y, z need not appear consecutively in these individuals' rankings. Furthermore, ρ^* as constructed above still 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 ϕ by submitting ballots as in ρ^*.

We will now construct a new $\rho^o \in \mathcal{S}^n$ and show that ϕ CSP for ρ implies that ϕ not CSP for ρ^o. This will take several intermediate steps. First, consider the ordering over alternatives induced by the preferences of individuals in A under ρ; we will call this ordering Q_A, and we know that under this ordering $x >_{Q_A} y >_{Q_A} z$. Construct a new profile $\hat{\rho}$ where $\hat{P}_i = P_i$ 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 $j \notin A$, 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, $\hat{\rho} \in \mathcal{S}^n$.

We know that coalition $N \setminus A$ can submit ballots as in ρ^* 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 $i \in N$, $y\hat{P}_iz$. And by CSP of ρ, $\phi(\hat{\rho}) \neq y$, because then ρ would be manipulable by coalition $N \setminus A$ submitting ballots as in $\hat{\rho}$. It follows that $\phi(\hat{\rho}) = w$, with all members $j \in N \setminus A$ ranking $y\hat{P}_jw\hat{P}_jz\hat{P}_jx$. However, this implies that for all $i \in A$, $w\hat{P}_iy$, 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 $x >_{Q_A} w >_{Q_A} y >_{Q_A} z$.

Now consider a new profile $\hat{\rho}^1$ in which everyone in A ranks the alternatives as in $\hat{\rho}$. For all $j \in N \setminus A$,

[10]This immediately implies that ϕ does not satisfy Pareto efficiency on the full domain $\mathcal{P}^n$, because ρ^* can be constructed so as to have all individuals not in $A \cup B$ place x and y at the top of their ballots, and so $x, y > z$ on all ballots ρ^*_i and yet $\phi(\rho^*) = z$.

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 $x >_{Q_A} w >_{Q_A} y >_{Q_A} z$. Thus, $\hat{\rho}^1 \in \mathcal{S}^n$. Using the same logic as above, we get that $\phi(\hat{\rho}^1) \neq x$ or w (else $N \setminus A$ manipulates with ρ^* 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 ρ. Thus, $\phi(\hat{\rho}^1) = w^1$, with individuals $j \in N \setminus A$ ranking $y\hat{P}_j^1 w^1 \hat{P}_j^1 z\hat{P}_j^1 w\hat{P}_j^1 x$. Again, it follows that for all $i \in A$, $\hat{P}_i^1$ 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 $x >_{Q_A} \{w^1, w\} >_{Q_A} y >_{Q_A} z$ (the w^1, w ranking need not be specified in the proof, although by coalition $N \setminus A$'s preferences, the ranking must have $w >_{Q_A} w^1$).

Repeat the above steps for $k = 2, ..., |X| - 4$ by choosing a new $\hat{\rho}^k$ with $y\hat{P}_j^k z\hat{P}_j^k ...$ for all $j \in N \setminus A$, and $\hat{P}_i^k = P_i$ 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 $x >_{Q_A} \{w, w^1, w^2, ...w^{K-4}\} >_{Q_A} y >_{Q_A} z$. At this point, construct ρ^o so that P_j^o is such that $yP_j^o zP_j^o \{w, w^1, ..., w^{|X|-4}\} P_j^o x$ for all $j \in N \setminus A$ and $P_i^o = P_i$. Again, ρ^o is single-peaked with respect to Q_A. CSP requires $\phi(\rho^o) P_j^o z$ for all $j \in N \setminus A$, otherwise this coalition would manipulate with ballots as in ρ^*. CSP of ϕ at ρ requires $\phi(\rho^o) \neq y$. And these two statements imply a contradiction. Thus, ϕ is not CSP. It follows that ϕ CSP implies ϕ monotonic when $\mathcal{D} = \mathcal{S}^n$. □

Theorem 4. *Let F be a weakly Paretian and IIA preference aggregation rule. If $\mathcal{D} = \mathcal{S}^n$, then F is neutral.*

Proof: See Gailmard et al. (2008).

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

Definition 11 (Blocking coalition for (x, y))**.** *A coalition $L \subseteq N$ is a blocking coalition for (x, y) if for all*

$\rho = (P_1, ..., P_n) \in \mathcal{D}$ *such that* xP_iy *for all* $i \in L$ *and* yP_jx *for all* $j \notin L$*,* $\phi(\rho) \neq y$*.*

Definition 12 (Blocking coalition)**.** *A coalition* $L \subseteq N$ *is a blocking coalition if for all* $\rho = (P_1, ..., P_n) \in \mathcal{D}$ *and all pairs* $(a, b) \in X^2$*,* aP_ib *for all* $i \in L \Rightarrow \phi(\rho) \neq b$*.*

Lemma 3. *ϕ coalitionally strategy-proof implies that if there exists one* $\rho \in \mathcal{S}^n$ *with* xP_iy *for all* $i \in L$ *and* yP_jx *for all* $j \notin L$ *and* $\phi(\rho) = x$*, then* L *is a blocking coalition.*

Proof: Let ρ be such that xP_iy for all $i \in L$ and yP_jx 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 $\rho' \in \mathcal{S}^n$ with xP'_iy for all $i \in L$ and yP'_jx 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 $\rho \in \mathcal{S}^n$ where $i \in L$ rank the alternatives xP_iyP_ia, $j \notin L$ rank them yP_jaP_jx and all $k \in N$ have cP_kd when $c \in \{a, x, y\}$ and $d \notin \{a, x, y\}$. Then $\phi(\rho) = x$, by L blocking for (x, y) and by weak Pareto. Thus,L is blocking for (x, a).

Now construct a $\rho' \in \mathcal{S}^n$ with $P'_j = P_j$ for $j \notin L$ and with $i \in L$ having aP_ixP_iy and cP_id when $c \in \{a, x, y\}$ and $d \notin \{a, x, y\}$. 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 (a, b) implies that at any profile $\rho \in \mathcal{S}^n$ in which aP_ib for all $i \in L$, then $\phi(\rho) \neq b$. Suppose not; assume that $\phi(\rho) = b$. Let Q be the ordering that

ρ is single-peaked with respect to. Now consider a $\rho' \in \mathcal{S}_Q^n$ where for each $j \notin L$, P_j is replaced by P_j', in which b is top-ranked under P_j'. By monotonicity, $\phi(\rho') = b$. However, under ρ' we have aP_ib for all $i \in L$ and bP_ja for all $j \notin L$. $\phi(\rho') = b$ contradicts L blocking for (a, b). Thus, L is a blocking coalition. □

Theorem 5. *Let ϕ be a coalitionally strategy-proof collective choice function. If $\mathcal{D} = \mathcal{S}^n$, then ϕ is dictatorial.*

Proof: Consider three profiles $\rho_1, \rho_2, \rho_3 \in \mathcal{S}^n$ 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. By monotonicity, we can consider such profiles without loss of generality.

Let $L \in \mathcal{L}$ 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 ρ_1, ρ_2, ρ_3 as follows:

	ρ_1	ρ_2	ρ_3
i	$x \succ z \succ y$	$x \succ y \succ z$	$x \succ y \succ z$
$L \setminus \{i\}$	$y \succ z \succ x$	$y \succ x \succ z$	$y \succ z \succ x$
$N \setminus L$	$z \succ x \succ y$	$z \succ x \succ y$	$z \succ y \succ x$

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.

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 ρ_1 and ρ_3. Thus, $\phi(\rho_3) = x$ would violate monotonicity.

Now consider an insincere ballot $\hat{\rho}$ that is identical to ρ_1, ρ_2, ρ_3 for all $w \notin \{x, y, z\}$ (i.e. these alternatives are, for every individual, ordered according to $Q_{-\{xyz\}}$), and with a Condorcet cycle over x, y, z at the top:

	$\hat{\rho}$
i	$x \succ y \succ z$
$L \setminus \{i\}$	$y \succ z \succ x$
$N \setminus L$	$z \succ x \succ y$

We know that $\phi(\hat{\rho}) = d \notin \{x, y, z\}$, otherwise ρ_1, ρ_2 or ρ_3 would be manipulable by an individual or coalition submitting ballots as in $\hat{\rho}$: $\phi(\hat{\rho}) = x \Rightarrow i$ manipulates ρ_1, $\phi(\hat{\rho}) = y \Rightarrow L \setminus \{i\}$ manipulates ρ_2, and $\phi(\hat{\rho}) = z \Rightarrow N \setminus L$ manipulates ρ_3.

Now, construct a new profile $\hat{\rho}_1 \in \mathcal{S}^n$ with the preferences of all $j \neq i$ identical to those given by $\hat{\rho}$. Player i's new preferences rank $d \succ x \succ z \succ y \succ$ ..., 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, and thus, a dictator, because all other players prefer x to d.

Case 2: Now suppose that $\phi(\rho_1) = x$. Then immediately, by Lemma 3, this implies that i is a dictator, because i is the unique person who prefers x to z at profile ρ_1. $\square$

When $n = 3$, Theorem 5 can be strengthened to say that if ϕ is a strategy-proof collective choice function and $\mathcal{D} = \mathcal{S}^n$, then ϕ is dictatorial. However, in we cannot weaken coalitional strategy-proofness to strategy-proofness when there is an odd number of players and $n \geq 5$. This is because, with a large enough collection of voters, it is always possible to detect a small set of "potential liars." In particular, 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.[11] In the absence of a core alternative in the submitted profile of ballots, removing these individuals would yield a profile 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 problems, however, 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. The following example demonstrates this, and shows that a choice function that selects elements of the core can still yield opportunities for manipulation when there is an even number of voters, and moreover, single-peakedness need not even be violated for such opportunities to exist.

Example 1. *Manipulating a multi-valued core.*

[11]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.

Theorem 6. *The "Drop 2" mechanism is strategy-proof on the single-peaked preference domain and unrestricted ballot domain when $n \geq 5$ and n is odd.*

Proof: Let $\tilde{Q}$ represent any ordering of the alternatives, with $\tilde{Q}$ ordering the alternatives $x_{\tilde{1}}, ..., x_{\tilde{k}}$. Consider the following "Drop 2" mechanism on $\rho \in \mathcal{P}^n$:

1. If ρ satisfies both worst-restriction and α-restriction, then $\phi(\rho) = C(\rho)$.

2. If ρ violates one of these conditions, then for each $i = 1, ..., n$ test whether ρ_{-i} satisfies both worst-restriction and $\alpha-$restriction. If so, let $i \in W$. In this case, $\phi(\rho) = \text{Argmin}_{x_{\tilde{j}} \in C(\rho_{-W})}\{j\}$ if $|W| < n$, and $\phi(\rho) = x_1$ otherwise.

To show that this mechanism is strategy-proof, we will consider the two cases covered by the mechanism (i.e. when a ballot profile is single-peaked, and when it is not). Let $\rho^* = (P_1, ..., P_n) \in \mathcal{S}^n$ be a sincere preference profile, with $\phi(\rho^*) = x^*$, and throughout, let $\succ_\rho$ denote the strict majority preference relation induced by a profile ρ.

First, suppose that i can profitably manipulate ρ^* with ballot P_i', and that (P_i', ρ_{-i}^*) is single-peaked. This implies that $x = C(P_i', \rho_{-i}^*) \neq C(\rho^*)$, and that xP_ix^*. However, we know that $x^* \succ_{\rho^*} x$, and, because xP_ix^*, that $x^* \succ_{(P_i', \rho_{-i}^*)} x$, contradicting the fact that x is the core of (P_i', ρ_{-i}^*).

Second, suppose that (P_i', ρ_{-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 $\rho \in \mathcal{S}^n$ then $|W| \leq 2$.[12] For this manipulation to be profitable for i, we know that $\phi(P_i', \rho_{-i}^*) = C(\rho_{-W}) = x$,

[12]This is because with $n \geq 5$, at least two people will rank any element of a triple last; thus worst-restriction could be violated by at most two ballots. Furthermore, $\alpha-$restriction can only be violated by a pair of individuals, and the true manipulator will be in any such pair. Thus, $\alpha-$restriction can be violated by at most two individuals.

and xP_ix^*. Let $W = \{i\}$ or $W = \{i, j\}$. Because $x^* \succ_{\rho^*} x$, then $x^* \succ_{\rho_{-W}} x$ 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, $x^* \succ_{\rho_{-W}} x$, contradicting $x \in C(\rho_{-W})$. $\square$

4 Examples of Manipulation in One Dimension

In this section, we briefly consider two well-known models of policymaking in one dimension. In both examples the incentive to manipulate manifests itself as an individual or coalition attempting to make the policy space appear multidimensional. These examples demonstrate several features of our results that may not be apparent from the theoretical sections of this paper, and that are important to note. First, as discussed in 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 $\mathcal{S}^n$, IIA and weak Pareto only imply neutrality of preference aggregation rules on $\mathcal{S}^n$. Section 4.1 illustrates this distinction by providing an example of a generally non-neutral institution that happens to be neutral on $\mathcal{S}^n$, that satisfies IIA and weak Pareto on $\mathcal{S}^n$, and that is also manipulable on $\mathcal{S}^n$.

Second, our theoretical framework leans on an assumption that individuals may submit ballots that are not single-peaked with respect to the true underlying ordering of alternatives. In Section 4.2

We do not intend for any of these examples to be surprising; in fact, the results will appear quite obvious to any reader familiar with the one-dimensional spatial model. Our intention is rather to demonstrate that when viewed as examples of manipulable collective choice functions over a single-peaked preference domain, a common logic explains them.[13]

[13]The points made in this context are essentially extending arguments of Schofield (1995) and Austen-Smith and Banks (1998).

4.1 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.[14] 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. Thus, over the domain $\mathcal{S}^n$, amendment agendas *are* neutral: if the collection of ballots an amendment agenda is given is single-peaked, then so is any permutation of that collection, and the outcome of voting will be the Condorcet winner (and the permuted 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.[15] In this section we consider amendment agendas to be choice functions, ϕ_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[16] yields the same outcome. However, the sequence of votes that individuals cast will differ, and at a

[14]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.

[15]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.

[16]Sophisticated voting in this context refers to individuals playing subgame-perfect Nash equilibrium strategies.

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.

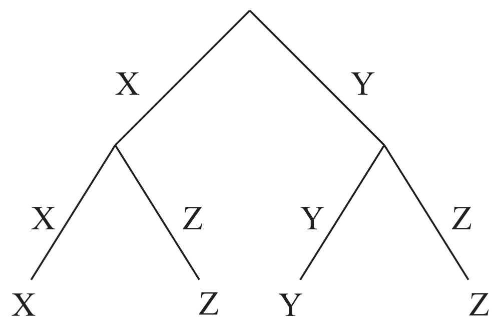

Figure 1: A two-stage amendment agenda

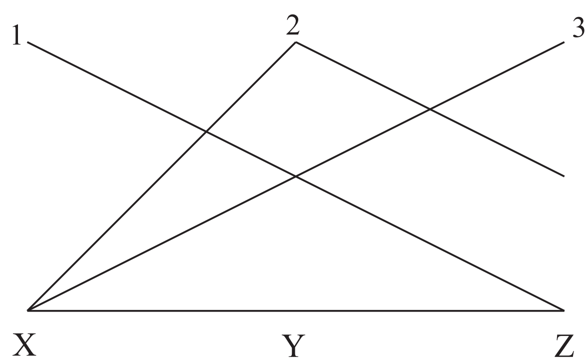

Figure 2: A single-peaked preference profile

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 $zP_3'xP_3'y$. Under this ballot profile our amendment agenda now yields z as the winner: x defeats y at Stage 1 by the (sincere) vote of Player 1 and the (insincere) vote of Player 2, 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 *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 (2002a). 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. Amendment agendas, and binary voting processes in general, are not strategy-proof.

4.2 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. For example, deliberative democracy has been forwarded as a means of divining the "best" choice,[17] as a legitimating device,[18] and as a means by which individuals' preferences may be brought more

[17]Information aggregation citations to be inserted.

[18]Legitimacy citations to be inserted.

in line with each other.[19]

Dryzek and List (2002) explore 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 consistency 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 desirable 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."[20]

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.[21] Our arguments in this paper provide some insight into the conditions that one must impose on the structuration process to guarantee that a democratic decision-making process offers no benefit from misrepresentation, and shows that these conditions are very restrictive. Indeed, the conditions are far more restrictive than has been claimed elsewhere.[22] Specifically, we have shown that a democratic decision-making institution can be ensured to

[19]Preference structuration citations to be inserted.

[20]Dryzek and List (2002b): 14.

[21]Theories discussing the production of structuration through deliberation include Mansbridge (1983), Goodin (1986), and Miller (1992), with empirical studies of the emergence of structuration by Radcliff (1993), and Farrar et al. (2006) – add these references to bibliography and include a more thorough discussion here.)

[22]In addition to Dryzek and List (2002b), our results call into question some of the relevant claims of Grofman and Feld (1988),

offer no benefit from misrepresentation by a coalition only if the details of the structure of the individual's preferences are written into the rules of the institution itself (i.e. the institution can utilize the underlying ordering of alternatives), or if the institution is dictatorial.

When looking at the incentives for individual misrepresentation of preferences our results are somewhat less bleak. Theorem 6 describes a mechanism that is strategy-proof when preferences are single-peaked, and ultimately this mechanism involves detecting potential liars, dropping their input from the deliberative process, and implementing a Condorcet winner. We leave open the question of whether we can design a non-dictatorial and strategy-proof institution that can implement anything other than a Condorcet winner.

5 Conclusions

Theorems 4 and 5 imply that one must be careful in interpreting collective will in any real-world policy-making 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)). Specifically, appeals to aggregate outcomes as indicators of collective will are not necessarily well-founded even when the majority will is assumed to exist. "Faithful representation" of the majority will within non-dictatorial institutions will occasionally take an insincere form. Accordingly, the normative, prescriptive, descriptive, and inferential issues raised by Arrow's theorem and the Gibbard-Satterthwaite theorem are more than simple mathematical curiosities dreamed up for the purpose of scholarly debate.

Single-peakedness does not solve problems of cycling in the real-world because policymaking institutions are generally not neutral. Specifically, the presumption that individuals have single-peaked preferences is not sufficient to assume that the result of their aggregation is a well-defined collective will. Single-

Miller (1992), and Mackie (2003), among others.

peakedness does not eliminate the possibility of gains through strategic manipulation within real-world institutions, because few (if any) policymaking institutions are dictatorial, and few (if any) policymaking institutions limit the preferences that groups of individuals can *claim* to have.

References

Arrow, K. J. (1951). *Social Choice and Individual Values*. John Wiley and Sons, New York, NY.

Austen-Smith, D. (1987). Sophisticated sincerity: Voting over endogenous agendas. *American Political Science Review*, 81(4):1323–1330.

Austen-Smith, D. and Banks, J. S. (1998). Social choice theory, game theory, and positive political theory. *Annual Review of Political Science*, 1(1):259–287.

Austen-Smith, D. and Banks, J. S. (2004). *Positive Political Theory II: Strategy & Structure*. University of Michigan Press, Ann Arbor, MI.

Ballester, M. A. and Haeringer, G. (2007). A Characterization of the Single-Peaked Domain. Working paper, Universitat Autònoma de Barcelona.

Black, D. (1948). On the Rationale of Group Decision-making. *Journal of Political Economy*, 56:23–34.

Blin, J. M. and Satterthwaite, M. A. (1976). Strategy-proofness and single-peakedness. *Public Choice*, 26(1):51–58.

Davis, O., Hinich, M., and Ordeshook, P. (1970). An Expository Development of a Mathematical Model of the Electoral Process. *American Political Science Review*, 54:426–448.

Denzau, A., Riker, W., and Shepsle, K. (1985). Farquharson and Fenno: Sophisticated Voting and Home Style. *American Political Science Review*, 79(4):1117–1134.

Dowding, K. and Van Hees, M. (2007). In Praise of Manipulation. *British Journal of Political Science*, 38(1):1–15.

Downs, A. (1957). *An Economic Theory of Democracy*. Harper and Row, New York.

Dryzek, J. and List, C. (2002a). Social Choice Theory and Deliberative Democracy: A Reconciliation. *British Journal of Political Science*, 33(1):1–28.

Dryzek, J. S. and List, C. (2002b). Social Choice Theory and Deliberative Democracy: A Reconciliation. *British Journal of Political Science*, 33(1):1–28.

Dummett, M. and Farquharson, R. (1961). Stability in Voting. *Econometrica*, 29(1):33–43.

Feld, S. and Grofman, B. (1988). Ideological Consistency as a Collective Phenomenon. *American Political Science Review*, 82(3):773–788.

Gailmard, S., Patty, J., and Penn, E. M. (2008). Arrow's theorem on single-peaked domains. *Mimeo*.

Gibbard, A. (1973). Manipulation of Voting Schemes: A General Result. *Econometrica*, 41(4):587–601.

Grofman, B. and Feld, S. L. (1988). Rousseau's General Will: A Condorcetian Perspective. *American Political Science Review*, 82(2):567–576.

Krehbiel, K. (1998). *Pivotal Politics: A Theory of U.S. Lawmaking*. University of Chicago Press, Chicago, IL.

Mackie, G. (2003). *Democracy Defended*. Cambridge University Press, New York, NY.

McCubbins, M., Noll, R., and Weingast, B. (1994). Legislative Intent: The Use of Positive Political Theory in Statutory Interpretation. *Journal of Law and Contemporary Problems*, 57(1):3–37.

Miller, D. (1992). Deliberative Democracy and Social Choice. *Political Studies*, 40(5):54–67.

Miller, N. R. (1977). Graph-Theoretical Approaches to the Theory of Voting. *American Journal of Political Science*, 21:769–803.

Pattanaik, P. (1976). Collective rationality and strategy-proofness of group decision rules. *Theory and Decision*, 7(3):191–203.

Poole, K. and Rosenthal, H. (1997). *Congress: A Political-Economic History of Roll-Call Voting*. Oxford University Press, New York, NY.

Reny, P. J. (2000). Arrow's theorem and the gibbard-satterthwaite theorem: A unified approach. *Mimeo*, Department of Economics, University of Chicago.

Riker, W. (1992). The Justification of Bicameralism. *International Political Science Review/ Revue internationale de science politique*, 13(1):101–116.

Satterthwaite, M. A. (1975). Strategy-Proofness and Arrows Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions. *Journal of Economic Theory*, 10:187–217.

Schofield, N. (1995). Coalition Politics: A Formal Model and Empirical Analysis. *Journal of Theoretical Politics*, 7:245–281.

Sen, A. K. (1966). A Possibility Theorem on Majority Decisions. *Econometrica*, 34(2):491–499.

Sen, A. K. and Pattanaik, P. K. (1969). Necessary and Sufficient Conditions for Rational Choice Under Majority Decision. *Journal of Economic Theory*, 1(2):178–202.

Ubeda, L. (2003). Neutrality in arrow and other impossibility theorems. *Economic Theory*, 23(1):195.

Young, H. P. (1995). Optimal Voting Rules. *Journal of Economic Perspectives*, 9(1):51–64.