UNIVERSITÀ DI SIENA
1240

QUADERNI DEL DIPARTIMENTO
DI ECONOMIA POLITICA E STATISTICA

Stefano Vannucci

AGENDA MANIPULATION-PROOFNESS, STALEMATES, AND THE WORTH OF REDUNDANT ELICITATION IN PREFERENCE AGGREGATION.

EXPOSING THE BRIGHT SIDE OF ARROW'S THEOREM.

n. 881 – Maggio 2022

AGENDA MANIPULATION-PROOFNESS, STALEMATES, AND THE WORTH OF REDUNDANT ELICITATION IN PREFERENCE AGGREGATION.

EXPOSING THE BRIGHT SIDE OF ARROW'S THEOREM.

STEFANO VANNUCCI

Abstract. This paper provides a general framework to explore the possibility of agenda manipulation-proof and proper consensus-based preference aggregation rules, so powerfully called in doubt by a disputable if widely shared understanding of Arrow's 'general possibility theorem'. We consider two alternative versions of agenda manipulation-proofness for social welfare functions, that are distinguished by 'parallel' vs. 'sequential' execution of agenda formation and preference elicitation, respectively. Under the 'parallel' version, it is shown that a large class of anonymous and idempotent social welfare functions that satisfy both agenda manipulation-proofness and strategy-proofness on a natural domain of single-peaked 'meta-preferences' induced by arbitrary total preference preorders are indeed available. It is only under the second, 'sequential' version that agenda manipulation-proofness on the same natural domain of single-peaked 'meta-preferences' is in fact shown to be equivalent to the classic Arrowian 'independence of irrelevant alternatives' for social welfare functions. In particular, it is shown that combining such 'sequential' version of agenda manipulation-proofness with a very minimal requirement of distributed responsiveness results in a characterization of the 'global stalemate' social welfare function, the constant function which invariably selects universal social indifference. It is also argued that, altogether, the foregoing results provide new significant insights concerning the actual content and the *constructive* implications of Arrow's 'general possibility theorem' from a mechanism-design perspective.

2000 Mathematics Subject Classifications. Primary 05C05; Secondary 52021, 52037.

JEL Classification D71

Keywords: Agenda manipulation, strategy-proofness, social welfare functions, aggregation, median join-semilattices.

1991 *Mathematics Subject Classification.* Mathematics Subject Classification. Primary 05C05; Secondary 52021, 52037.

Key words and phrases. Agenda manipulation, strategy-proofness, social welfare function, aggregation, median join-semilattice.

1. Introduction

The *agenda* is a key part of any decision problem, and is specified by its two components. The first component, its *content*, is defined by the alternative options it contains. The second component, its *structure*, is the rule governing the process of option-scrutiny that is required in order to solve the given decision problem. Thus, such a structure amounts to a *total preorder* of the available options, resulting in an ordered partition of options that channels their *sequential* scrutiny. It follows that *agenda control* is perforce a crucial issue, and has at least *two* dimensions, namely agenda-content and agenda-structure control. Therefore *agenda manipulation,* or the exercise of agenda control to influence the final decision, is also a 'two-dimensional' activity which must be taken into account, to be possibly prevented or restrained.

As it happens, concerns for *agenda manipulation along both of its dimensions* have always played a distinguished role in the literature on collective decision-making, but have scarcely if ever been the target of explicit treatment within formal models of preference aggregation. To be sure, it is well-known that under transitivity of the relevant preference relation, preference-maximizing choices are *path-independent,* namely do not depend on the sequence of intermediate choices and rejections out of the sequence of subsets which is dictated by the agenda-structure. Hence, whenever the aggregation rule is a *social welfare function*, meaning that both the individual preference relations to be aggregated and the 'social' or aggregate ones are *transitive* (and *total*), agenda-structure control is virtually inconsequential, and agenda manipulation along the agenda-structure dimension is automatically prevented. On the contrary, agenda-content manipulation is of course still possible even if all the relevant preference relations are transitive. Since we are going to focus precisely on social welfare functions, in the sequel we shall only discuss agenda-content manipulation, disregarding entirely agenda-structure manipulation. Therefore, in the rest of the present paper we shall simply *identify agendas with their contents*[1]. Accordingly, 'agenda manipulation' is henceforth used, for the sake of simplicity, as a synonym for 'agenda-content manipulation'.

Indeed, in his classic *Social Choice and Individual Values* (1963) Arrow used precisely *the need to prevent agenda manipulation* as the

[1]By contrast, most contributions to agenda control and manipulation in the political science literature are focussed on *agenda-structure* control and manipulation (see e.g. Austen-Smith, Banks (2005) for an extensive treatment of agenda-structure control models in political science).

main argument in favor of his own 'Independence of Irrelevant Alternatives' (IIA), as a key condition for *proper* (or non-trivial[2]) *consensus-based social welfare functions,* which are the main focus of that work. Since then, the notion that IIA should be regarded as a basic 'non-manipulability property' has been further reinforced by the rise and enormous proliferation of models focussing on *strategic manipulation* issues in preference aggregation, to become eventually almost common place. Yet, within standard models of preference aggregation including social welfare functions it is *just preference relations* that agents provide as *inputs* while the relevant agenda is a *parameter* of the aggregation rule. But then, such an aggregation rule is nothing else than the relevant strategic *game-form.* It follows that agenda-manipulation amounts to a *structural manipulation* of the very 'aggregation game', literally a game-changer. Hence, the exact connection between agenda manipulation-proofness and IIA is *not* amenable to a proper game-theoretic scrutiny *unless* the preference aggregation model is expanded to involve the *agenda formation process* itself. In particular, such an expanded model is needed to establish whether the full force of IIA is actually necessary to prevent agenda manipulation for social welfare functions that are at least minimally outcome-unbiased and agent-inclusive.

Thus, some explicit formulation of the agenda formation process has to be introduced in the relevant preference aggregation model. In the present work *two main types of agenda formation protocols* are considered. Both of them rely on a prespecified *admissible set of* outcomes out of which the *actual agenda* has to be defined. Moreover, in order to avoid any sort of infinite regress, we can safely assume that outcome-admissibility is established by another (possibly 'democratic', but *distinct*) procedure[3]. In the first agenda formation protocol, however, agents provide *at once* both their preferences on admissible outcomes and their proposals concerning the agenda. In the second one, on the contrary, a *first stage* is devoted to *specifying the actual agenda,* and is

[2]A dictatorial social welfare function is of course unanimity-respecting and thus in a sense also consensus-based, but only trivially so.

[3]It is worth recalling here that Dahl (1956) famously suggested to label as 'populist democracy' the doctrine that advocates reliance on the simple majority rule to settle every issue including the identification of the *admissible issues* for a possible public agenda. Arguably, one might invoke a generalized notion of 'populist democracy' as the advocacy of a unique 'democratic' decision rule to settle every issue including every aspect pertaining to agenda control. In that connection, assuming that admissibility of outcome sets is subject to a *distinct protocol* (if possibly also 'democratic' in some appropriate sense) amounts to preventing any 'populist' interpretation of the overall decision mechanism.

followed by a *second, preference-elicitation* stage where the agents express their preferences on the previously chosen agenda. Accordingly, two distinct formulations of *Agenda Manipulation-Proofness* (AMP) are introduced, and their distinctive impact on the design of preference aggregation rules that guarantee at least a minimal amount of *outcome-unbiased distributed responsiveness* to individual preferences is explored and discussed at length. In particular, it is shown that under the *first* formulation, AMP *and* Strategy-Proofness on a comprehensive single-peaked domain of 'meta-preferences' most naturally induced by basic preferences on outcomes are shared by a large class of preference aggregation rules on the full domain of *arbitrary* profiles of total preference preorders[4]. Such a class of strategy-proof aggregation rules includes social welfare functions which indeed satisfy a remarkable combination of valuable properties. Namely, anonymity, monotonicity and a *basic* version of Pareto-optimality, possibly even (weak) neutrality, though occasionally producing *stalemates*[5] as an output to some preference profiles exhibiting certain *specific* patterns of strong conflict (e.g. Condorcet cycles)[6]. By contrast, under the *second version* of AMP a couple of much weaker, indeed minimal, requirements of unbiased and distributed responsiveness jointly provide a characterization of the *Global Stalemate constant social welfare function*, namely the social welfare functions which has *universal indifference* as its *unique possible output.*[7] Moreover, if the Weak Pareto property or even just idempotence (namely, 'respect for unanimity') is adjoined to IIA, an Arrowian impossibility result is obtained. Thus, in order to secure agenda manipulation-proofness of a social welfare functions one may consider *two* basic alternative approaches having strikingly *different* consequences. One of those approaches amounts to the introduction of IIA: it was correctly identified by Arrow's seminal contribution, and paves the way to his classic characterization of *dictatorial* social welfare

[4]As mentioned below, such a result relies heavily on the main theorem of Savaglio, Vannucci (2021) concerning strategy-proof aggregation rules in median join-semilattices.

[5]A stalemate is defined as 'social indifference' among a set of alternative social states including a pair x, y such that x is unanimously strictly preferred to y. Thus, by definition, a stalemate admits of violations of the Weak Pareto principle (which enforces strict social preference under the aforementioned situation).

[6]The above mentioned social welfare functions are the quota rules, including the Condorcet-Kemeny median rule which is indeed neutral when the number of agents is odd.

[7]Such a result, which amounts to a considerable strenghtening of a previous characterization of the same constant social welfare function due to Hansson (1969), relies heavily on Wilson (1972) and Savaglio,Vannucci (2021).

functions. That result signals an important obstruction to the design of social welfare functions as democratic preference aggregation protocols. The other approach, however, has no connection whatsoever to IIA, and is consistent with a large class of anonymous, unanimity-respecting social welfare functions that also retain a basic version of Pareto optimality involving nonstrict preferences. We argue that the very contrast between those two approaches and their respective results makes it possible to single out and appreciate the *constructive* implications of Arrowian 'impossibility theorems' concerning the design of preference aggregation rules, as a significant part of their actual meaning and content.

The rest of this paper is organized as follows: section 2 collects the formal description of the model and the results; section 3 is devoted to a detailed if highly selective discussion of the massive amount of related work; section 4 provides some concluding remarks, and prospects for future research.

2. Model and results

Let A be a nonempty finite set of alternative social states with $|A| \geq 3$, $\mathcal{R}_A$ the set of all total preorders (i.e. reflexive, transitive and connected binary relations) on A, $\mathcal{L}_A \subseteq \mathcal{R}_A$ the set of all linear orders (or *antisymmetric* total preorders on A), and $\mathcal{P}(A)$ the set of parts of A, or possible *agendas* from A. Let $N = \{1, .., n\}$ denote a *finite* population of agents/voters. We assume that $n \geq 3$ in order to avoid tedious qualifications. The subsets of N are also referred to as *coalitions,* and $(\mathcal{P}(N), \subseteq)$ denotes the partially ordered set of coalitions induced by set-inclusion. An *order filter* of $(\mathcal{P}(N), \subseteq)$ is a set $F \subseteq \mathcal{P}(N)$ of coalitions such that for any $S \in F$ and any $T \subseteq N$, if $S \subseteq T$ then $T \in F$. The *basis* of order filter F is the set of inclusion-minimal elements/coalitions of F, and is denoted by $F^{\min}$.

Each agent $i \in N$ is endowed with a total preference preorder $R_i \in \mathcal{R}_A$ (whose *asymmetric* component or *strict preference* is denoted by $P(R_i)$), and proposes an agenda $A_i \subseteq A$. A *social welfare function* for (N, A) if a function $f : \mathcal{R}_A^N \to \mathcal{R}_A$. We shall also consider two types of social welfare functions enriched with an *endogenous agenda formation* process. According to the class of *parallel* rules *agents release concurrently their entire inputs* consisting of a total preorder on the set of *all* admissible alternatives and of a proposed agenda: preference aggregation and agenda formation are also mutually concurrent

processes. Notice that this also entails that typically, namely whenever the selected agenda is a *proper* subset of A, the elicited individual preferences turn out to be *redundant.*

Thus, a *parallel agenda-formation-enriched (PAFE) social welfare function* for (N, A) is an aggregation rule $\mathbf{f} : (\mathcal{P}(A) \times \mathcal{R}_A)^N \to \mathcal{P}(A) \times \mathcal{R}_A$ (with projections $\mathbf{f}_1$ and $\mathbf{f}_2$ on $\mathcal{P}(A)$ and $\mathcal{R}_A$, respectively). In particular, such a PAFE $\mathbf{f}$ is said to be *decomposable* if and only if it can be decomposed into two component aggregation rules: an agenda formation rule $f^{(1)} : \mathcal{P}(A)^N \longrightarrow \mathcal{P}(A)$ and a social welfare function $f^{(2)} : \mathcal{R}_A^N \to \mathcal{R}_A$ [8]: hence, one can also write $\mathbf{f} \simeq f^{(1)} \times f^{(2)}$. Observe that a decomposable PAFE $\mathbf{f}$ also induces a family of functions $\mathcal{F}_{\mathbf{f}} := \left\{ f_B^{(2)} : \mathcal{R}_A^N \to \mathcal{R}_B \mid B \in f^{(1)}[\mathcal{P}(A)] \right\}$ where $f_B^{(2)}(R_N) := (f^{(2)}(R_N))_{|B}$ for any $R_N \in \mathcal{R}_A^N$. Accordingly, the possible values of functions in $\mathcal{F}_{\mathbf{f}}$ are given by a family of total preorders, namely $\left\{ f_B^{(2)}(R_N) \right\}_{R_N \in \mathcal{R}_A^N, B \in \mathcal{P}(A)]}$: thus, for every $R_N \in \mathcal{R}_A^N$ and $B \in \mathcal{P}(A)$, $f_B^{(2)}(R_N) \in \mathcal{R}_B \subseteq \bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B$. In particular, we shall focus on *sovereign* agenda-formation rules $f^{(1)} : \mathcal{P}(A)^N \longrightarrow \mathcal{P}(A)$ (namely, such that for every $C \subseteq A$ there exists $B_N \in \mathcal{P}(A)^N$ with $f^{(1)}(B_N) = C$).

Claim 1. *Let f' be an agenda formation rule for (N, A) and f a social welfare function for (N, A), namely $f' : \mathcal{P}(A)^N \longrightarrow \mathcal{P}(A)$ and $f : \mathcal{R}_A^N \to \mathcal{R}_A$. Then, there exists a decomposable PAFE social welfare function $\mathbf{f} \simeq f' \times f$.*

Proof. Trivial: just take $\mathbf{f}((B_i, R_i)_{i \in N}) = (f'(B_N), f(R_N))$ for each $(B_i, R_i)_{i \in N} \in (\mathcal{P}(A) \times \mathcal{R}_A)^N$. □

Hence, *any* social welfare function can be regarded as a component of a decomposable PAFE social welfare function by combining it with an agenda formation rule.

[8] In other terms, the two projections of $\mathbf{f}$, namely $\mathbf{f}_1 : (\mathcal{P}(A) \times \mathcal{R}_A^T)^N \longrightarrow \mathcal{P}(A)$ and

$\mathbf{f}_2 : (\mathcal{P}(A) \times \mathcal{R}_A^T)^N \longrightarrow \mathcal{R}_A^T$ are such that for every $R_N, R'_N \in (\mathcal{R}_A^T)^N$ and $B_N, B'_N \in (\mathcal{P}(A))^N$:

$\mathbf{f}_1(B_N, R_N) = \mathbf{f}_1(B_N, R'_N) := f^{(1)}(B_N)$

and

$\mathbf{f}_2(B_N, R_N) = \mathbf{f}_2(B'_N, R_N) := f^{(2)}(R_N)$.

Under the class of *sequential* rules, on the contrary, *agents release their inputs in two steps*: first they provide concurrently their proposed agendas to be aggregated into a shared agenda, then they submit concurrently their preferences on the previously determined actual agenda as their input to preference aggregation itself. Notice that in this case, no redundancy in preference elicitation is to be expected. Thus, a *sequential agenda-formation-enriched (SAFE) social welfare function* for (N, A) is in fact an *agenda-contingent social welfare function,* namely a pair $\widehat{\mathbf{f}} = (\widehat{f}^1, \mathcal{F}(\widehat{f}^1))$ consisting of an agenda formation rule $\widehat{f}^1 : \mathcal{P}(A)^N \longrightarrow \mathcal{P}(A)$, and a family $\mathcal{F}(\widehat{f}^1) = \left\{\widehat{f}_B : \mathcal{R}_B^N \to \mathcal{R}_B\right\}_{B \in \widehat{f}^1[\mathcal{P}(A)]}$ of possible social welfare functions, one for each possible agenda selected by $\widehat{f}^1$. Hence, again, the values of possible social welfare functions according to $\widehat{\mathbf{f}}$ are in $\bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B$.

As a consequence, SAFE and (decomposable) PAFE social welfare functions essentially share $\bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B$ as their common outcome space.

Therefore, it transpires that SAFE and PAFE social welfare functions for a pair (N, A) consist of functions whose domains and codomains result from various combinations of 'building blocks' chosen from the collection given by $\mathcal{P}(A)$ and the sets $\mathcal{R}_B$ of all total preorders on B, for any $B \subseteq A$. As it turns out, such 'building blocks' share a common structure: all of them are *median join-semilattices,* and that fact will play a key role in the subsequent analysis of the behaviour of PAFE and SAFE social welfare functions. Accordingly, we turn now to providing a precise definition of median join-semilattices, and establishing the previous claim on $\mathcal{P}(A)$ and the sets of the family $\{\mathcal{R}_B\}_{B \subseteq A}$.

Definition 1. *A (finite)* ***join-semilattice*** *is a pair $\mathcal{X} = (X, \leqslant)$ where X is a (finite) set and $\leqslant$ is a partial order (i.e. a reflexive, transitive and antisymmetric binary relation) such that the least upper bound or join $x \vee y$ (with respect to $\leqslant$) is well-defined in X for all $x, y \in X$ and thus $\vee : X \times X \to X$ is a well-defined associative and commutative function that also satisfies idempotency, namely $x \vee x = x$ for every $x \in X$.*

Remark 1. *Thus a join-semilattice $\mathcal{X} = (X, \leqslant)$ can also be regarded as a pair $(X, \vee)$ where $\vee : X \times X \longrightarrow X$ is an associative, commutative and idempotent operation such that, for any $x, y \in X$, $x \vee y = x$ iff $y \leqslant x$. Note that a partial meet-operation $\wedge : X \times X \longrightarrow X$ is also*

definable in $\mathcal{X}$ by means of the following rule: for any $x, y \in X$, $x \wedge y$ is the (necessarily unique, whenever it exists) $z \in X$ such that: (i) $x \vee z = x$, $y \vee z = y$, and (ii) $v \vee z = z$ for every $v \in X$ which satisfies (i).

Observe that a finite join-semilattice $\mathcal{X} = (X, \leqslant)$ has a (unique) universal upper bound or *top element* $\mathbf{1} = \vee X = \wedge \varnothing$, and its *co-atoms* are those elements $x \in X$ such that $x \ll \mathbf{1}$ (i.e. $x < \mathbf{1}$ and there is no $z \in X$ such that $x < z < \mathbf{1}$): the set of co-atoms of $\mathcal{X} = (X, \leqslant)$ is denoted by $\mathcal{C}_{\mathcal{X}}$. An element $x \in X$ is *meet-irreducible* if for any $Y \subseteq X$, $x = \wedge Y$ entails $x \in Y$. Moreover, for any $Y \subseteq X$, $\vee Y$, respectively is well-defined if and only if there exists $z \in X$ such that $y \leqslant z$ for all $y \in Y$, namely the elements of Y have a *common upper bound*. The set of all meet-irreducible elements of $\mathcal{X} = (X, \leqslant)$ will be denoted by $M_{\mathcal{X}}$. Notice that, by construction, for every $x \in X$, $x = \wedge M(x)$ where $M(x) := \{m \in M_{\mathcal{X}} : x \leqslant m\}$. By construction, a co-atom is also a meet-irreducible element, but the converse need not be true. When co-atoms and meet-irreducibles do in fact *coincide* the join-semilattice is said to be *coatomistic*.[9]

Definition 2. ***(Median join-semilattice)*** *A (finite) join-semilattice $\mathcal{X} = (X, \leqslant)$ is a* ***median join-semilattice*** *if it also satisfies the following pair of conditions:*

i) ***upper distributivity****: for all $u \in X$, and for all $x, y, z \in X$ such that u is a lower bound of $\{x, y, z\}$, $x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$ (or, equivalently, $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$) holds i.e. $(\uparrow u, \leqslant_{|\uparrow u})$ -where $\leqslant_{|\uparrow x}$ denotes the restriction of $\leqslant$ to $\uparrow u$- is a distributive lattice*[10]*;*

---

[9]Dually, a (finite) **meet-semilattice** is a pair $\mathcal{X} = (X, \leqslant)$ where X is a (finite) set, and partial order $\leqslant$ is such that the greatest lower bound or *meet* $x \wedge y$ (with respect to $\leqslant$) is well-defined in X for all $x, y \in X$ and thus $\wedge : X \times X \to X$ is a well-defined associative and commutative function that also satisfies idempotency, namely $x \wedge x = x$ for every $x \in X$.

An element $x \in X$ of a meet-semilattice is *join-irreducible* if for any $x = \vee Y$ entails $x \in Y$ for any (finite) $Y \subseteq X$ such that $\vee Y$, is well-defined. The set of all join-irreducible elements of $\mathcal{X} = (X, \leqslant)$ is denoted by $J_{\mathcal{X}}$. The *atoms* of $\mathcal{X}$ are those elements $x \in X$ such that $\mathbf{0} \ll x$ (i.e. $\mathbf{0} < x$ and there is no $z \in X$ such that $\mathbf{0} < x < z$ where $\mathbf{0} = \wedge X = \vee \varnothing$): the set of atoms of $\mathcal{X}$ is denoted by $\mathcal{A}_{\mathcal{X}}$. Clearly, $\mathcal{A}_{\mathcal{X}} \subseteq J_{\mathcal{X}}$. The semilattice $\mathcal{X}$ is *atomistic* if $\mathcal{A}_{\mathcal{X}} = J_{\mathcal{X}}$.

Notice that, by construction, for every $x \in X$, $x = \vee J(x)$ where $J(x) := \{j \in J_{\mathcal{X}} : j \leqslant x\}$.

If a join-semilattice $\mathcal{X} = (X, \leqslant)$ is *also* a meet-semilattice then $\mathcal{X}$ is a **lattice** and *absorption laws* hold, namely for any $x, y \in X$, $x \vee (y \wedge x) = x = x \wedge (y \vee x)$.

[10]A partially ordered set $(Y, \leqslant)$ is a *distributive lattice* iff, for any $x, y, z \in X$, $x \wedge y$ and $x \vee y$ exist, and $x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z)$ (or, equivalently,

(ii) ***co-coronation*** *(or meet-Helly property): for all $x, y, z \in X$ if $x \wedge y$, $y \wedge z$ and $x \wedge z$ exist, then $(x \wedge y \wedge z)$ also exists.*

In fact, it is easily checked that if $\mathcal{X} = (X, \leqslant)$ is a median join-semilattice then the partial function $\mu_{\mathcal{X}} : X^3 \to X$ defined as follows: for all $x, y, z \in X$, $\mu_{\mathcal{X}}(x, y, z) = (x \vee y) \wedge (y \vee z) \wedge (x \vee z)$

is in fact a *well-defined ternary operation* on X, the **median** of $\mathcal{X}$ which satisfies the following two characteristic properties (see Sholander (1952, 1954)):

(μ_1) $\mu_{\mathcal{X}}(x, x, y) = x$ for all $x, y \in X$

(μ_2) $\mu_{\mathcal{X}}(\mu_{\mathcal{X}}(x, y, v), \mu_{\mathcal{X}}(x, y, w), z) = \mu_{\mathcal{X}}(\mu_{\mathcal{X}}(v, w, z), x, y)$

for all $x, y, v, w, z \in X$.

A pair (X, μ) where μ is a ternary operation on X that satisfies (μ_1) and (μ_1) is also said to be a *median algebra*.

Relying on $\mu_{\mathcal{X}}$, a ternary (median-induced) **betweenness** relation

$B_{\mu_{\mathcal{X}}} := \{(x, z, y) \in X^3 : z = \mu_{\mathcal{X}}(x, y, z)\}$ can also be defined on X. [11] The pair $(X, B_{\mu_{\mathcal{X}}})$ is also said to be a *median (ternary) space.*

Remark 2. *It is worth emphasizing here that any finite median join-semilattice is naturally endowed with two equivalent* ***metrics***[12]*, and that a further betweenness relation can be defined on it relying on such metrics. However, it turns out that such a metric-based betweenness is in fact equivalent to the median-based betweenness $B_{\mu_{\mathcal{X}}}$ introduced above in the text (see e.g. Sholander (1954), Avann (1961)).*

$x \vee (y \wedge z) = (x \vee y) \wedge (x \vee z)$). Moreover, a (distributive) lattice $\mathcal{X}$ is said to be *lower (upper) bounded* if there exists $\perp \in X$ ($\top \in X$) such that $\perp \leqslant x$ ($x \leqslant \top$) for all $x \in X$, and *bounded* if it is both lower bounded and upper bounded. A bounded distributive lattice $(X, \leqslant)$ is *Boolean* if for each $x \in X$ there exists a *complement* namely an $x' \in X$ such that $x \vee x' = \top$ and $x \wedge x' = \perp$.

[11] It should be recalled that such a median operation μ is also well-defined in any *distributive lattice* $\mathcal{X} = (X, \leqslant)$. Thus, every (finite) distributive lattice is in particular a (finite) median join-semilattice (and a (finite) median meet-semilattice as well).

[12] In a (finite) median join-semilattice $\mathcal{X} = (X, \leqslant)$ a *metric* $d_r : X \times X \to \mathbb{Z}_+$ can be defined in a natural way by the following rule: for any $x, y \in X$, $d_r(x, y) = 2r(x \vee y) - r(x) - r(y)$, where r is a *rank function* of $\mathcal{X}$, namely a function $r : X \to \mathbb{Z}_+$ such that, for any $x, y \in X$, $r(y) = r(x) + 1$ whenever x is an immediate $\leqslant$-predecessor of y.

This metric turns out to be equivalent to the metric $\delta_{C(\mathcal{X})}$ induced on $\mathcal{X}$ by the length of shortest path between any two elements on the graph defined by the Hasse diagram of $\mathcal{X}$ (the simple undirected graph having X as its set of vertices, with edges connecting each pair consisting of a vertex and one of its immediate $\leqslant$-predecessors).

Thus, $B_{\mu_{\mathcal{X}}}$ is indeed a most natural 'intrinsic' betweenness relation and can also be regarded as 'the' natural metric betweenness attached to $\mathcal{X}$.

Relying on such a betweenness $B_{\mu_{\mathcal{X}}}$, a 'natural' notion of *single-peakedness* for preference preorders on $\mathcal{X} = (X, \leqslant)$ can be defined as follows.

Definition 3. *(**Single-peaked preference preorders on a median join-semilattice**). Let $\mathcal{X} = (X, \leqslant)$ be a finite median join-semilattice and $\succcurlyeq$ a preorder i.e. a reflexive and transitive binary relation on X (we shall denote by $\succ$ and $\sim$ its asymmetric and symmetric components, respectively). Then, $\succcurlyeq$ is said to be **single-peaked** with respect to betweenness relation $B_{\mu_{\mathcal{X}}}$ (or $B_{\mu_{\mathcal{X}}}$-single-peaked) if and only if*

U-(i) there exists a unique maximum of $\succcurlyeq$ in X, its top outcome -denoted top($\succcurlyeq$)- and

U-(ii) for all $x, y, z \in X$, if $x = top(\succcurlyeq)$ and $z = \mu_{\mathcal{X}}(x, y, z)$ then not $y \succ z$.

We denote by U_{B_μ} the set of all $B_{\mu_{\mathcal{X}}}$-*single-peaked* preorders on X. An N-profile of B_μ-single-peaked preorders is a mapping from N into U_{B_μ}. We denote by $U^N_{B_{\mu_{\mathcal{X}}}}$ the set of all N-profiles of B_μ-lu preorders.

Moreover, a set $D \subseteq U^N_{B_{\mu_{\mathcal{X}}}}$ of preorders which are single-peaked w.r.t. $B_{\mu_{\mathcal{X}}}$ is a **rich single-peaked domain** for $\mathcal{X}$ if for all $x, y \in X$ there exists $\succcurlyeq \in D$ such that $top(\succcurlyeq) = x$ and $UC(\succeq, y) = \{z \in X : z = \mu(x, y, z)\}$ (where $UC(\succeq, y) := \{y \in X : x \succcurlyeq y\}$ is the upper contour of $\succcurlyeq$ at y).

An aggregation rule f for (N, X) is **strategy-proof** *on* $U^N_{B_{\mu_{\mathcal{X}}}}$ iff for all $B_{\mu_{\mathcal{X}}}$-single-peaked N-profiles $(\succcurlyeq_i)_{i \in N} \in U^N_{B_{\mu_{\mathcal{X}}}}$, and for all $i \in N$, $y_i \in X$, and $(x_j)_{j \in N} \in X^N$ such that $x_j = top(\succcurlyeq_j)$ for each $j \in N$, *not* $f((y_i, (x_j)_{j \in N \setminus \{i\}})) \succ_i f((x_j)_{j \in N})$. Finally, an aggregation rule $f : X^N \to X$ is $B_{\mu_{\mathcal{X}}}$-**monotonic** iff for all $i \in N$, $y_i \in X$, and $(x_j)_{j \in N} \in X^N$,

$f((x_j)_{j \in N}) = \mu_{\mathcal{X}}(x_i, f((x_j)_{j \in N}), f(y_i, (x_j)_{j \in N \setminus \{i\}}))$.[13]

In particular, let $\mathcal{X} = (X, \leqslant)$ be a finite *join-semilattice* and $M_{\mathcal{X}}$ the set of its meet-irreducible elements, and for any $x^N \in X^N$, and any $m \in M_{\mathcal{X}}$, posit $N_m(x^N) := \{i \in N : x_i \leqslant m\}$. Then, the following properties of an aggregation rule can also be introduced:

[13] $B_{\mu_{\mathcal{X}}}$-monotonicity of f amounts to requiring all of its projections f_i to be *gate maps* to the image of f (see van de Vel (1993), p.98 for a definition of gate maps). The introduction of $B_{\mu_{\mathcal{X}}}$-monotonic functions in a strategic social choice setting is essentially due to Danilov (1994).

$M_{\mathcal{X}}$-Independence: an aggregation rule $f : X^N \to X$ is **$M_{\mathcal{X}}$-independent** if and only if for all $x_N, y_N \in X^N$ and all $m \in M_{\mathcal{X}}$: if $N_m(x_N) = N_m(y_N)$ then $f(x_N) \leqslant m$ if and only if $f(y_N) \leqslant m$.

Isotony: an aggregation rule $f : X^N \to X$ is **Isotonic** if $f(x_N) \leqslant f(x'_N)$ for all $x_N, x'_N \in X^N$ such that $x_N \leqslant x'_N$ (i.e. $x_i \leqslant x'_i$ for each $i \in N$).

It can be easily shown (see Monjardet (1990)) that the *conjunction* of **$M_{\mathcal{X}}$-Independence** and **Isotony** is equivalent to the following condition:

Monotonic $M_{\mathcal{X}}$-Independence: An aggregation rule $f : X^N \to X$ is **monotonically $M_{\mathcal{X}}$-independent** if and only if for all $x_N, y_N \in X^N$ and all $m \in M_{\mathcal{X}}$: if $N_m(x_N) \subseteq N_m(y_N)$ then $f(x_N) \leqslant m$ implies $f(y_N) \leqslant m$.[14]

We are now ready to establish the following claim.

Claim 2. *$(\mathcal{P}(A), \subseteq)$, $(\mathcal{R}_A, \subseteq)$, $(\mathcal{R}_B, \subseteq)$ for any $B \subseteq A$, and*

$$\bigcup_{B \in \mathcal{P}(A)} (\mathcal{R}_B, \subseteq) \text{ are } \textit{median join-semilattices.}$$

Proof. Let us define the join of two total preorders $R, R' \in \mathcal{R}_A$ as the *transitive closure* $\overline{\cup}$ of their set-theoretic union. Then, by construction, $\mathcal{X} := (\mathcal{R}_A, \overline{\cup})$ is a *join-semilattice*, and satisfies both *upper distributivity* (by Claim (P.1) of Janowitz (1984)), and *co-coronation* (by Claims (P.3) and (P.5) of Janowitz (1984)). It follows that $(\mathcal{R}_A, \overline{\cup})$ thus defined is indeed a *median join-semilattice* (whose median ternary operation is denoted here μ'), and its meet-irreducibles are the *total preorders $R_{A_1A_2} \in \mathcal{R}_A$ having just two (non-empty) indifference classes* A_1, A_2 such that (i) (A_1, A_2) is a two-block ordered partition of A, written $(A_1, A_2) \in \Pi_A^{(2)}$, namely $A_1 \cup A_2 = A$, $A_1 \cap A_2 = \emptyset$ and (ii) $[xR_{A_1A_2}y$ and *not* $yR_{A_1A_2}x]$ if and only if $x \in A_1$ and $y \in A_2$. It can be easily checked that such total preorders $R_{A_1A_2}$ with $(A_1, A_2) \in \Pi_A^{(2)}$ are also the *co-atoms of* $(\mathcal{R}_A, \overline{\cup})$. Of course, the very same argument applies to $(\mathcal{R}_B, \overline{\cup})$, for every $B \subseteq A$. Moreover, the partially ordered set $\mathcal{X}' := (\mathcal{P}(A), \subseteq)$ of agendas is of course a bounded distributive lattice with respect to set-theoretic union $\cup$ and intersection $\cap$.

[14]The notions of $J_{\mathcal{X}}$-Independence and Monotonic $J_{\mathcal{X}}$-Independence are defined similarly by dualization for a finite median meet-semilattice $\mathcal{X} = (X, \leqslant)$ as follows: for all $x_N, y_N \in X^N$ and all $j \in J_{\mathcal{X}}$, if
$N_j(x_N) := \{i \in N : j \leqslant x_i\} \subseteq$
$\subseteq N_j(y_N) := \{i \in N : j \leqslant y_i\}$
then $j \leqslant f(x_N)$ implies $j \leqslant f(y_N)$.

Hence $(\mathcal{P}(A), \cup)$ is in particular a median join-semilattice. As a consequence, the product join-semilattice $\mathcal{X} \times \mathcal{X}' := (\mathcal{R}_A \times \mathcal{P}(A), \overline{\cup} \times \cup)$ is also a *median join-semilattice:* indeed, the ternary product-operation $\mu_{\mathcal{X}} \times \mu_{\mathcal{X}'} : (\mathcal{R}_A^T \times \mathcal{P}(A))^3 \longrightarrow \mathcal{R}_A \times \mathcal{P}(A)$ inherits the characteristic median properties $\mu(i), \mu(ii)$ (as previously defined above) from its components. Finally, $\bigcup_{B \in \mathcal{P}(A)} (\mathcal{R}_B, \subseteq)$ is a median join-semilattice with join $\overline{\cup}$, because it is a (disjoint) *sum* (or co-product) of the family $\{(\mathcal{R}_B, \subseteq)\}_{B \subseteq A}$ of median join-semilattices, and its median operation $\mu^{\sqcup}$ is defined as follows: for any $B, C, D \in \mathcal{P}(A)$, and $R^B \in \mathcal{R}_B$, $R^C \in \mathcal{R}_C$, $R^D \in \mathcal{R}_D$,

$$\mu^{\sqcup}(R^B, R^C, R^D) = (R^B \overline{\cup} R^C) \cap (R^C \overline{\cup} R^D) \cap (R^B \overline{\cup} R^D). \qquad \square$$

Therefore, in particular, $\mathcal{X}^* := (\bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B, \overline{\cup})$ is also endowed with a 'natural' metric d (namely $d = d_r = \delta_{C(\mathcal{X}^*)}$). But then, any preference relation $R \in \bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B$ induces in a 'natural' way a reflexive preference relation $\mathbf{R}_R$ on $\bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B$ which has R itself as its unique maximum and is *single-peaked* with respect to d, being defined as follows: for any $R', R'' \in (\bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B) \smallsetminus \{R\}$, $R' \mathbf{R}_R R''$ holds if and only if R' lies on a geodesic from R to R'' on the Hasse diagram $C(\mathcal{X}^*)$. Moreover, it can be shown that any such $\mathbf{R}_R$ is also *transitive*[15].

Thus, it turns out that $\bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B$ can also be 'naturally' endowed with a set $\mathcal{D}^{sp(d)} := \left\{ \mathbf{R}_R : R \in \bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B \right\}$ of *preorders* on $\bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B$ which are *single-peaked* with respect to the 'intrinsic' metric d of $\bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B$ itself.[16]

[15]See e.g. Sholander (1954), Section 3, property 3.6 for a proof.

[16]Notice that $\mathcal{D}^{sp(d)}$ includes the set $\mathcal{R}^{sp(d)}$ of all *total preorders* (hence in particular all the *linear orders*) on $\bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B$ which are single-peaked with respect to d. Moreover, $\mathcal{R}^{sp(d)}$ includes in turn the subclass $\mathcal{R}^{msp(d)}$ of all *metric* single-peaked

We are now eventually ready to provide two distinct definitions of agenda-manipulation proofness to be matched, respectively, to PAFE and SAFE social welfare functions.

Definition *Agenda Manipulation-Proofness of a PAFE social welfare function (AMP$_P$).* A PAFE social welfare function $\mathbf{f} : (\mathcal{P}(A)^N \times \mathcal{R}_A)^N \longrightarrow \mathcal{P}(A) \times \mathcal{R}_A$ with projections $\mathbf{f}_1, \mathbf{f}_2$ is AMP$_P$ if for all $i \in N$, $R_N \in \mathcal{R}_A^N$, and $B_N, B'_N \in \mathcal{P}(A)^N$ such that $C = \mathbf{f}_1(B_N, R_N) \subseteq \mathbf{f}_1(B'_N, R_N) = D$,

$\mathbf{f}_2(B_N, R_N)_{|C} \mathbf{R}_{R_i} \mathbf{f}_2(B'_N, R_N)_{|C}$ iff $\mathbf{f}_2(B'_N, R_N)_{|C} \mathbf{R}_{R_i} \mathbf{f}_2(B_N, R_N)_{|C}$. A *social welfare function* $f : \mathcal{R}_A^N \to \mathcal{R}_A$ is said to be AMP$_P$ iff it is a component of a PAFE which is AMP$_P$, namely there exist a *sovereign* agenda formation rule $f^0 : \mathcal{P}(A)^N \longrightarrow \mathcal{P}(A)$ and a PAFE social welfare function $\mathbf{f}$ which is AMP$_P$ and such that $\mathbf{f} \simeq f^0 \times f$.

In other words, AMP$_P$ requires that at every preference profile $R_N = (R_i)_{i \in N}$ on the entire set A of admissible alternatives each agent i be *indifferent* (according to her preference $\mathbf{R}_{R_i}$on preference preorders on A as induced by her actual preference R_i on A) between the *restriction* of the *social* preference $f(R_N)$ to an arbitrary agenda C, *no matter if that agenda is the actually selected agenda D or just a subagenda of D.*

Definition *Agenda Manipulation-Proofness of a SAFE social welfare function (AMP$_S$)*

A SAFE social welfare function $\widehat{\mathbf{f}} = (\widehat{f}^1, \mathcal{F}(\widehat{f}^1))$ (with

$\mathcal{F}(\widehat{f}^1) = \left\{ \widehat{f}_B : \mathcal{R}_B^N \to \mathcal{R}_B \right\}_{B \in \widehat{f}^1[\mathcal{P}(A)]}$ as defined above) is AMP$_S$ if for all $i \in N$, $R_N \in \mathcal{R}_A^N$, and $B_N, B'_N \in \mathcal{P}(A)^N$, $C, D \in \mathcal{P}(A)$ such that $C = \widehat{f}^1(B_N) \subseteq \widehat{f}^1(B'_N) = D$,

$\widehat{f}_C((R_N)_{|C}) \mathbf{R}_{R_i} (\widehat{f}_D((R_N)_{|D})_{|C}$ iff $\widehat{f}_D((R_N)_{|D})_{|C} \mathbf{R}_{R_i} \widehat{f}_C((R_N)_{|C})$.

A *social welfare function* $f : \mathcal{R}_A^N \to \mathcal{R}_A^T$ is AMP$_S$ iff it is a component of a SAFE which is AMP$_S$, namely there exist a *sovereign* agenda formation rule $f^0 : \mathcal{P}(A)^N \longrightarrow \mathcal{P}(A)$ and a SAFE social welfare function $\widehat{\mathbf{f}}$ which is AMP$_S$ and such that $\widehat{\mathbf{f}} = (f^0, \mathcal{F}(f^0))$ and $f \in \mathcal{F}(f^0)$.

Thus, AMP$_S$ requires that at every preference profile $R_N = (R_i)_{i \in N}$ on the entire set A of admissible alternatives each agent i be *indifferent* (according to her preference $\mathbf{R}_{R_i}$on preference preorders on A as

---

total preorders on $\bigcup_{B \in \mathcal{P}(A)} \mathcal{R}_B$ (namely those total preorders which are *entirely* determined by the d-distance from the peak).

induced by her actual preference R_i on A) between the *social* preference $f_C((R_N)_{|C})$ at the *restriction* of R_N to any selected agenda C, and the *restriction* to C of the *social* preference $f_D((R_N)_{|D})$ at the *restriction* of R_N to any other selected agenda $D \supseteq C$.

As mentioned above, in his classic work (Arrow (1963)) Arrow refers to the need to prevent agenda manipulation as the main argument to support the requirement of Independence of Irrelevant Alternatives for social welfare functions, that is defined as follows.

Definition *Independence of Irrelevant Alternatives (IIA).*
A social welfare function $f_A : \mathcal{R}_A^N \to \mathcal{R}_A$ satisfies IIA iff for all $R_N, R'_N \in \mathcal{R}_A^N$, and $B \in \mathcal{P}(A)$, $(f(R_N))_{|B} = (f(R'_N))_{|B}$ whenever $(R_N)_{|B} = (R'_N)_{|B}$.

Therefore, we have just introduced *three* distinct conditions that are meant to address the same problem, namely preventing agenda manipulation. A first fact about such conditions is worth mentioning at the outset: when regarded as conditions on social welfare functions both AMP_P and AMP_S only make reference to an arbitrary *single preference profile* on A, while IIA concerns an arbitrary *pair of preference profiles* on A. That contrast is quite remarkable, because reference to a *single* preference profile is a feature that seems to make full sense, in view of Arrow's overt intention to put aside all the issues related to possible strategic misrevelation of preferences. Notice, however, that in Arrow's work the notion of agenda manipulation-proofness is only introduced in a quite informal way. Accordingly, our next task is to explore the precise relationship of IIA to each one of the agenda manipulation-proofness properties introduced above. Indeed, our first finding is that IIA is *not at all* related to AMP_P.

Proposition 1. *Let $f : \mathcal{R}_A^N \to \mathcal{R}_A$ be a social welfare function and $\mathbf{f} : (\mathcal{R}_A \times \mathcal{P}(A))^N \to \mathcal{R}_A \times \mathcal{P}(A)$ a decomposable PAFE social welfare function for (N, A) such that $\mathbf{f} \simeq f^0 \times f$ where f^0 is a sovereign agenda-formation rule. Then, $\mathbf{f}$ is AMP_P (and consequently f is also AMP_P, by definition).*

Proof. Straightforward: let $R_N \in \mathcal{R}_A^N$, and $B_N, B'_N \in \mathcal{P}(A)^N$ such that $C = \mathbf{f}_1(B_N, R_N) \subseteq \mathbf{f}_1(B'_N, R_N) = D$. By decomposability of f, $\mathbf{f}_2(B_N, R_N) = \mathbf{f}_2(B'_N, R_N) = f(R_N)$. Hence, for every $i \in N$, both $\mathbf{f}_2(B_N, R_N)_{|C} \mathbf{R}_{R_i} \mathbf{f}_2(B'_N, R_N)_{|C}$ and
$\mathbf{f}_2(B'_N, R_N)_{|C} \mathbf{R}_{R_i} \mathbf{f}_2(B_N, R_N)_{|C}$ hold by reflexivity of $\mathbf{R}_{R_i}$, and the thesis follows. □

Observe that, when formally considered as a condition for a PAFE social welfare function $\mathbf{f}$, AMP_P is in fact an *interprofile* condition[17] because it involves *two* profiles $(B_N, R_N), (B'_N, R_N)$ in $(\mathcal{R}_A \times \mathcal{P}(A))^N$. However, the projection of AMP_P to the social welfare component f of $\mathbf{f}$ collapses in fact to an *intraprofile* condition since it involves a single profile $R_N \in \mathcal{R}_A^N$. Now, IIA is of course an *interprofile* condition for social welfare functions involving arbitrary *pairs* of profiles in $\mathcal{R}_A^N$. Therefore, ostensibly, AMP_P and IIA are *mutually unrelated* as conditions for social welfare functions. Our next result shows that, on the contrary, IIA is *tightly* connected to AMP_S as established by the following proposition.

Proposition 2. *Let $\widehat{\mathbf{f}} = (\widehat{f}^1, \mathcal{F}(\widehat{f}^1))$ be a SAFE social welfare function for (N, A) such that $\widehat{f}^1$ is a sovereign agenda formation rule. Then $\widehat{\mathbf{f}}$ is AMP_S if and only if $\widehat{f}_A : \mathcal{R}_A^N \to \mathcal{R}_A$ satisfies IIA.*

Proof. $\Longleftarrow$ Let $f : \mathcal{R}_A^N \to \mathcal{R}_A$ be a social welfare function for (N, A) that satisfies IIA. Then, take any sovereign agenda-formation rule $f' : \mathcal{P}(A)^N \longrightarrow \mathcal{P}(A)$ and consider SAFE social welfare function $\widehat{\mathbf{f}} = (f', \mathcal{F}(f'))$ with $\mathcal{F}(f') := \left\{f_B : \mathcal{R}_B^N \to \mathcal{R}_B\right\}_{B \subseteq A}$ defined as follows: for every $B \subseteq A$, and $R_N \in \mathcal{R}_B^N$, $f_B(R_N) := (f(R'_N))_{|B}$ where $R'_N \in \mathcal{R}_A$ is such that $(R'_N)_{|B} = R_N$. Clearly, such an f_B is well-defined precisely because f satisfies IIA. But then, take any $R_N \in \mathcal{R}_A^N$, and $B_N, B'_N \in \mathcal{P}(A)^N$, $C, D \in \mathcal{P}(A)$ such that $C = f'(B_N) \subseteq f'(B'_N) = D$, and consider $f_C((R_N)_{|C})$ and $(f_D((R_N)_{|D}))_{|C}$. By definition $f_C((R_N)_{|C}) = (f(R_N))_{|C} = (f(R_N)_{|D})_{|C} = (f_D((R_N)_{|D}))_{|C}$ whence $f_C((R_N)_{|C})\mathbf{R}_{R_i}(f_D((R_N)_{|D}))_{|C}$ iff $(f_D((R_N)_{|D}))_{|C}\mathbf{R}_{R_i}\ f_C((R_N)_{|C})$ for all $i \in N$ i.e. AMP_S holds.

$\Longrightarrow$ By contraposition. Suppose that f violates IIA. Thus, there exist $R_N, R'_N \in \mathcal{R}_A^N$ and $B \subseteq A$ such that $(R_N)_{|B} = (R'_N)_{|B}$ yet $(f(R_N))_{|B} \neq (f(R'_N))_{|B}$. Now, consider any (single-peaked) $\mathbf{R}_{R_i} \in \mathcal{D}^{sp(d)}$ which is in fact a linear order. Clearly, either

$(f(R_N))_{|B}\mathbf{R}_{R_i}(f(R'_N))_{|B}$ or $(f(R'_N))_{|B}\mathbf{R}_{R_i}(f(R_N))_{|B}$ (but *not* both of them) hold true whence AMP_S fails as required. □

Finally, we can proceed to the next main task of the present analysis, which is to explore the class of social welfare functions which are agenda manipulation-proof *and* do satisfy at least some *minimal* combination

[17]See Fishburn (1973) for a careful classification of structural, interprofile and intraprofile conditions for social welfare functions and related constructs.

of *outcome-unbiasedness* and *distributed responsiveness* to agents' preferences, as specified below.

A basic unbiasedness requirement is embodied in the standard sovereignty property, as defined below.

Sovereignty (S) A social welfare function $f : \mathcal{R}_A^N \to \mathcal{R}_A$ for (N, A) is **sovereign** if for each $R \in \mathcal{R}_A$ there exists $R_N \in \mathcal{R}_A^N$ such that $f(R_N) = R$.

A further, and weaker, unbiasedness condition is implicit in the following property.

Weak Sovereignty (WS) A social welfare function $f : \mathcal{R}_A^N \to \mathcal{R}_A$ for (N, A) is **weakly sovereign** if for any $x, y \in A$ there exists $R_N \in \mathcal{R}_A^N$ such that $xf(R_N)y$.

Clearly, WS ensures a *minimal degree of outcome-unbiasedness and responsiveness* of the relevant social welfare function, but it is consistent both with fairly distributed responsiveness-patterns involving a large number of agents, and with extremely concentrated responsiveness-patterns involving very few agents, or even just a single agent.

In order to make precise such distributed responsiveness requirement, we introduce the *responsiveness correspondence* of a social welfare function as defined below.

Responsiveness Correspondence of a social welfare function Let $f : \mathcal{R}_A^N \to \mathcal{R}_A$ be a social welfare function for (N, A). Then, its *responsiveness correspondence* $F_f : A \times A \rightrightarrows \mathcal{P}(N)$ is defined as follows: for every $x, y \in A$,

$$F_f(x,y) := \left\{ \begin{array}{r} S \subseteq N : \text{there exists } R_S^{xy} \in \mathcal{R}_A^S \text{ such that for all } R_N \in \mathcal{R}_A^N, \\ \text{if } [xR_iy \text{ iff } xR_i^{xy}y \text{ for every } i \in S] \text{ then } xf(R_N)y \end{array} \right\}.$$

Minimally Distributed Responsiveness (MDR) A social welfare function $f : \mathcal{R}_A^N \to \mathcal{R}_A$ for (N, A) satisfies **minimally distributed responsiveness** if whenever $\{i\} \in F_f(x, y)$ for some $i \in N$ and some pair of *distinct* $x, y \in A$ it must be the case that there exist $S \subseteq N \setminus \{i\}$ and $v, z \in A$, $v \neq z$ such that $S \in F_f(v, z)$.

In plain words, if the nonstrict preference of a single agent i between two distinct alternatives x, y has to be accepted as part of the social preference, then the nonstrict preference between two distinct alternatives v and z of *some other coalition not including* i is also entitled to

acceptance as part of the social preference. Thus, arguably, the *combination* of WS and MDR amounts in fact to an appropriate *minimal* requirement of *unbiased distributed responsiveness*.

Let us now recall a few (mostly classic) requirements for social welfare functions.

A social welfare function $f : \mathcal{R}_A^N \to \mathcal{R}_A$ satisfies

Anonymity (AN) if for every $R_N \in \mathcal{R}_A^N$ and every permutation σ of N, $f(x^N) = f(R_{\sigma(N)})$ (where $R_{\sigma(N)} = (R_{\sigma(1)}, ..., R_{\sigma(n)})$);

Idempotence (ID) if $f(R_N) = R$ whenever R_N is such that $R_i = R$ for each $i \in N$;

Neutrality (NT) if $[xf(R_N)y$ iff $yf(R'_N)x]$ for any $x, y \in A$ and $R_N, R'_N \in \mathcal{R}_A^N$ such that xR_iy iff yR'_ix for each $i \in N$;

Weak Neutrality (WNT) if $[f(R_N) \subseteq R$ iff $f(R_N) \subseteq R']$ for any two-indifference-class $R, R' \in \mathcal{R}_A$ and $R_N \in \mathcal{R}_A^N$ such that $R_i \subseteq R$ iff $R_i \subseteq R'$ for every $i \in N$;

Weak Pareto Principle (WP) if for every $x, y \in A$ and $R_N \in \mathcal{R}_A^N$, if $xP(R_i)y$ for every $i \in N$ then $xP(f(R_N))y$;

Basic Pareto Principle (BP) if for every $x, y \in A$ and $R_N \in \mathcal{R}_A^N$, if xR_iy for every $i \in N$ then $xf(R_N)y$;

Local Separation (LS) if for every $x, y \in A$ there exist $R_N, R'_N \in \mathcal{R}_A^N$ such that $f(R_N)_{|\{x,y\}} \neq f(R'_N)_{|\{x,y\}}$.

It should be emphasized, for future reference, that LS implies WS, while WP and LS are mutually independent.

Moreover, for any domain $\mathbf{D}$ of (preference) preorders on $\mathcal{R}_A$, a social welfare function $f : \mathcal{R}_A^N \to \mathcal{R}_A$ is **strategy-proof** on $\mathbf{D}$ iff for every $i \in N$, $\mathbf{R}_N \in \mathbf{D}^N$, $R_N \in \mathcal{R}_A^N$ and $R' \in \mathcal{R}_A$, $f(R_N)\mathbf{R}_i f((R'_i, R_{N \setminus \{i\}})$.

Moreover, a social welfare function $f : \mathcal{R}_A^N \to \mathcal{R}_A$ is said to be

dictatorial (respectively, **inversely dictatorial**) if there exists $i \in N$ such that for all $R_N \in \mathcal{R}_A^N$ and $x, y \in A$, $xf(R_N)y$ only if xR_iy (respectively, yR_ix), **weakly paretian** if it satisfies **WP,** and **weakly anti-paretian** if $yP(f(R_N))x$ for every $x, y \in A$ and $R_N \in \mathcal{R}_A^N$ such that $xP(R_i)y$ for every $i \in N$, **consensual** if it satisfies **ID**, and **properly consensual** if it satisfies **ID** and **MDR**.

The **global stalemate** social welfare function for (N, A) is the *constant* function $f^{U_A} : \mathcal{R}_A^N \to \mathcal{R}_A$ such that $f^{U_A}(R_N) = U_A$ for every $R_N \in (\mathcal{R}_A^T)^N$, where $U_A := A \times A$, the *universal indifference* relation.

We are now ready to establish which social welfare functions among those that ensure at least a modicum of unbiased distributed responsiveness do also satisfy the agenda manipulation-proofness requirements AMP_P and AMP_S, respectively.

Concerning AMP_P social welfare functions, we can rely on the following recent result (see Savaglio, Vannucci (2019, 2021)).

Theorem 1. *(Savaglio, Vannucci (2021)) Let $\mathcal{X} = (X, \leqslant)$ be a finite median join-semilattice, $M_{\mathcal{X}}$ the set of its meet-irreducible elements, B_μ its median-induced betweenness, and $f : X^N \to X$ an aggregation rule. Then, the following statements are equivalent:*

(i) f is strategy-proof on D^N for every rich domain $D \subseteq U_{B_\mu}$ of locally unimodal preorders on w.r.t. B_μ on X;

(ii) f is monotonically $M_{\mathcal{X}}$-independent;

(ii) there exists a family $\mathcal{F}_{M_{\mathcal{X}}} = \{F_m : m \in M_{\mathcal{X}}\}$ of order filters of $(\mathcal{P}(N), \subseteq)$ such that

$f(x_N) = f_{\mathcal{F}_{M_{\mathcal{X}}}}(x_N) := \bigwedge \{m \in M_{\mathcal{X}} : N_m(x_N) \in F_m\}$ for all $x_N \in X^N$.

Remark 3. *Thus, in particular, let $\mathcal{X} = (\mathcal{R}_A^T, \overline{\cup}))$ be the join-semilattice of total preorders on finite set A, μ its median ternary operation and B_μ the corresponding betweenness as previously defined, $\Pi_A^{(2)}$ the set of all total preorders on A with two indifference classes and $f : \mathcal{R}_A^N \to \mathcal{R}_A$ an aggregation rule (namely, a social welfare function) for $(N, \mathcal{R}_A)$. Then, $M_{\mathcal{X}} = \Pi_A^{(2)}$ and f is strategy-proof on D^N for every rich domain $D \subseteq U_{B_\mu}$ of locally unimodal preorders w.r.t. B_μ on $\mathcal{R}_A$ iff there exists a family $\mathcal{F}_{M_{\mathcal{X}}} = \left\{F_{A_1A_2} : (A_1, A_2) \in \Pi_A^{(2)}\right\}$ of order filters of $(\mathcal{P}(N), \subseteq)$ such that*

$$f(R_N) = f_{\mathcal{F}_{M_{\mathcal{X}}}}(R_N) := \bigwedge \{R_{A_1A_2} \in M_{\mathcal{X}} : \{i \in N : R_i \subseteq R_{A_1A_2}\} \in F_{A_1A_2}\}$$

for all $R_N \in \mathcal{R}_A^N$.

Remark 4. *Hence, the collection of such strategy-proof social welfare functions includes the following subclasses:*

- *Inclusive quorum systems , namely functions $f_{\mathcal{F}_{M_{\mathcal{X}}}}$ such that every order filter $F_{R_{A_1A_2}}$ is transversal i.e. $S \cap T \neq \varnothing$ for all $S, T \in F_{R_{A_1A_2}}$ and $\bigcup_{R_{A_1A_2} \in M_{\mathcal{X}}} F^{\min}_{R_{A_1A_2}} = N$ (observe that such a class includes any rule such that for every $R_{A_1A_2} \in M_{\mathcal{X}}$, $F_{R_{A_1A_2}}$ is simple-majority collegial i.e. there exists a minimal simple majority coalition $S_{A_1A_2} \subseteq N$, $|S_{A_1A_2}| = \left\lfloor \frac{|N|+2}{2} \right\rfloor$ with $F_{R_{A_1A_2}} = \{T \subseteq N : S_{A_1A_2} \subseteq T\}$). Generally speaking, inclusive quorum systems need not be anonymous or neutral.*

- *Outcome-biased aggregation rules, namely functions* $f_{\mathcal{F}_{M_{\mathcal{X}}}}$ *where* $F_{R_{A_1A_2}} = \varnothing$ *for some* $R_{A_1A_2} \in M_{\mathcal{X}}$ *(observe that they include the subclass of those aggregation rules such that for some total preorder* $\overline{R} \in \mathcal{R}_A$*, including possibly a linear order,* $F_{R_{A_1A_2}} = \varnothing$ *for every* $R_{A_1A_2} \in M_{\mathcal{X}}$ *such that* $\overline{R} \subseteq R_{A_1A_2}$*).*
- *Weakly-neutral aggregation rules, namely functions* $f_{\mathcal{F}_{M_{\mathcal{X}}}}$ *where* $F_{R_{A_1A_2}} = F_{R_{A'_1A'_2}}$ *whenever* $R_{A_1A_2} \wedge R_{A'_1A'_2}$ *exists.*
- *Quota aggregation rules, i.e. functions* $f_{\mathcal{F}_{M_{\mathcal{X}}}}$ *such that for each* $R_{A_1A_2} \in M_{\mathcal{X}}$ *there exists an integer* $q_{[R_{A_1A_2}]} \leq |N|$ *with*
  $F_m = \left\{T \subseteq N : q_{[R_{A_1A_2}]} \leq |T|\right\}$ *(such rules are clearly anonymous, but not necessarily weakly-neutral: they are of course weakly-neutral as well if, furthermore,* $F_{R_{A_1A_2}} = F_{R_{A'_1A'_2}}$ *whenever* $R_{A_1A_2} \wedge R_{A'_1A'_2}$ *exists). Quota aggregation rules are said to be positive if* $q_{[R_{A_1A_2}]} > 0$ *for every* $R_{A_1A_2} \in M_{\mathcal{X}}$*. The subclass of positive and weakly-neutral quota aggregation rules includes as a prominent example the co-majority social welfare function* $f^{\partial maj}$ *defined as follows: for every* $R_N \in \mathcal{R}_A$
  $f^{\partial maj}(R_N) := \wedge_{S \in \mathcal{W}^{maj}}(\vee_{i \in S} R_i)$
  where $\mathcal{W}^{maj} := \left\{S \subseteq N : |S| \geq \frac{n+1}{2}\right\}$.
- *The global stalemate social welfare function* f^{U_A} *for* (N, A) *which obtains when* $F_m = \varnothing$ *for all* $m \in M_{\mathcal{X}}$ *.*

It is worth noticing that a large subclass of such social welfare functions $f_{\mathcal{F}_{M_{\mathcal{X}}}}$ (including *positive quota aggregation rules* and *inclusive quorum systems*) satisfy the Basic Pareto Principle (BP), as made precise by the following claim.

Claim 3. *Let* $f_{\mathcal{F}_{M_{\mathcal{X}}}}$ *be a social welfare function as defined above such that* $F_{R_{A_1A_2}}$ *is a nontrivial proper order filter (i.e.* $\varnothing \notin F_{R_{A_1A_2}} \neq \varnothing$*) for every* $R_{A_1A_2} \in M_{\mathcal{X}}$*. Then* $f_{\mathcal{F}_{M_{\mathcal{X}}}}$ *satisfies BP.*

Proof. Suppose that $x, y \in A$ and $R_N \in \mathcal{R}_A^N$ are such that xR_iy for every $i \in N$*, yet not* $xf_{\mathcal{F}_{M_{\mathcal{X}}}}y$. Namely, by construction,
$(x, y) \notin \bigwedge \{R_{A_1A_2} \in M_{\mathcal{X}} : \{i \in N : R_i \subseteq R_{A_1A_2}\} \in F_{A_1A_2}\}$.
Hence, there exists $R_{A_1A_2} \in M_{\mathcal{X}}$ such that $\{i \in N : R_i \subseteq R_{A_1A_2}\} \in F_{A_1A_2}$ and $(x, y) \notin R_{A_1A_2}$. However, by assumption , $F_{A_1A_2}$ is nonempty and *every* $T \in F_{A_1A_2}$ is itself nonempty: thus, $N \in F_{A_1A_2}$. But then $(x, y) \in R_i \subseteq R_{A_1A_2}$ for any $i \in T$, a contradiction. □

Remark 5. *It should be emphasized that BP and WP are independent alternative ways of weakening the (strong) Pareto principle*[18]*. To see that, just consider social welfare functions* f^{UN} *and* $f^{\mathbf{L}x^*}$ *for* (N, A) *defined informally as follows: if* R_N *is such that* $R_i = R$ *for each* $i \in N$ *then* $f^{UN}(R_N) = R$*, otherwise* $f^{UN}(R_N)$ *is the universal indifference relation on* A *while* $f^{\mathbf{L}x^*}(R_N)$ *-where* $\mathbf{L}$ *is a linear order of* $\mathcal{R}_A$*- is the* $\mathbf{L}$*-minimum linear order* L *of* A *having* x^* *as its top element and such that* $L \supseteq \cap_{i\in N} P(R_i)$ *if there is no* y *such that* $(y, x^*) \in \cap_{i\in N} P(R_i)$*, and the* $\mathbf{L}$*-minimum linear order* L *of* A *such that* $L \supseteq \cap_{i\in N} P(R_i)$ *otherwise. Clearly, neither of them satisfy the (strong) Pareto principle: however,* f^{UN} *satisfies BP and violates WP while* $f^{\mathbf{L}x^*}$ *satisfies WP and violates BP. Nevertheless, WP has been widely used in the extant literature, whereas BP has been rarely if ever explicitly employed.*

Proposition 3. *There exist social welfare functions* $f : \mathcal{R}_A^N \to \mathcal{R}_A$ *which satisfy* AMP_P*, AN, ID, WNT, BP and are strategy-proof on the domain* $\mathcal{D}^{sp(d)}$ *of single-peaked preorders on* $\bigcup_{B\in\mathcal{P}(A)} \mathcal{R}_B$ *.*

Proof. Since $(\bigcup_{B\in\mathcal{P}(A)} \mathcal{R}_B, \subseteq)$ is a median join-semilattices by Claim 2 above, it follows that Theorem 1 above applies hence positive weakly-neutral quota social welfare functions as defined above satisfy AN, ID, WNT and are strategy-proof on the domain $\mathcal{D}^{sp(d)}$ of single-peaked *preorders* on $\bigcup_{B\in\mathcal{P}(A)} \mathcal{R}_B$. Moreover, they also satisfy AMP_P by Claim 1, and BP by Claim 3, and the thesis is established. □

Remark 6. *Observe that several domains of preference profiles are being considered here. The first one consists of profiles* $R_N = (R_i)_{i\in N}$ *of arbitrary total preorders on the set* A *of basic alternatives. The second domain consist of profiles* $\mathbf{R}_N = (\mathbf{R}_i)_{i\in N}$ *of single-peaked (partial) preorders on the ground set* $\mathcal{R}_A$ *of the median join-semilattice of total preorders on* A *(with single-peakedness induced by the median betweenness of* $\mathcal{R}_A$*, and* $\mathbf{R}_N$ *induced by* R_N*). The third domain amounts to the subdomain of the second one which only includes the metric single-peaked profiles* $\widehat{\mathbf{R}}_N = (\widehat{\mathbf{R}}_i)_{i\in N}$ *of preorders on* $\mathcal{R}_A$ *that are entirely*

[18]A social welfare function $f : \mathcal{R}_A^N \longrightarrow \mathcal{R}_A$ satisfies the *(strong) Pareto principle* iff $xP(f(R_N))y$ for any $x, y \in A$ and $R_N \in \mathcal{R}_A^N$ such that xR_iy for every $i \in N$, and $xP(R_j)y$ for some $j \in N$.

determined by profiles R_N through their graphic distance from the top. Accordingly, we can consider different notions of WP and BP: namely,

(1) WP (BP) of each 'social preference' $f(R_N)$ with respect to R_N: the requirement that, at any R_N, $f(R_N)$ should faithfully reflect any unanimous preference for an alternative in A over another one;

(2) WP (BP) of each 'social preference' $f(R_N)$ with respect to $\mathbf{R}_N$ (or $\widehat{\mathbf{R}}_N$): the requirement that, at any R_N, $f(R_N)$ should be consistent with unanimous preferences according to $\mathbf{R}_N$ (or $\widehat{\mathbf{R}}_N$), which means that there should be no alternative 'social preference' $R' \in \mathcal{R}_A$ that is unanimously preferred over $f(R_N)$ according to $\mathbf{R}_N$ (or $\widehat{\mathbf{R}}_N$).

Two most remarkable points are to be made here concerning the social welfare functions mentioned in the previous Proposition. First, such social welfare functions fail to satisfy WP with respect to the first and second domains, consisting respectively of arbitrary profiles R_N of total preorders on A, and of single-peaked profiles $\mathbf{R}_N$ of preorders on $\mathcal{R}_A$. Second, the very same social welfare functions do satisfy WP with respect to the domain consisting of metric single-peaked profiles $\widehat{\mathbf{R}}_N$ of total preorders on $\mathcal{R}_A$.

Remark 7. *It is worth noticing that all of the anonymous, idempotent and strategy-proof social welfare functions mentioned in the previous proposition (including those which satisfy the Basic Pareto Principle BP e.g. positive quota social welfare functions) admit a stalemate as one of the possible outcomes, arising from certain specific patterns of strong conflict among individual preferences. By definition, such (contingent) stalemates give rise to violations of the Weak Pareto principle (WP) by the chosen 'social preference' both with respect to the 'basic' preference profiles R_N of total preorders on the set A of alternatives, and with respect to general single-peaked domains $\mathbf{R}_N$ on $\mathcal{R}_A$ induced by the former R_N profiles. Thus, the foregoing social welfare functions may also be regarded as valuable sources of information and advice concerning the 'general interest' (or 'common good'). In many cases, they provide an explicit description of the alternatives that best represent the 'common good', or define anyway clear improvements on the status quo. But occasionally they may also help to pursue the 'general interest' by pointing to situations of pathologically strong social conflict: they do that precisely by returning outcomes that allow for 'inefficient' choices when fed with inputs encoding such a sort of social conflict*[19]*. To put*

[19]See also Saari (2008) on the connection between conflict, cycling and inefficiency.

it in other terms, any violation of WP by such social welfare functions might be regarded as a sort of 'error message' calling for public intervention (e.g. promoting an improved access to key relevant information for the general public, implementing some appropriate redistribution policies, or just relying on some contingent agenda manipulation activities of the sort thoroughly analyzed and discussed in Schwartz (1986)[20] *in order to ensure outcome-efficiency).*

Concerning the study of AMP_S social welfare functions we can rely on the following well-known results due to Wilson (1972) and Hansson (1973)[21], respectively.

Theorem 2. *(i)(Wilson (1972)) Let $f : \mathcal{R}_A^N \to \mathcal{R}_A$ be a social welfare function which satisfies IIA and WS. Then, f is either dictatorial, or inversely dictatorial, or else $f = f^{U_A}$ i.e. f is the global stalemate social welfare function for (N, A);*

(ii) (Hansson (1973)) Let $f : \mathcal{R}_A^N \to \mathcal{R}_A$ be a social welfare function which satisfies IIA and WS. Then, f is either dictatorial, or inversely dictatorial, or else it violates LS.

The following alternative characterization of the global stalemate social welfare function combines the AMP_S formulation of IIA with two weak conditions following from anonymity and neutrality such as WS and MDR to the effect of emphasizing the inordinate strength of AMP_S (or IIA).

[20]Notice, however, that in IIA-based models of collective choice as advocated by Schwartz (1986) such agenda-structure manipulation activities are treated as *normal* and endemic to every democratic aggregation protocol. Of course, that is pretty much the same conclusion as that typically suggested by authors that regard Arrow's theorem as an indictment of democratic preference aggregation protocols, à la Riker (1982). By contrast, within the IIA-free models considered in the present work, such agenda-structure manipulation processes can (and should) be considered as *local, contingent* subroutines appended to general democratic aggregation protocols in order to increase their effectiveness to cope with certain *specific* sorts of conflicts related to Condorcet cycles.

[21]It should be emphasized that Hansson's theorem (which was established independently of Wilson Theorem) amounts to replacing the 'global stalemate' clause of the Wilson Theorem with a weaker clause (violation of the LS condition i.e. of 'Strong Non-Constancy' in Hansson's own original terminology). Incidentally, a close inspection of Hansson's proof shows that it can also be deployed to imply the stronger Wilson's 'global stalemate' clause. See also Malawski, Zhou (1994) and Cato (2012) for related work on preference aggregation without WP.

Proposition 4. *A social welfare function $f : \mathcal{R}_A^N \to \mathcal{R}_A^T$ satisfies AMP$_S$, WS and MDR if and only if f is the global stalemate social welfare function i.e. $f = f^{U_A}$.*

Proof. $\Longrightarrow$ Suppose that social welfare function f satisfies AMP$_S$, WS and MDR. Hence f also satisfies IIA by Proposition 2. But then, it follows from Wilson's Theorem as mentioned above that f is dictatorial, inversely dictatorial, or the global stalemate constant function f^{U_A}. However, a dictatorial social welfare function clearly violates MDR: indeed, suppose $i \in N$ is such that, for every $x, y \in A$ and $R_N \in \mathcal{R}_A^N$, $xf(R_N)y$ entails xR_iy. Moreover, by WS, for every $x, y \in A$ there exists $R_N \in \mathcal{R}_A^N$ such that $xf(R_N)y$, whence xR_iy holds. Then, by MDR there exists $S \subseteq N \setminus \{i\}$ and a pair of distinct $v, z \in A$ such that $S \in F_f(v, z)$ i.e. there exists $R_S^{vz} \in \mathcal{R}_A^S$ such that $vf(R_N)z$ for every $R_N \in \mathcal{R}_A^N$ with $R_{S|\{v,z\}} = R_{S|\{v,z\}}^{vz}$. Now, consider a profile $R_N \in \mathcal{R}_A^N$ such that zR_iv, *not* vR_iz and $R_{S|\{v,z\}} = R_{S|\{v,z\}}^{vz}$. By definition of F_f, $vf(R_N)z$. However, since f is dictatorial, *not* vR_iz implies *not* $vf(R_N)z$, a contradiction. Thus, f is *not* dictatorial, as required.

Similarly, suppose there exists $i \in N$ such that for every $x, y \in A$ and $R_N \in \mathcal{R}_A^N$, $xf(R_N)y$ entails yR_ix. Again, it follows from WS that for every $x, y \in A$ there exists $R_N \in \mathcal{R}_A^N$ such that $xf(R_N)y$, whence yR_ix holds, by our assumption. Then, by MDR there exists $S \subseteq N \setminus \{i\}$ and a pair of distinct $v, z \in A$ such that $S \in F_f(v, z)$. Now, consider a profile $R_N \in \mathcal{R}_A^N$ such that vR_iz, *not* zR_iv and $R_{S|\{v,z\}} = R_{S|\{v,z\}}^{vz}$. By definition of F_f, $vf(R_N)z$. However, by assumption, *not* zR_iv implies *not* $vf(R_N)z$, a contradiction. Thus, f is *not* inversely dictatorial either. Therefore, it follows from Wilson's Theorem that $f = f^{U_A}$, the global stalemate function.

$\Longleftarrow$ It can be easily shown that the global stalemate social welfare function f^{U_A} satisfies AMP$_S$, WS and MDR. Indeed, take any sovereign agenda formation rule f, posit $\mathcal{F}(f) := \left\{f_B := f^{U_B}\right\}_{B \subseteq A}$, and consider the corresponding PAFE social welfare function $\mathbf{f} = (f, \mathcal{F}(f))$. By definition, $f_A := f^{U_A}$ which obviously satisfies IIA, being a constant function defined on $\mathcal{R}_A^N$. Thus, by Proposition 2, $\mathbf{f}$ satisfies AMP$_S$ and consequently, by definition, $f^{\ U_A}$ also satisfies AMP$_S$. Moreover, for any $x, y \in A$ and $R_N \in \mathcal{R}_A^N$, $xf^{U_A}(R_N)y$ hence WS is trivially satisfied by f^{U_A}. Finally, observe that the responsiveness correspondence $F_{f^{U_A}}$ is such that $F_{f^{U_A}}(x, y) = \mathcal{P}(N)$ for all $x, y \in A$. But then, for any $x, y \in A$ and any $i, j \in N$, $\{N, \{i\}, \{j\}\} \subseteq F_{f^{U_A}}(x, y)$. It follows that f^{U_A} also satisfies MDR. $\square$

Remark 8. *Notice that AN and NT do indeed imply WS and MDR, while the converse does not hold: to see this, consider the social welfare function f^* such that for some $1, 2, 3 \in N$, and for every $x, y \in A$, $R_N \in \mathcal{R}_A^N$, $xf^*(R_N)y$ iff either $xR_{\{1,2\}}y$ or [not $xR_{\{1,2\}}y$ and xR_3y]. Then, in view of the equivalence of AMP_S and IIA established by Proposition 2, it follows that Proposition 4 amounts to an extension of Hansson's characterization of the Global Stalemate social welfare function f^{U_A} via AN, NT and IIA (Hansson(1969a)).*

Corollary 1. *There is no social welfare function $f : \mathcal{R}_A^N \to \mathcal{R}_A^T$ that satisfies AMP_S, S and MDR. Thus, in particular, there is no idempotent social welfare function that satisfies AMP_S and MDR.*

Proof. Suppose that on the contrary there exists a social welfare function f which satisfies AMP_S, S, and MDR. Since S clearly implies WS, it follows from Proposition 4 above that $f = f^{U_A}$, a contradiction because by definition f^{U_A}does *not* satisfy S. The second statements follows trivially since any idempotent social welfare function does satisfy S. □

So, we have an impossibility result that follows just from the combination of AMP_S and MDR, with no role at all for the Weak Pareto Principle.

Furthermore, we also have the following two characterizations of *dictatorial* social welfare functions. The first one is of course just a reformulation of Arrow's 'general possibility theorem' as an immediate consequence of Proposition 2 and the abovementioned Wilson's Theorem. The second one is essentially a similar reformulation of Hansson's Non-Constancy Theorem (Hansson (1973)) as presented above.

Corollary 2. *(i) (Arrow's Theorem) A social welfare function $f : \mathcal{R}_A^N \to \mathcal{R}_A$ satisfies AMP_S and WP if and only if f is dictatorial;*

(ii) (Hansson' s Non-Constancy Theorem) A social welfare function $f : \mathcal{R}_A^N \to \mathcal{R}_A$ satisfies AMP_S, LS and is ***not*** *inversely dictatorial if and only if f is dictatorial.*

Proof. (i) Immediate. In view of Proposition 2, f is AMP_S if and only if it satisfies IIA. Now, if f satisfies WP then it is not inversely dictatorial, and it can't be the Global Stalemate social welfare function. Hence, by Wilson's Theorem as introduced above it must be a dictatorial social welfare function. Conversely, if f is dictatorial then it clearly satisfies WP and IIA (and thus AMP_S, by Proposition 2).

(ii) Also immediate, in view of Proposition 2, Hansson's Theorem, and the simple observation that a dictatorial social welfare function is clearly not inversely dictatorial, and satisfies both IIA and LS. □

Remark 9. *Notice that the original proof of Hansson's Non-Constancy Theorem in Hansson (1973) relies in fact on Arrow's Theorem. Moreover, a careful inspection of that proof makes it clear that the only social welfare function that is neither dictatorial nor inversely dictatorial and satisfies AMP_S is the Global Stalemate function. In other terms, Hansson's proof shows that Arrow's Theorem implies Wilson's Theorem (it should be recalled here that Hansson's Non-Constancy Theorem was first published in a 1972 working paper, independently of Wilson's Theorem). But then, since Wilson's Theorem obviously implies Arrow's Theorem, the foregoing Corollary confirms that the two of them are in fact equivalent.*

It should also be mentioned that, relying on other well-known results from the extant literature, further elaborations on the role of AMP_S established by the foregoing propositions can be easily produced. For instance, a further result in Wilson (1972) shows that *any* social welfare function which satisfies IIA (hence AMP_S by Proposition 2) must produce 'social preferences' invariably composed by *some* combination of *at most five* different types of patches corresponding respectively to 'locally imposed strict preferences', 'minimal (local) stalemates', 'non-minimal (local) stalemates', 'locally dictatorial preferences' and 'locally inversely-dictatorial preferences'[22]. Furthermore, it is also well-known that there exists a quite general model-theoretic rationale underlying such results (see e.g. Lauwers, Van Liedekerke (1995) for details). Namely, it is sufficient to join *either* IIA and the Weak Pareto Principle (WP) *or* IIA and Weak Sovereignty (WS) to force the set of *decisive*

[22]See Wilson (1972), Theorem 5. Binmore (1976) is an interesting further extension of that theorem, showing that in order to avoid the grim consequences of the latter the domain of a social welfare function which satisfies IIA should be dramatically restricted. Specifically, its domain should not include *all* the preference profiles of total preorders consistent with *at least one* arbitrarily fixed 'party structure' (namely, a partition of agents into 'parties' defined as sets of agents whose preferences are al least partially concordant on *every* pair of alternatives).

coalitions[23] of a social welfare function f for (N, A) to be an *ultrafilter*[24] on N, a fact which in turn implies that f is *dictatorial* since N is by assumption finite.

Thus, it is abundantly clear that AMP_S (or IIA) is a powerful obstruction to *each one* of the following basic requirements for any properly consensus-based social welfare function, namely (i) *weak Pareto-optimality*, (ii) *sovereignty, and* (iii) *minimally distributed responsiveness*. The inescapable conclusion is that it is *precisely* the attempt to prevent agenda manipulation of a social welfare function via AMP_S that results in the classic limitative or 'impossibility' theorems. Hence, insisting on AMP_S (i.e. on IIA) to ensure agenda manipulation-proofness gives rise *by itself* to a bleak scenario concerning the construction of properly consensual social welfare functions.

Let us then briefly summarize the widely shared views on the import of Arrow's theorem which typically follow from a firm endorsement of IIA as grounded on the assumption that IIA amounts to a key requirement for *any* reasonable voting rule in order to prevent agenda manipulation. Since Arrow's IIA is a property shared by several commonly used voting rules including the simple majority rule (arguably, a paragon of 'democratic' voting rules), it seems to follows that Arrow's theorem does validate the following challenging, momentous statement. Namely, the assertion that *any attempt* to use *democratic* voting rules to articulate a *consistent formulation of the collective interest* with a view to identify and select policies which best promote it *is doomed to failure*. That is so precisely because Arrow's theorem shows that under IIA (*weak) Pareto optimality can only be achieved through dictatorship*. Therefore, since dictatorship is obviously to be rejected as a means to define proper consensus-based 'social preferences', insisting to prevent agenda manipulation (specifically, agenda-content manipulation) entails reliance on some aggregation rule which might license 'social preferences' that reverse unanimously held individual strict preferences between some pairs of alternatives. And that is also deemed to be not acceptable. But then, the only alternative left is *to allow for agenda manipulation* (specifically, agenda-content manipulation) to

[23]A *decisive coalition* of a social welfare function f is any coalition $C \subseteq N$ that can enforce the unanimous preference of its members between each pair of alternatives as the actual social preference.

[24]An *ultrafilter* (or *maximal lattice-filter* on N) is a nonempty set $\mathcal{F} \subseteq \mathcal{P}(N) \setminus \{\varnothing\}$ such that for every $C, D \in \mathcal{F}$: (i) $C \cap D \in \mathcal{F}$ and (ii) either $C \in \mathcal{F}$ or $N \setminus C \in \mathcal{F}$. Since N is finite, every ultrafilter $\mathcal{F}$ on N is *principal*, namely $\mathcal{F} = \{C \subseteq N : i \in C\}$ for some $i \in N$. It follows that $\{i\}$ is a decisive coalition for f, and consequently f is a *dictatorial* social welfare function.

the effect of undermining *reliability* of the aggregation rule, since the representation of the 'general interest' provided by such a rule might typically reflect just successful manipulatory activities, and possibly nothing else. Either way, the aim of producing a consistent, faithful and credible formulation of the 'general interest' cannot be apparently fulfilled. To put it bluntly, majority voting cannot be relied upon to discover the public interest or 'general will' because of its possible cycles, and *nothing else can work* because of the same vulnerability to agenda manipulation. As a consequence, in actual practice there is not such a thing as a 'general interest' to discover, formulate and implement as a guide or benchmark for public policy. It also follows, accordingly, that there is no way to feed work in mechanism design and institutional design aimed at improving the effectiveness of democratic institutions with well-grounded and reliable criteria summarizing the 'general interest'.[25]

That scenario has been variously described as the impossibility of rational collective decisions[26], or the impossibility of a reliable and significant consensus-based expression of the 'collective good' (or 'general

[25]Riker (1982) is probably the most outspoken and consistent presentation of such a view available in print. While Dahl (1956) tentatively labeled 'populist democracy' the doctrine that identifies the exercise of sovereignty with *exclusive and unrestrained reliance on majority voting*, Riker assumes that Arrow's theorem licences the imputation of the same, allegedly hopeless, limitations of the majority rule to *every possible 'democratic' preference aggregation rule*. Consequently, virtually all the social-choice-theoretic work in mechanism design is uncerimoniously identified with 'populism' and dismissed as a hopeless endeavour. Riker's suggested and clearly preferred alternative is acceptance of the 'Schumpeterian' view that modern 'democratic politics' is -and *has to be*- nothing else than (i) competitive electoral selection of the ruling elite, and (ii) pervasive and relentless activities of agenda manipulation on the part of elected officials and representatives, in view of more or less special interests, with no effective role left for a public representation of the common interest as a shared, consensus-based benchmark. Accordingly, analyzing 'democratic politics' from such a perspective is regarded as the main task of 'democratic theory' proper, and the plain endorsement of that view is the defining feature of 'liberalism'(see Riker (1982), and Schofield (1985) for a thorough technical treatment of a wide array of social choice theoretic models that shares at least some of Riker's views, while refraining from the latter's most ideological overtones).

[26]See e.g. Buchanan (1954) for an early example of that view, assuming the alleged impossibility of collective decisions to replicate the 'rationality' of individual decisions. A more balanced view, advocating a combination of IIA with less demanding criteria of 'rationality' for *both* collective and individual decisions is advanced by Schwartz (1986).

interest' or 'general will'[27]). This is in fact the most familiar understanding of Arrow's 'impossibility theorem', an interpretation which projects a 'dark' view of its content and significance concerning the viability and effectiveness of democratic protocols. That is so because such an interpretation suggests that there is no way to use voting methods, decision systems or preference aggregation rules of any sort to help improving the effectiveness and deliberative quality of current democratic protocols[28]. However, again, all of the above rests crucially on the understanding that IIA is both reasonable and virtually inescapable[29]. But then , Propositions 1 and 3 show that *there is in fact an alternative way to achieve agenda manipulation-proofness via* AMP_P: such an alternative makes it possible to devise anonymous and idempotent social welfare functions that satisfy a basic version of the Pareto principle, and are -in a compelling sense- *both agenda manipulation-proof and strategy-proof* . Thus, there is indeed an effective way out of the strictures identified by Arrow's theorem. From that perspective, Arrow's result actually provides *constructive* information about the design of social welfare functions and preference aggregation rules: in that sense, there is also a *bright side* of Arrow's theorem. Access to the latter requires three basic steps:

(i) reliance on (possibly *redundant*) *preference elicitation concerning an entire set of prefixed admissible alternatives* in order to ensure *agenda manipulation-proofness without any recourse to IIA*;

(ii) a *mild relaxation of the Pareto Principle to BP allowing for occasional stalemates* (namely, social indifference over a set of alternatives including Pareto-dominated outcomes), and a concurrent reinterpretation of possible violations of WP and other, stronger, versions of the Pareto principle as 'warning signals' pointing to the need for remedial actions including policies to correct blatant disparities of access to information and/or other key resources;

(iii) refocussing on *a further condition ('monotonic* $M_{\mathcal{X}}$*-independence')* in order to address strategic manipulation issues: such a condition can be regarded as a combination of a mild monotonicity property and a considerably *weakened* version of IIA.

Broadly speaking, some form of each one of the foregoing steps was previously considered or at least evoked in the extant literature (a

[27]See e.g. the highly influential Riker (1982) as discussed above (note 25).

[28]An influential tentative list of basic, *substantive* requirements for democratic decision protocols including 'political equality', 'deliberation', 'participation', and 'agenda control' is due to Dahl (1979).

[29]See for instance Schwartz (1986) p.33 for a remarkably clear, adamant endorsement of that view.

detailed discussion of that matter is provided below in the next section of the present work). What is new here is, arguably, *their joint consideration as made possible by a model that combines agenda formation and preference aggregation.* The resulting analysis shows that a sound *parallel-coupling* of social welfare functions to their own agenda-formation processes makes it possible to jointly achieve anonymity, idempotence, agenda manipulation-proofness and a very basic form of Pareto efficiency, together with strategy-proofness on a suitably large and natural domain of single-peaked 'meta-preferences'(or 'preferences on preferences'). As observed above, it is also remarkable that the latter strategy-proofness property turns out to be essentially equivalent to an 'independence' condition which amounts to a *considerable weakening of IIA.*

3. Related work

As mentioned in the introduction, agenda manipulation-proofness was used by Arrow as the main motivation for introducing IIA as a basic requirement for social welfare functions in his seminal work (Arrow (1963)). Since then, it has become quite common to use Arrowian as a qualifier for social welfare functions or aggregation rules which satisfy some version of 'independence of irrelevant alternatives' (and possibly some further basic requirement such as idempotence i.e. 'respect for unanimity')[30]. Hence, the amount of literature which is broadly related to the topics covered by the present paper is simply enormous. Therefore, we shall confine the ensuing review to the most strictly relevant previous contributions, collecting them in two distinct subsections that correspond to two focal points of the present analysis which are both related to IIA, namely 'agenda manipulation-proofness and IIA' and 'weakening IIA and strategy-proofness'.

[30]See, among many others, Aleskerov (1999), Sethuraman, Teo, Vohra (2003), Nehring, Puppe (2010). An alternative (and perhaps more appropriate) usage is the one that rather contrasts Arrowian (or multi-profile) and Bergson-Samuelson (or single-profile) social welfare functions, and goes as follows. Let N, A be two (finite) sets and $\mathcal{R}_A$ the set of all total preorders on A. An *Arrowian social welfare function* for (N, A) is a function

$f : \mathcal{R}_A^N \to \mathcal{R}_A$,

while a *Bergson-Samuelson social welfare function* for (N, A) is a function

$f : \{r^N\} \to \mathcal{R}_A$, with $r^N \in \mathcal{R}_A$.

Moreover, their *strict* counterparts are obtained by replacing $\mathcal{R}_A$ with $\mathcal{L}_A$ i.e. the set of all *linear orders* (namely, *antisymmetric* total preorders) on A.

(I) *Agenda manipulation-proofness and IIA.*

The role of preference elicitation on the *entire* set of admissible alternatives in order to ensure transitivity properties of the 'social preferences' and the resulting violation of IIA has been repeatedly pointed out, and contrasted with *peacemeal* elicitation of preferences on specific agendas of admissible alternatives, which is conducive to IIA-consistency and violation of transitivity properties of 'social preferences'(see e.g. Sen (1977) where that contrast is discussed with reference to several versions of the simple majority rule and the Borda Count scoring rule). Unfortunately, preference elicitation and agenda formation are typically *not* modelled together in the extant literature: specifically, agenda manipulation is usually left unmodelled and thus given a quite informal treatment[31]. As a consequence, even the agenda-content and agenda-structure dimensions of agenda manipulation are typically *not* neatly and consistently distinguished. Thus, following the lead of a much discussed, misleading example proposed in Arrow (1963), use of the label 'IIA' has also been occasionally stretched to also refer (improperly) to requirements on social preference rankings on a certain subset of alternatives across several *distinct* admissible agendas for a *fixed* profile of individual preferences on the largest admissible agenda (see e.g. Ray (1973), Fishburn (1973), Sen (1977), Schwartz (1986), Young (1995)[32] for discussions of such topic). Furthermore, without an explicit joint modelling of agenda formation and preference

[31]A partial exception due to Dietrich (2016) is available in the related framework of *judgment aggregation* (to be discussed below) where agenda manipulation is modeled as sensitivity of aggregate judgments on issues to agenda-content alterations (*including expansions*), with no explicit role for preferences. Notably, that notion of agenda manipulation-proofness is shown to be tightly connected to a version of IIA, and results in a characterization of dictatorial judgment aggregation rules when combined with a unanimity-respecting condition for general finite agendas.

[32]Actually, Young (1995) also discusses at length a condition he calls 'local stability' or *local independence of irrelevant alternatives (LIIA)* which is also satisfied *by median-based aggregation rules* (but not by positionalist rules such as the Borda Count). When applied to a social welfare function $f : \mathcal{R}_A^N \to \mathcal{R}_A$ LIIA may be formulated as follows:

$f((R_N)_{|B}) = (f(R_N))_{|B}$ for all $R_N \in \mathcal{R}_A$, and all $B \in \mathcal{I}_{f(R_N)}$ where, for any $R \in \mathcal{R}_A$,

$$\mathcal{I}_R := \left\{ \begin{array}{c} I \subseteq A : I = \{z : xRzRy\} \\ \text{for some } x, y \in A \end{array} \right\}.$$

Thus LIIA is in fact an *intraprofile* property, rather than an *interprofile* property like IIA and its relaxed versions (again, see Fishburn (1973) for a classic, exhaustive classification of standard social choice-theoretic properties for preference aggregation rules).

elicitation it is virtually impossible to distinguish not only between parallel coupling and sequential coupling of those two processes, but also between preference-first and agenda-first sequential coupling. In the previous section of the present work it has been shown that parallel-coupling allows for agenda manipulation-proofness without IIA, while under agenda-first sequential coupling agenda manipulation-proofness does in fact *amount to IIA.* But then, what about *preference-first sequential coupling* of agenda formation and preference elicitation? In that connection, the main theorem of Hansson (1969b) concerning *generalized social choice correspondences (GSCCs)*[33], and its reformulation and extension due to Denicolò (2000) are indeed relevant and most helpful. In fact, Hansson's result relies on an extended (indeed, *strenghtened*) version of IIA for GSCCs that are not required to be generated through maximization of the total preorders which represent social preferences. Specifically, it implies that any *social choice correspondence* F on a set A (with $|A| \geq 3$) that satisfies WP and such an extended IIA property can be represented as the choice of maxima of the social preferences in the range of a social welfare function f which satisfies IIA and WP if and only if both F and f are dictatorial (see Hansson (1969b), Theorem 3 [34]). Accordingly, it can also be shown that agenda manipulation-proofness of a social welfare function is in fact equivalent to IIA *even under preference-first sequential coupling* of agenda formation and preference elicitation[35].

(II) *Weakening IIA and strategy-proofness.* The other major theme in the present work is that, once agenda manipulation-proofness of properly consensus-based social welfare function is secured through parallel-coupling of agenda formation and preference elicitation (with no role at all for IIA), the strategy-proofness issue for such social welfare

[33]Or 'group decision functions' in the original terminology of Hansson (1969b).

A *generalized social choice correspondence* for (N, A) is a function

$f : \mathcal{R}_A^N \longrightarrow \mathcal{C}_\mathcal{A}$ where $\mathcal{A} \subseteq \mathcal{P}(A) \setminus \{\emptyset\}$ with $A \in \mathcal{A}$ and $\mathcal{C}_\mathcal{A}$ is the set of all functions $C : \mathcal{A} \longrightarrow \mathcal{P}(A) \setminus \{\emptyset\}$ such that $C(B) \subseteq B$ for every $B \in \mathcal{A}$.

A *social choice correspondence* for (N, A) is a function $f : \mathcal{R}_A^N \longrightarrow \mathcal{P}(A) \setminus \{\emptyset\}$ i.e. a generalized social choice correspondence such that $\mathcal{A} = \{A\}$.

A social choice correspondence for (N, A) whose range consists of *singleton-sets* is also said to be a *social choice function,* and usually written $f : \mathcal{R}_A^N \longrightarrow A$.

[34]See Denicolò (2000) for a simplified presentation of Hansson's theorem, and a detailed formulation of its consequences for social choice correspondences and social welfare functions as just mentioned in the text.

[35]The proof is along the same lines of the proof of Proposition 2. Details are available from the author upon request.

functions does also admit a sensible formulation and a positive solution. Specifically, the latter requires just (a) focussing on the 'right' individual preferences (which *must* be preferences on the outcomes of a social welfare function, hence *preferences on social preferences over outcomes* i.e. ultimately 'meta-preferences on basic preferences') and (b) observing that *basic preferences on alternatives induce in a natural way single-peaked 'meta-preferences' on the 'preference space' which in turn ensure strategy-proofness of the properly consensus-based social welfare functions* mentioned above. Moreover, it turns out that in such a setting strategy-proofness is in fact equivalent to the combination of a very mild *monotonicity* condition on the influence of coalitions (namely the requirement that adding support to a previously positive decision on a certain binary issue should never result in a decision reversal) and an *independence* condition that amounts to a *much weakened version of IIA.*

Thus, in a sense, a certain version of IIA ultimately reenters the picture but (i) in a *much weakened* and *very specific* form [36] and (ii) with reference to *strategy-proofness*, an issue that (as opposed to agenda manipulation-proofness) was explicitly *put aside* in the original Arrowian analysis of social welfare functions (see Arrow (1963), p.7).

Now, both of those tenets run counter to some views that are apparently still widely held in the literature, and to which we now turn. To begin with, the exceptional strength of IIA is sometimes downplayed or in any case not fully appreciated. One reason for that may be the (correct) perception of the relationship of IIA to agenda-content manipulation-proofness as combined with the (incorrect) view that sequential-coupling of agenda formation to preference elicitation is the only available possibility[37]. It is also possibly the case that IIA is occasionally confused with its earlier counterpart named 'Postulate of

[36]To be sure, it is also well-known that IIA is *so strong* that there are also weakened versions of IIA which imply *dictatorship* for preference aggregation rules when coupled with the Weak Pareto condition or indeed *any non-constancy constraint.* Notice that this fact holds not only for social welfare functions, but also for preference aggregation rules admitting *any total binary relation* as their output (with no transitivity requirement at all!). That is the case of so-called *Independent Decisiveness* of aggregation rule f requiring that any coalition which happens to be able to enforce its strict preference over a certain ordered pair of alternatives (x, y) for *some* preference profile no matter what the preferences of others over x, y happen to be must also be *decisive* for (x, y): see Sen (1993), Denicolò (1998), Quesada (2002). For another example of an Arrow-like theorem for aggregation rules in the same vein (albeit in a much more general setting) see Daniëls, Pacuit (2008).

[37]Indeed, there is arguably no other way to make full sense of the following statement from a well-known and highly respected scholar: 'Independence of Irrelevant

Relevancy' which is due to Huntington (1938), and is explicitly quoted by Arrow himself as a source of inspiration and 'a condition analogous to' IIA (Arrow (1963), p. 27). Notice, however, that while Huntington's 'Postulate of Relevancy' may well be quite similar in spirit to IIA, it is in fact *much weaker* than the latter because it relies on a *common language of linearly ordered grades* (indeed, numbers[38]) to express *absolute judgments* (as opposed to merely comparative ones)[39]. A third line of reasoning in support of IIA originates from a misleading interpretation of a well-known theorem due to Satterthwaite (1975) that establishes a tight connection between strategy-proof *strict* social choice functions and *strict* social welfare functions that satisfy IIA[40]. Indeed, Satterthwaite himself claims that such a theorem 'creates a strong new justification for [WP and] IIA as conditions that an ideal social welfare function should satisfy' (Satterthwaite (1975), p. 207, with some minor editing of mine)[41]. Notice, however, that the 'IIA-nonmanipulability' connection identified and discussed by Satterthwaite concerns IIA as a property of a strict social welfare function and nonmanipulability of the strict social choice function attached to the former, and such nonmanipulability amounts to *strategy-proofness* (and obviously *not*

Alternatives and therewith Binary Independence are eminently reasonable assumptions to make in a realistic study of collective choice. I know of no real-world collective-choice process that violates either condition. Both formalize the idea that collective choices depend only on such preferential data as could be revealed by *voting.'* (Schwartz (1986), p.33).

[38]The examples considered by Huntington (1938) concern in fact competing teams, and the relevant numbers/scores are uniquely determined by the measurement of individual performances of each team's members.

[39]*Majority judgment* as recently introduced by Balinski and Laraki (Balinski, Laraki (2011)) denotes a family of aggregation and voting mechanism which typically satisfy the 'Postulate of Relevancy' while violating IIA (see also Vannucci (2019) for a detailed discussion of strategy-proofness properties of majority judgment).

[40]Specifically, Satterthwaite's theorem establishes a one-to-one correspondence between sovereign strategy-proof strict social choice functions and sovereign strict social welfare functions that satisfy IIA and the Weak Pareto principle, whenever the size of the set A of alternatives is not smaller than three (see Satterthwaite (1975), Theorem 2). A *strict* social choice function (social welfare function, respectively) is a social choice function (social welfare function, respectively) whose domain is restricted to arbitrary profiles of *linear orders* of A.

[41]Even a most respected, highly influential author such as Saari opts on his part for the following concise rendition of the upshot of the Satterthwaite's theorem mentioned in the text: 'A nonmanipulable system satisfies IIA' (Saari (2008), p.60).

agenda manipulation-proofness) of the latter. To be sure, further interesting elaborations on such connections between IIA and nonmanipulable aggregation rules are provided in Sato (2015). Specifically, Sato considers *four* notions of *nonmanipulability* for strict social welfare functions in order to formulate both agenda manipulation-proofness and strategy-proofness requirements, respectively. Then, relying on the Kendall metric for linear orders, he introduces a weak continuity condition for strict social welfare functions called *Bounded Response* [42]. The main result of Sato (2015) implies *the equivalence* of the following statements concerning a *strict* social welfare function f for (N, A): (1) f satisfies Bounded Response and *at least one* of the four distinct agenda manipulation-proofness or strategy-proofness conditions mentioned above; (2) f satisfies Bounded Response and *each one* of the foregoing nonmanipulability conditions; (3) f satisfies Adjacency-restricted Monotonicity (AM)[43] and the Arrowian *IIA* condition. Thus, even factoring in AM (a very mild requirement that is virtually undisputable) it turns out that IIA is in particular a necessary condition of strategy-proofness *only* for a *specific* class of '*weakly continuous*' and *strict* social welfare functions[44]. To put it simply, a closer inspection of both Satterthwaite (1975) and Sato (2015) confirms that the most interesting results they contribute are in fact *silent* on necessary and/or sufficient conditions for strategy-proofness of *general*, *unrestricted social welfare functions*.

[42]The Kendall distance d_K between rankings is given by the minimal number of transpositions of adjacents elements that is necessary to obtain one linear order starting from another one.

A strict social welfare function f satisfies Bounded Response if $d_K(f(R_N), f(R'_N)) \leq 1$ whenever two preference profiles R_N, R'_N are the same except for the preference of a single agent i, and R_i and R'_i are *adjacent* (i.e. R'_i is obtained from R_i by permuting the R_i-ranks of *a single* pair of alternatives with *consecutive* R_i-ranks).

[43]The *Adjacency-Restricted Monotonicity* condition for strict social welfare functions simply requires that for any pair of 'adjacent' profiles R_N, R'_N and any $x, y \in A$, if $[yR_ix$, xR'_iy and $xf(R_N)y]$ then $xf(R'_N)y$ as well.

[44]It should also be emphasized that, when it comes to preference aggregation problems, there is no reason to consider continuity conditions as essentially 'technical' and innocuous. It is indeed well-known that anonymity and idempotence of a social welfare function (or indeed of virtually any preference aggregation rule for arbitrary profiles of total preorders) are inconsistent with preservation of 'preference proximity' (see Baigent (1987)). See also Lauwers, Van Liedekerke (1995) and Saari (2008) for more general considerations on the difficulties raised by continuity properties for aggregation rules. Clearly enough, requiring proper 'responsiveness' of a preference aggregation rule is one thing, and insisting on its 'continuity' quite another.

All of the above implies that both agenda manipulation-proofness and strategy-proofness of a proper consensus-based social welfare function do indeed require that IIA be *either just dropped or at the very least considerably relaxed.*

The independence condition used in the present paper, namely $M_{\mathcal{X}}$-Independence, can be indeed regarded as a drastic relaxation of IIA when applied to social welfare functions. It was first introduced by Monjardet (1990) and explicitly related to IIA and Arrowian aggregation models, but not at all to strategy-proofness issues (or, for that matter, to agenda manipulation-proofness issues) [45] .

Unsurprisingly, several alternative weakenings of IIA have been proposed in the earlier literature. An entire set of substantially relaxed versions of IIA was first introduced and discussed by Hansson (1973) with no reference whatsoever to nonmanipulability issues of any sort. The strongest of them (i.e. the *least* dramatic relaxation of IIA, denoted by Hansson as Strong Positionalist Independence (SPI)[46]) requires invariance of aggregate preference between any two alternatives x, y for any pair of preference profiles such that their *restrictions to* $\{x, y\}$ *are identical, and for every agent/voter the supports of the respective closed preference intervals having* x *and* y *as their extrema are also identical.* Incidentally, SPI has been recently rediscovered, relabeled as Modified IIA, and provided with a new motivation by Maskin (2020). Indeed, Maskin points out that SPI enforces resistance of the relevant aggregation rule to certain sorts of 'vote splitting' effects, thereby connecting SPI to manipulation issues, including strategic manipulation. Notice, however, that Maskin's proposal is aimed at strategy-proofness of the '*maximizing' social choice function* induced by a certain social welfare function (as opposed to strategy-proofness of the social welfare function itself). Another weakening of IIA that is even stronger than SPI has also been proposed by Saari under the label 'Intensity form of IIA' (IIIA): it requires invariance of aggregate preference between any two alternatives x, y for any pair of preference profiles such that *for every*

[45]It should be noted, however, that conditions strictly related to $M_{\mathcal{X}}$-Independence are deployed in Dietrich, List (2007b) to study strategy-proofness properties in judgment aggregation as discussed below in the present section.

[46]The label comes from the fact that SPI is of course satisfied by 'positionalist' or score-based aggregation rules including the Borda Count rule (which assigns to every alternative x a score given by the sum of its individual ranks, defined as the sizes of the sets of alternatives which are classified as strictly worse than x itself).

agent/voter the rank (or score) difference between x and y is left unchanged from one profile to the other (see Saari (1995) and (1998))[47].

Remarkably, even at a first glance one conspicuous difference between $M_{\mathcal{X}}$-Independence and SPI (or IIIA) stands out immediately: the former relies heavily on the structure of the outcome set, while SPI and IIIA only impinge upon the relevant preference profiles, completely disregarding any specific feature of the underlying outcome set.

This crucial difference and its significant import can be further clarified and fully appreciated by reconsidering all the relaxations of IIA mentioned above from the common perspective of '*aggregation by binary issues*' that encompasses them all.

The *binary aggregation model* originates with Wilson (1975) and has been further extended by Rubinstein, Fishburn (1986)[48]: a finite number of k issues are considered for a collective yes/no judgment (the output) to be based on some profile of individual yes/no judgments on each issue (the input), under some *feasibility constraints* (usually the same, but possibly different) imposed, respectively, on inputs and outputs[49]. Thus, the basic aggregation rules for a set N of agents and a set $K = \{1, ..., k\}$ of binary issues are given by functions $f : X^N \to X$ with $X \subseteq \{0, 1\}^K$. This model has also be shown to be equivalent to the basic model of *judgment aggregation* where the judgments to be aggregated amount to acceptance/rejection of every element of an *agenda* of interconnected formulas of a suitable formal language representing propositions (see Dokow, Holzman (2009)). Indeed, Arrow's general (im)possibility theorem for social welfare functions has been explicitly shown to follow as a special interesting case under both the *feasible binary aggregation* and the *judgment aggregation* frameworks (see e.g. Dokow, Holzman (2010a, 2010b) for the former and Dietrich,

[47]Arguably, Saari's IIIA can also be regarded as a formalization of the criticism of IIA originally advanced by Dahl (1956) with his advocacy of aggregation rules based on intensity of individual preferences. Notice that IIIA is indeed satisfied by some positional aggregation rules such as the Borda Count but also by majority judgment as discussed above.

[48]To be sure, the original work by Wilson only considers the *finite* case, but Wilson's framework can also be extended to an infinite number of issues, and to non-binary issues (see Dokow, Holzman (2010c)). Indeed, one such extension is covered in Rubinstein, Fishburn (1986). Since the present paper is only concerned with finite social welfare functions, however, we shall only consider the basic binary aggregation model with a *finite* number of issues.

[49]In more recent contributions coming from the computational social choice and artificial intelligence research communities '*integrity constraints*' is the most commonly used label to denote such constraints (see e.g. Grandi, Endriss (2013)).

List (2007a), Mongin (2008), Daniëls, Pacuit (2008), Porello (2010) for the latter)[50].

An additional and most convenient perspective for the finite version of the binary aggregation model of our concern here is provided by some joint work of Nehring and Puppe (see in particular Nehring, Puppe (2007),(2010)). To be sure, Nehring, Puppe (2007) is mainly concerned with *strategy-proof social choice functions* as defined on profiles of total preorders on finite sets. Conversely, Nehring, Puppe (2010) is focussed on an 'abstract' class of Arrowian aggregation problems including preference aggregation and, more specifically, social welfare functions, but it does *not* address issues concerning their strategy-proofness properties. However, social choice functions with the top-only property[51] may be regarded as aggregation rules endowed with a specific domain of total preorders, and the class of Arrowian aggregation rules considered in Nehring, Puppe (2010) does include the case of preference aggregation rules in finite median semilattices. Specifically, Nehring and Puppe attach to any *finite* outcome space a certain finite hypergraph $\mathbb{H} = (X, \mathcal{H})$ denoted as *property space,* where the set $\mathcal{H} \subseteq \mathcal{P}(X) \setminus \{\varnothing\}$ of (nonempty) hyperedges or *properties* of outcomes/states in X is *complementation-closed* and *separating* (namely $H^c := X \setminus H \in \mathcal{H}$ whenever $H \in \mathcal{H}$, and for every two *distinct* $x, y \in X$ there exists $H_{x^+y^-} \in \mathcal{H}$ such that $x \in H_{x^+y^-}$and $y \notin H_{x^+y^-}$). Such a property space $\mathbb{H}$ models the set of all *binary* properties of outcomes that are regarded as relevant for the decision problem at hand. Thus, *binary issues are modeled here as pairs* (H, H^c) *of complementary properties* and, as it is easily checked, *both the feasible binary aggregation and the judgment aggregation models can be immediately reformulated as aggregation models in property spaces.* Then, a betweenness relation $B_{\mathbb{H}} \subseteq X^3$ is introduced by stipulating that $B_{\mathbb{H}}(x, y, z)$ holds precisely

[50]In particular, Mongin (2008) introduces a specific *weakening of IIA* for the standard judgment aggregation model, by restricting the scope of IIA to atomic propositional formulas, and still obtains a version of Arrow's (im)possibility theorem under WP. As previously mentioned, Daniëls, Pacuit (2008) offers another characterization of dictatorial rules in a quite general judgment aggregation framework using just some consequences of IIA as combined with non-constancy and neutrality conditions. Furthermore, it has been shown that Arrowian characterizations of dictatorial aggregation rules by IIA and idempotence hold for other disparate domains including arbitrary single-valued choice functions on finite sets (Shelah (2005)) and task assignments (Dokow, Holzman (2010c)).

[51]A social choice function for (N, A) is a function $f : \mathcal{D}^N \to A$ where $\mathcal{D} \subseteq \mathcal{R}_A^T$: it satisfies the *top-only property* if $f(R_N) = f(R'_N)$ whenever $t(R_i) = t(R'_i)$ for each $i \in N$, and $|t(R_i)| = |t(R'_i)| = 1$ for all $i \in N$ (with $t(R_i) := \{x \in A : xR_iy \text{ for all } y \in A\}$).

when y satisfies all the properties shared by x and z [52]. Moreover, *single-peaked* preference domains on X can be defined relying on $B_{\mathbb{H}}$. In particular, $B_{\mathbb{H}}$ is said to be *median* if for every $x, y, z \in X$ there exists a unique $m_{xyz} \in X$ such that $B_{\mathbb{H}}(x, m_{xyz}, y)$, $B_{\mathbb{H}}(x, m_{xyz}, z)$, and $B_{\mathbb{H}}(y, m_{xyz}, z)$ hold[53]. Of course, the main advantage of that second-order representation of binary issues is the possibility to *focus on several different property spaces which are defined on the very same ground set of alternative states.*

The following key results are obtained by Nehring and Puppe: (i) the class of all idempotent social choice functions which are strategy-proof on the domain of single-peaked preferences thus defined are characterized in terms of voting by binary issues through a certain combinatorial property[54] of the families of winning coalitions for the relevant issues and (ii) *if the property space is median* then such combinatorial property is definitely met, and consequently non-dictatorial neutral and/or anonymous strategy-proofs aggregation rules including the simple majority rule are available (Nehring, Puppe (2007), Theorems 3 and 4). Furthermore, in Nehring, Puppe (2010) the very same theoretical framework is deployed to analyze preference aggregation and social welfare functions. In particular, several 'classical' properties for social welfare conditions including the Arrowian Independence of Irrelevant Alternatives (IIA) property can be reformulated in more general terms which depend on the specification of the relevant property space[55]: it follows that several versions of IIA can be considered. But then, as it turns out, (iii) the *versions of IIA attached to median property*

[52]In particular, a nonempty subset $Y \subseteq X$ is said to be *convex* for $\mathbb{H} = (X, \mathcal{H})$ if for every $x, y \in Y$ and $z \in X$, if $B_{\mathbb{H}}(x, z, y)$ then $z \in Y$, and *prime* (or a *halfspace*) for $\mathbb{H}$ if both Y and $X \setminus Y$ are convex for $\mathbb{H}$ and $\{Y, X \setminus Y\} \subseteq \mathcal{H}$.

[53]In that case, $\mathbb{H}$ is said to be a *median property space,* $(X, m^{\mathbb{H}})$ (where $m^{\mathbb{H}} : X^3 \to X$ is defined by the rule $m^{\mathbb{H}}(x, y, z) = m_{xyz}$ for every $x, y, z \in X$) is a *median algebra*, and for each $u \in X$ the pair $(X, \vee_u)$ (where $x \vee_u y = y$ iff $m^{\mathbb{H}}(x, y, u) = y$ for some $u \in X$) is a *median join-semilattice* having u as its maximum.

[54]The combinatorial property mentioned in the text is the so-called 'Intersection Property' which requires that for every minimal inconsistent set of properties, it must be the case that any selection of winning coalitions for the corresponding binary issues has a non-empty intersection.

[55]Specifically, given a property space $\mathbb{H} = (\mathcal{R}_A^T, \mathcal{H})$, such a generalized IIA for a social welfare function $f : (\mathcal{R}_A^T)^N \longrightarrow \mathcal{R}_A^T$ can be defined as follows: for every $H \in \mathcal{H}$ and $R_N, R'_N \in (\mathcal{R}_A^T)^N$ such that $\{i \in N : R_i \in H\} = \{i \in N : R'_i \in H\}$,

if $f(R_N) \in H$ then $f(R'_N) \in H$ as well. Of course the original Arrowian version of such a generalized IIA is obtained by taking $\mathcal{H} := \{H_{(x,y)} : x, y \in A\}$ with $H_{(x,y)} := \{R \in \mathcal{R}_A^T : xRy\}$.

spaces are consistent with anonymous and (weakly) neutral social welfare functions including those induced by majority-based aggregation rules (Nehring, Puppe (2010), Theorem 4). Interestingly, a specific example of a median property space for the set of all total preorders is also provided by Nehring and Puppe, namely the one whose issues consist in asking for each non-empty $Y \subseteq X$ and any total preorder R whether or not Y is a *lower contour* of R with respect some outcome $x \in X$.[56] By contrast, it can be easily checked that when translated into the property-space framework SPI and IIIA correspond to *non-median property spaces.*[57]

The overlappings between such results and those presented here are remarkable, along with some sharp differences which make them mutually independent. Since any finite median semilattice is indeed an example of a finite median algebra[58], and is consequently representable as a median property space[59], all of the Nehring and Puppe's results

---

[56]Thus, the property space suggested here is $\mathbb{H}^\circ := \left\{\mathcal{R}_A^T, \mathcal{H}^\circ\right\}$, where $\mathcal{H}^\circ := \{H_L : \varnothing \neq L \subseteq A\}$ and $H_L := \left\{ \begin{array}{c} R \in \mathcal{R}_A^T : \text{for some } x \in A \\ L = \{y \in A : xRy\} \end{array} \right\}$.

[57]Indeed, the most natural property-space attached to SPI is $\mathcal{H}_{SPI} := \left\{ \begin{array}{c} H_{(x,y,B)} : x, y \in A \\ B \subseteq A \setminus \{x, y\} \end{array} \right\}$ with $H_{(x,y,B)} := \left\{ \begin{array}{c} R \in \mathcal{R}_A^T : \ \{a\} \times \ B \subseteq R \text{ and} \\ B \times \{b\} \subseteq R \\ \text{if } \{a, b\} = \{x, y\} \end{array} \right\}$.

Similarly, the most natural property-space attached to IIIA is $\mathcal{H}_{IIIA} := \left\{ \begin{array}{c} H_{(x,y,k)} : x, y \in A, \\ k \leq |A| - 2 \end{array} \right\}$ with $H_{(x,y,k)} := \left\{ \begin{array}{c} R \in \mathcal{R}_A^T : \text{ either } I_{x,y} = \{z \in A : xP(R)zP(R)y\} \text{ or } k = |I_{x,y}| \text{ where either} \\ \text{or } I_{x,y} = \{z \in A : yP(R)zP(R)x\} \text{ or } k = |I_{x,y}| \text{ w} \\ \text{and } |I_{x,y}| = k \end{array} \right\}$.

It can be shown that $\mathcal{H}_{SPI}$ and $\mathcal{H}_{IIIA}$ are *not* median property spaces since both of them contain minimal inconsistent subsets of properties of size three. To check validity of that statement, just consider any triplet $\left\{H_{x,y,\emptyset}, H_{y,z,\emptyset}, H_{z,x,\emptyset}\right\} \subseteq \mathcal{H}_{SPI}$ and $\{H_{x,y,0}, H_{y,z,0}, H_{z,x,0}\} \subseteq \mathcal{H}_{IIIA}$ with $x \neq y \neq z \neq x$.

[58]Specifically, a finite median join-semilattice can be regarded as a generic instance of a finite median algebra with one of its elements singled out (that point corresponds to the top element of the semilattice).

[59]For instance, it is *always* possible to represent a (finite) median algebra as a (finite) property space by taking as properties its *prime* sets as defined through its median betweenness (see e.g. Bandelt, Hedlíková (1983), Theorem 1.5, and note 52 above for a definition of prime sets). It is important to observe that in general a finite median algebra or ternary space admits of several representations by distinct median property spaces. By contrast, a ternary (finite) algebra or space which

mentioned above *do apply* to finite median semilattices as a special case. Notice however that our results provide a characterization of (finite) social welfare functions which is both *more explicit* (it includes a polynomial description of some such rules) and *more comprehensive* (it is a complete characterization in that it is not limited to sovereign and idempotent ones). Moreover, our treatment of social welfare functions can also be translated in terms of a *median* property space, though a *different one* from that considered by Nehring and Puppe. In fact, in our case the set of relevant properties corresponds to the meet-irreducibles of the of total preorders, namely the total preorders having just *two* indifference classes, or equivalently the binary ordered classifications of outcomes as *good* or *bad*, respectively. Accordingly, the collection of relevant issues consist in asking, for each binary good/bad classification of outcomes and any total preorder R, whether the latter is consistent with the given binary classification[60]. Thus, proper consensus-based social welfare functions that are agenda manipulation-proof and even strategy-proof can be defined by *binary aggregation*, provided that the set of relevant binary issues is carefully selected, and in fact *expanded* if the basic alternatives are *more than three*: notice that, when expressed in terms of properties of $\mathcal{R}_A$ with $|A| = m$, the size of the set of actually relevant binary issues (for each individual preference relation of any profile) is $m(m-1)$ for IIA, and $2(2^{m-1}-1)$ for $M_{\mathcal{X}}$-Independence. This point is strongly consonant with one of the main arguments in Saari (2008), lamenting the enormous loss of information enforced by the Arrowian IIA. It also amounts to a special instance of a recurrent theme in the social choice-theoretic literature, namely emphasizing the link between Arrow's theorem and the strictures of the preference-information base enforced by IIA and the other Arrowian axioms (see e.g. the classic Sen (2017) for extensive elaborations on that topic). Notice, however, that while changes and/or enrichments of the Arrowian input-format figure prominently

---

is not median can only be represented by (finite) property spaces which are *not* median.

[60]Thus, the appropriate version of generalized IIA in our own model is $\mathbb{H}^* := (\mathcal{R}_A^T, \mathcal{H}^*)$ with $\mathcal{H}^* := \left\{ \begin{array}{c} H_{A_1A_2} : A_1 \neq \varnothing \neq A_2 \\ A_1 \cap A_2 = \varnothing,\ A_1 \cup A_2 = A \end{array} \right\}$

$H_{A_1,A_2} := \left\{ R \in \mathcal{R}_A^T : R \subseteq R_{A_1A_2} \right\}$

and $R_{A_1A_2}$ is of course the two-indifference-class total preorders having A_1 and A_2 as top and bottom indifference classes, respectively. Notice that both $\mathbb{H}^*$ and Nehring-Puppe's $\mathbb{H}^\circ$as previously defined (see note 56 above) are *median* property spaces, while the original Arrowian $\mathbb{H}$ is not.

among the invoked remedies for the aforementioned strictures, the relaxations of IIA we have been considering only apply to a *fixed, standard input-format consisting of profiles of total preorders.*

4. Concluding remarks

The IIA condition for preference aggregation rules was introduced by Arrow in order to ensure their agenda manipulation-proofness, but when combined with a few minimal reasonable conditions it results in a characterization of dictatorial social welfare functions. That is the content of Arrow's general possibility theorem. Under the assumption that IIA is indeed the only way to block agenda manipulation, that theorem does also imply that reliable and proper consensus-based social welfare functions do not exist. Now, that interpretation projects a negative, disagreeable shadow on the perceived consequences of Arrow's theorem since it suggests that *no meaningful consensus-based formulation of 'general interest' is available* as a guide and benchmark to promote and assess public decisions and policies, and improve the design and/or implementation of democratic protocols. Arguably, it is not unfair to describe all that as the '*dark*' side of Arrow's theorem. The present work shows however that, as a matter of fact, agenda manipulation-proofness of a social welfare function is indeed available *without any appeal to IIA* provided that agenda formation and preference elicitation are *simultaneous*. In the latter case some *anonymous, idempotent, agenda manipulation-proof, minimally efficient and weakly-neutral social welfare functions do exist*. Moreover, a *much relaxed independence condition* that they do satisfy ensures their *strategy-proofness* as well. But then, from the perspective provided by such *positive* results, Arrow's theorem may also be regarded as a most *constructive* contribution to the design of preference aggregation rules, in that it warns us that agenda formation and preference elicitation are *not* to be coupled sequentially. That is precisely the '*bright*' side of Arrow's theorem that the present work is meant to highlight and emphasize.

To be sure, the consensus-based, agenda manipulation-proof, and strategy-proof preference aggregation rules we have shown to be available require a significant increase of the amount of information to be extracted from preference profiles and processed. Thus, reliance on such aggregation rules also involves a careful consideration of *computational complexity* issues. Moreover, while *individual* strategy-proofness issues concerning social welfare functions have been also considered in the present work, the further problems arising from *coalitional strategy-proofness* requirements for preference aggregation rules have been left

untouched.[61] Such most significant issues are however best left as two challenging topics for future research.

References

[1] Aleskerov F. (1999): *Arrovian Aggregation Models.* Kluwer, Dordrecht.

[2] Arrow K.J. (1963): *Social Choice and Individual Values*, 2nd edn. Yale University Press, New Haven.

[3] Austen-Smith D., J. Banks (2005): *Positive Political Theory II: Strategy and Structure.* University of Michigan Press, Ann Arbour.

[4] Avann S.P. (1961): Metric ternary distributive semilattices, *Proceedings of the American Mathematical Society* 12, 407-414.

[5] Baigent N. (1987): Preference proximity and anonymous social choice, *Quarterly Journal of Economics* 102, 161-170.

[6] Balinski M, R. Laraki, (2011): *Majority Judgment. Measuring, Ranking, and Electing.* MIT Press, Cambridge MASS.

[7] Bandelt H.J., J. Hedlíková (1983): Median algebras, *Discrete Mathematics* 45, 1-30.

[8] Binmore K.G. (1976): Social choice and parties, *Review of Economic Studies* 43, 459-464.

[9] Buchanan J.M. (1954): Social choice, democracy, and free markets, *Journal of Political Economy* 62, 114-123.

[10] Cato S. (2012): Social choice without the Pareto principle: a comprehensive analysis, *Social Choice and Welfare* 39, 869-889.

[11] Dahl R.A. (1956): *A Preface to Democratic Theory.* University of Chicago Press, Chicago.

[12] Dahl R.A. (1979): Procedural democracy, in P. Laslett, J. Fishkin (eds.): *Philosophy, Politics & Society (Fifth Series).* Basil Blackwell, Oxford.

[13] Daniëls T.R., E. Pacuit (2008): A general approach to aggregation problems, *Journal of Logic and Computation* 19, 517-536.

[14] Danilov V.I. (1994): The structure of non-manipulable social choice rules on a tree, *Mathematical Social Sciences* 27, 123-131.

[15] Denicolò V. (1998): Independent decisiveness and the Arrow theorem, *Social Choice and Welfare* 15, 563-566.

[16] Denicolò V. (2000): Independence of irrelevant alternatives and consistency of choice, *Economic Theory* 15, 221-226.

[17] Dietrich F. (2016): Judgment aggregation and agenda manipulation, *Games and Economic Behavior* 95, 113-136.

[18] Dietrich F., C. List (2007a): Arrow's theorem in judgment aggregation, *Social Choice and Welfare* 29, 19-33.

[19] Dietrich F., C. List (2007b): Strategy-proof judgment aggregation, *Economics and Philosophy* 23, 269-300.

[61] Concerning the relationships between individual and coalitional strategy-proofness for general aggregation rules see e.g. Vannucci (2016).

[20] Dokow E., R. Holzman (2009): Aggregation of binary evaluations for truth-functional agendas, *Social Choice and Welfare* 32, 221-241.
[21] Dokow E., R. Holzman (2010a): Aggregation of binary evaluations, *Journal of Economic Theory* 145, 495-511.
[22] Dokow E., R. Holzman (2010b): Aggregation of binary evaluations with abstentions, *Journal of Economic Theory* 145, 544-561.
[23] Dokow E., R. Holzman (2010c): Aggregation of non-binary evaluations, *Advances in Applied Mathematics* 45, 487-504.
[24] Fishburn P.C. (1973): *The Theory of Social Choice.* Princeton University Press, Princeton NJ.
[25] Grandi U., U. Endriss (2013): Lifting integrity constraints in binary aggregation, *Artificial Intelligence* 199-200, 45-66.
[26] Hansson B. (1969a): Group preferences, *Econometrica* 37, 50-54.
[27] Hansson B. (1969b): Voting and group decision functions, *Synthese* 20, 526-537.
[28] Hansson B. (1973): The independence condition in the theory of social choice, *Theory and Decision* 4, 25-49.
[29] Huntington E.V. (1938): A paradox in the scoring of competing teams, *Science* 88, 287-288.
[30] Janowitz M.J. (1984): On the semilattice of weak orders of a set, *Mathematical Social Sciences* 8, 229-239.
[31] Lauwers L., L. Van Liedekerke (1995): Ultraproducts and aggregation, *Journal of Mathematical Economics* 24, 217-237.
[32] Malawski M., L. Zhou (1994): A note on social choice theory without the Pareto principle, *Social Choice and Welfare* 11, 103-107.
[33] Maskin E. (2020): A modified version of Arrow's IIA condition, *Social Choice and Welfare* 54, 203-209.
[34] Mongin P. (2008): Factoring out the impossibility of logical aggregation, *Journal of Economic Theory* 141, 100-113.
[35] Monjardet B.(1990): Arrowian characterizations of latticial federation consensus functions, *Mathematical Social Sciences* 20, 51-71.
[36] Nehring K., C. Puppe (2007): The structure of strategy-proof social choice-part I: general characterization and possibility results on median spaces, *Journal of Economic Theory* 135, 269-305.
[37] Nehring K., C. Puppe (2010): Abstract Arrowian aggregation, *Journal of Economic Theory* 145, 467-494.
[38] Porello D. (2010): Ranking judgments in Arrow's setting, *Synthese* 173, 199-210.
[39] Quesada A. (2002): Power of enforcement and dictatorship, *Theory and Decision* 52, 381-387.
[40] Ray P. (1973): Independence of irrelevant alternatives, *Econometrica* 41, 987-991.
[41] Riker W.H. (1982): *Liberalism against Populism. A Confrontation between the Theory of Democracy and the Theory of Social Choice.* W.H. Freeman, San Francisco.
[42] Rubinstein A., P.C. Fishburn (1986): Algebraic aggregation theory, *Journal of Economic Theory* 38, 63-77.
[43] Saari D.G. (1995): *Basic Geometry of Voting.* Springer, New York.

[44] Saari D.G. (1998): Connecting and resolving Sen's and Arrow's theorems, *Social Choice and Welfare* 15, 239-261.

[45] Saari D.G. (2008): *Disposing Dictators, Demystifying Voting Paradoxes. Social Choice Analysis.* Cambridge University Press, Cambridge UK.

[46] Sato S. (2015): Bounded response and the equivalence of nonmanipulability and independence of irrelevant alternatives, *Social Choice and Welfare* 44, 133-149.

[47] Satterthwaite M.A. (1975): Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions, *Journal of Economic Theory* 10, 187-217.

[48] Savaglio E., S. Vannucci (2019): Strategy-proof aggregation rules and single peakedness in bounded distributive lattices, *Social Choice and Welfare* 52: 295-327.

[49] Savaglio E., S. Vannucci (2021): Strategy-proof aggregation rules in median semilattices with applications to preference aggregation, WP DEPS 867, University of Siena.

[50] Schofield N.J. (1985): *Social Choice and Democracy.* Springer, Berlin.

[51] Schwartz T. (1986): *The Logic of Collective Choice.* Columbia University Press, New York.

[52] Sen A.K. (1977): Social choice theory: a re-examination, *Econometrica* 45, 53-88.

[53] Sen A.K. (1993): Internal consistency of choice, *Econometrica* 61, 495-521.

[54] Sen A.K. (2017): *Collective Choice and Social Welfare,* expanded edition. Penguin Books.

[55] Sethuraman J., C.-P. Teo, R.V. Vohra (2003): Integer programming and Arrovian social welfare functions, *Mathematics of Operations Research* 28, 309-326.

[56] Shelah S. (2005): On the Arrow property, *Advances in Applied Mathematics* 34, 217-251.

[57] Sholander M. (1952): Trees, lattices, order, and betweenness, *Proceedings of the American Mathematical Society* 3, 369-381.

[58] Sholander M. (1954): Medians and betweenness, *Proceedings of the American Mathematical Society* 5, 801-807.

[59] Van de Vel M.L.J. (1993): *Theory of Convex Structures.* North Holland, Amsterdam.

[60] Vannucci S. (2016): Weakly unimodal domains, anti-exchange properties, and coalitional strategy-proofness of aggregation rules, *Mathematical Social Sciences* 84, 56-67.

[61] Vannucci S. (2019): Majority judgment and strategy-proofness: a characterization, *International Journal of Game Theory* 48, 863-886.

[62] Wilson R. (1972): Social choice theory without the Pareto principle, *Journal of Economic Theory* 5, 478-486.

[63] Wilson R. (1975): On the theory of aggregation, *Journal of Economic Theory* 10, 89-99.

[64] Young, H.P. (1995): Optimal voting rules, *Journal of Economic Perspectives* 9, 51-64.

University of Siena, Department of Economics and Statistics