Parallel proofs of Arrow's and the Gibbard-Satterthwaite theorem

Antonio Quesada[†]

Departament d'Economia, Universitat Rovira i Virgili, Avinguda de la Universitat 1, 43204 Reus, Spain

2nd April 2005

80.19

Abstract
Arrow's and the Gibbard-Satterthwaite theorems are proved using a common proof strategy based on a dictatorship result for choice functions. One of the instrumental results obtained shows the inconsistency between the basic assumption in each of these theorems and a mild majority principle.

JEL Classification: D71

Keywords: Arrow's theorem; Gibbard-Satterthwaite theorem; Choice function; Majority.

[†] E-mail address: aqa@urv.net.

1. Introduction

It is known that Arrow's (1963, p. 97) theorem and the Gibbard (1973) - Satterthwaite (1975) theorem, two of the fundamental results in economic theory, are closely related results. Reny (2001), for instance, provides a single proof of both theorems, whereas Barberà (2001) presents a result of which the two theorems are special cases.

This note provides additional connections between the two theorems. A first connection concerns the theorems themselves, since both are proved following a common proof that relies on a dictatorial result for choice functions. A second connection refers to the conditions peculiar to each theorem, independence of irrelevant alternatives (IIA) in Arrow's theorem and stratregy-proofness in the Gibbard-Satterthwaite theorem. On the one hand, it is shown that both IIA and SP are inconsistent with a mild majority principle; and, on the other, that the same assumption can lead from IIA to the dictatorship result in Arrow's theorem and also from stratregy-proofness to the dictatorship result in the Gibbard-Satterthwaite theorem. These connections are briefly discussed in Section 5.

2. Notation and definitions

Let $N = \{1, 2, \dots , n\}$ be a non-empty finite set whose members designate individuals, A a finite set with $m \geq 1$ elements representing alternatives, L the set of rankings (complete, transitive and antisymmetric binary relations) that can be defined on A and L^n the set of profiles $(P_1, \dots , P_n)$, where, for all $i \in N$, $P_i \in L$. Rankings on A are identified with sequences $(x_1, \dots , x_m)$ such that $A = \{x_1, \dots , x_m\}$, the interpretation being that x_r is strictly preferred to x_s if, and only if, $r < s$. For $P \in L^n$, $\xi \in A^n$ and $J \subseteq N$, ξ_i and P_i denote the ith component of ξ and P, respectively, whereas P_J abbreviates $(P_j)_{j\in J}$. For $J \subseteq N$ and $i \in N$, $-J$ abbreviates $N\backslash J$ and $-i$ abbreviates $N\backslash\{i\}$.

Expression	*means*
${}^k r$	for $1 \leq k \leq m$, the kth alternative x_k in the ranking $r = (x_1, \dots , x_m) \in L$
$p(x, r)$	the position of $x \in A$ in ranking $r \in L$ (the number k such that ${}^k r = x$)
$r\vert_B$	the restriction of $r = (x_1, \dots , x_n) \in L$ to $B \subseteq A$, that is, the ranking $s = (y_1, \dots , y_b)$ such that $b = \vert B\vert$ and, for all $y \in B$, $p(y, s) \leq p(y, r)$
xry	$p(x, r) < p(y, r)$: $x \in A$ is above (or comes before) $y \in A\backslash\{x\}$ in $r \in L$
xP_Jy	for $J \subseteq N$ and $P \in L^n$, abbreviation of "xP_jy, for all $j \in J$"

r^{xy}	the ranking $s \in L$ obtained from $r \in L$ by exchanging $x \in A$ and $y \in A\backslash\{x\}$: $p(x, s) = p(y, r)$ and, for all $z \in A\backslash\{x, y\}$, $p(z, s) = p(z, r)$
$r^{\uparrow x}$	the ranking $s \in L$ obtained from $r \in L$ by ranking x the first: ${}^1s = x$ and, for all $y \in A\backslash\{x\}$ and $z \in A\backslash\{x, y\}$, $p(y, s) \le p(z, s) \Leftrightarrow p(y, r) \le p(z, r)$
(x^J, y^{-J})	the member ξ of A^n with, for all $j \in J$, $\xi_j = x$ and, for all $j \in N\backslash J$, $\xi_j = y$
(x^J, ζ_{-J})	the member ξ of A^n with, for all $j \in J$, $\xi_j = x$ and, for all $j \in N\backslash J$, $\xi_j = \zeta_j$
$P_J{}^{\uparrow x}$	for non-empty $J \subseteq N$ and $P \in L^n$, abbreviation of $(P_j{}^{\uparrow x})_{j\in J}$
x covers y in r	$p(x, r) + 1 = p(y, r)$, where $x \in A$, $y \in A\backslash\{x\}$ and $r \in L$

Definition 2.1. A social welfare function (SWF) is a mapping $f : L^n \to L$. A social choice function (SCF) is a mapping $g : L^n \to A$. A choice function (CF) is a mapping $h : A^n \to A$.

Definition 2.2. A set $J \subseteq N$ with $r \geq 1$ members is decisive if it is decisive for all $x \in A$ and it is decisive for x: (i) in a SWF f if, for all $P \in L^n$ and $y \in A\backslash\{x\}$, $xf(P)y$ when xP_Jy; (ii) in a SCF g if, for all $P \in L^n$, $g(P) = x$ when, for all $i \in J$, ${}^1P_i = x$; and (iii) in a choice function h if, for all $\xi_{-J} \in A^{n-r}$, $h(x^J, \xi_{-J}) = x$.

Definition 2.3. Social welfare function f [social choice function g, choice function h] is dictatorial if, for some $i \in N$, $\{i\}$ is decisive in f [g, h].

Definition 2.4. A social welfare function f satisfies: (i) PAR (the Pareto principle) if, for all $x \in A$, N is decisive for x; and (ii) IIA (independence of irrelevant alternatives) if, for all $P \in L^n$, $Q \in L^n$, $x \in A$ and $y \in A\backslash\{x\}$, $(P_1|_{\{x,y\}}, \ldots, P_n|_{\{x,y\}}) = (Q_1|_{\{x,y\}}, \ldots, Q_n|_{\{x,y\}})$ implies $f(P)|_{\{x,y\}} = f(Q)|_{\{x,y\}}$.

Definition 2.5. A social choice function g is: (i) onto if, for every $x \in A$, there is $P \in L^n$ such that $g(P) = x$; and (ii) strategy-proof (SP) if, for all $i \in N$, $(P_i, P_{-i}) \in L^n$ and $Q_i \in L$, it is not the case that $g(Q_i, P_{-i})P_ig(P_i, P_{-i})$.

3. A dictatorial result for choice functions

A CF $h : A^n \to A$ is interpreted as a voting rule: for $\xi \in A^n$ and $i \in N$, ξ_i is the alternative for which individual i votes and $h(\xi)$ is the alternative the choice function chooses. This section presents a result for CFs on which the common proof of Arrow's theorem and the Gibbard-Satterthwaite theorem in Section 4 hinges.

A1. For all $\xi \in A^n$, $h(\xi) \in \{\xi_1, \dots, \xi_n\}$.
A2. For all $\xi, \zeta \in A^n$, if $h(\xi) = x$ and $\{i \in N: \xi_i = x\} \subseteq \{i \in N: \zeta_i = x\}$ then $h(\zeta) = x$.
A2*. For all $\xi \in A^n$, if $h(\xi) = x$ then $\{i \in N: \xi_i = x\}$ is decisive for x.

By A1, if $h(\xi) = x$ then there is $i \in N$ such that $\xi_i = x$, so $h(\xi)$ is chosen from the set of alternatives some voter votes for. A2 can be viewed as an independence condition: if h chooses x then, as long as the set of individuals voting for x is not reduced, x is still the alternative h selects. A2 also admits a non-manipulability interpretation: when h chooses x, the members not voting for x cannot force h to choose another candidate by voting otherwise. Observe that A2 is equivalent to A2* above, which asserts that if $J \subseteq N$ can enforce alternative x once then J is decisive for x (this property is related to Denicolò's (1998) relational independent decisiveness condition).

Lemma 3.1. Let $n \geq 2 < m$, $\emptyset \neq J \subset N$, $x \in A$, $y \in A \backslash \{x\}$ and $h : A^n \to A$ satisfy A1 and A2. If $h(x^J, y^{-J}) = x$ then J is decisive.

Proof. Assume $h(x^J, y^{-J}) = x$. By A2, J is decisive for x. Consider next $z \in A \backslash \{x, y\}$. By A1, $h(z^J, x^{-J}) \in \{x, z\}$. If $h(z^J, y^{-J}) = y$ then, by A2, $h(x^J, y^{-J}) = y$: contradiction. Thus, $h(z^J, y^{-J}) = z$ and, by A2, J is decisive for z. Consider finally y. By A1, $h(y^J, x^{-J}) \in \{x, y\}$. Choose $z \in A \backslash \{x, y\}$. If $h(y^J, x^{-J}) = x$ then, by A2, $h(z^J, x^{-J}) = x$, contradicting the fact that J is decisive for z. Hence, by A1, $h(x^J, y^{-J}) = x$ and, by A2, J is decisive for y.■

Proposition 3.2. If $n \geq 2 < m$ then every $h : A^n \to A$ satisfying A1 and A2 is dictatorial.

Proof. Being N finite and, by A1, decisive, it suffices to show that $J \subseteq N$ decisive implies $\{i\} \subseteq J$ decisive or $J \backslash \{i\}$ decisive. Let $J \subseteq N$ be decisive, $i \in J$, $K = J \backslash \{i\}$, $x \in A$, $y \in A \backslash \{x\}$ and $z \in A \backslash \{x, y\}$. By A1, $h(x^i, y^K, z^{-J}) \in \{x, y, z\}$. If it is z then, by A2, $h(y^J, z^{-J}) = z$, contradicting decisiveness of J. If it is x, by A2, $h(x^i, y^{-i}) = x$ so $\{i\}$ is decisive by Lemma 3.1. And if it is y then, by A2 and Lemma 3.1, K is decisive.■

4. Parallel proofs of Arrow's theorem and the Gibbard-Satterthwaite theorem

For SWF $f : L^n \to L$, define $f^* : A^n \to\to A$ to be the mapping such that $f^*(x_1, \dots, x_n) := \{x \in A: x = {}^1f(P)$, for some $P \in L^n$ with $({}^1P_1, \dots, {}^1P_n) = (x_1, \dots, x_n)\}$. For SCF $g : L^n \to A$, let $g^* : A^n \to\to A$ be the mapping such that $g^*(x_1, \dots, x_n) := \{x \in A: x = g(P)$, for some $P \in L^n$ with $({}^1P_1, \dots, {}^1P_n) = (x_1, \dots, x_n)\}$. If $F : A^n \to\to A$ is such that, for all

$\xi \in A^n$, $F(\xi)$ has one element then F is identified with the function $G : A^n \to A$ such that, for all $\xi \in A^n$, $G(\xi) = x$, where $\{x\} = \{F(\xi)\}$.

The strategy to prove both Arrow's theorem and the Gibbard-Satterthwaite theorem is summarized below (the "$\Uparrow$" implications are not proved as they are easy to prove). On the one hand, it is shown that, when $n \geq 2 < m$, if a SWF f satisfies PAR and IIA then f^* is a choice function that satisfies A1 and A2. By Proposition 3.2, f^* is dictatorial. Since f^* dictatorial makes f dictatorial, it then follows Arrow's theorem. On the other hand, it is shown that, when $n \geq 2 < m$, if a SCF g is onto and strategy-proof then g^* is a choice function that satisfies A1 and A2. By Proposition 3.2, g^* is dictatorial. Since g^* dictatorial makes g dictatorial, it then follows the Gibbard-Satterthwaite theorem.

Arrow's theorem

With $n \geq 2 < m$ and $f : L^n \to L$

f satisfies PAR and IIA

$\Downarrow \quad \Uparrow$

${}^1f(P_1, \dots, P_n) \in \{{}^1P_1, \dots, {}^1P_n\}$

${}^1f(P_1, \dots, P_n) = x$ implies that
$\{i \in N\colon {}^1P_i = x\}$ is decisive for x in f

$\Downarrow \quad \Uparrow$

f^* such that $f^*(x_1, \dots, x_n) := \{x \in A\colon$
$x = {}^1f(P)$, for some $P \in L^n$ such that
$({}^1P_1, \dots, {}^1P_n) = (x_1, \dots, x_n)\}$
is single-valued and dictatorial

$\Downarrow \quad \Uparrow$

f dictatorial

Gibbard-Satterthwaite theorem

With $n \geq 2 < m$ and $g : L^n \to A$

g is onto and strategy-proof

$\Downarrow \quad \Uparrow$

$g(P_1, \dots, P_n) \in \{{}^1P_1, \dots, {}^1P_n\}$

$g(P_1, \dots, P_n) = x$ implies that
$\{i \in N\colon {}^1P_i = x\}$ is decisive for x in g

$\Downarrow \quad \Uparrow$

g^* such that $g^*(x_1, \dots, x_n) := \{x \in A\colon$
$x = g(P)$, for some $P \in L^n$ such that
$({}^1P_1, \dots, {}^1P_n) = (x_1, \dots, x_n)\}$
is single-valued and dictatorial

$\Downarrow \quad \Uparrow$

g dictatorial

In this respect, the Gibbard-Satterthwaite theorem could be viewed as an unsuccessful attempt to escape from Proposition 3.2 by transforming the domain of the CF from A^n

to L^n, whereas Arrow's theorem could be viewed as an unsuccessful attempt to escape from Proposition 3.2 by simultaneously transforming the domain of the CF from A^n to L^n and its codomain from A to L^n.

Lemma AR1 states that if x is at the top of the social ranking $f(P)$ and IIA holds then: (i) x remains at the top when an individual changes his preference concerning two alternatives that he ranks both above x or both below x; and (ii) if some individual exchanges in his ranking the position of x with a contiguous alternative y then the social top is x or y. Lemma GS1 states the same for a strategy-proof SCF g.

Lemma AR1. Let $n \geq 2 \leq m$, $f : L^n \rightarrow L$ satisfy IIA and, for a given $P \in L^n$, $^1f(P) = x$.

$$\text{If } Q_i \in L \text{ is such that } \{y \in A\text{: } xQ_iy\} = \{y \in A\text{: } xP_iy\} \text{ then } {}^1f(Q_i, P_{-i}) = x. \tag{1}$$
$$\text{If } x \text{ covers } y \text{ in } P_i \text{ and } Q_i := P_i^{xy} \text{ then } {}^1f(Q_i, P_{-i}) \in \{x, y\}. \tag{2}$$

Proof. (1) Since $^1f(P) = x$ and, for all $y \in A\backslash\{x\}$, $P_i|_{\{x,y\}} = Q_i|_{\{x,y\}}$, by IIA, x is above every $y \in A\backslash\{x\}$ in $f(Q_i, P_{-i})$, so $^1f(Q_i, P_{-i}) = x$. (2) As $^1f(P) = x$ and, for all $v \in A\backslash\{x, y\}$, $P_i|_{\{x,v\}} = Q_i|_{\{x,v\}}$, by IIA, v cannot be above x in $f(Q_i, P_{-i})$. Thus, $^1f(Q_i, P_{-i}) \in \{x, y\}$.■

Lemma GS1. Let $n \geq 2 \leq m$, $g : L^n \rightarrow A$ be SP and, for a given $P \in L^n$, $g(P) = x$.

$$\text{If } Q_i \in L \text{ is such that } \{y \in A\text{: } xQ_iy\} = \{y \in A\text{: } xP_iy\} \text{ then } g(Q_i, P_{-i}) = x. \tag{3}$$
$$\text{If } x \text{ covers } y \text{ in } P_i \text{ and } Q_i := P_i^{xy} \text{ then } g(Q_i, P_{-i}) \in \{x, y\}. \tag{4}$$

Proof. (3) Suppose $g(Q_i, P_{-i}) = y \neq x$. If xQ_iy then, as $g(P_i, P_{-i}) = x$, g is not SP. If yQ_ix then, as yP_ix, g is not SP. (4) Suppose $g(Q_i, P_{-i}) = z \in A\backslash\{x, y\}$. If zQ_iy then g is not SP because $g(Q_i, P_{-i})P_ig(P_i, P_{-i})$. If xQ_iz then g is not SP because $g(P_i, P_{-i})Q_ig(Q_i, P_{-i})$.■

Whereas Lemmas AR1 and GS1 are purely instrumental results, Lemmas AR2 and GS2 are arguably the most significant results in the paper; see Section 4. By Lemma AR2, if x is at the social top in $f(P)$, IIA holds and the set of individuals ranking some $y \neq x$ at the top is non-empty then, for at least one of those y, x remains at the social top when y is raised to the top in all the rankings in which x was not at the top. Lemma GS2 (whose proof mimics exactly the proof of Lemma AR2) expresses the same result for a strategy-proof SCF g.

Lemma AR2. If $n \geq 2 \leq m$ and $f : L^n \rightarrow L$ satisfies IIA then, for all $P \in L^n$, if $^1f(P) = x$ and $K := \{i \in N\text{: } {}^1P_i \neq x\} \neq \varnothing$ then, for some $v \in \{^1P_1, \ldots , {}^1P_n\}\backslash\{x\}$, $^1f(P_K^{\uparrow v}, P_{-K}) = x$.

Proof. There is nothing to prove if, for all $i \in K$ and $j \in K\backslash\{i\}$, ${}^1P_i = {}^1P_j$ so let y and z be two members of $\{{}^1P_i\}_{i\in K}\backslash\{x\}$. With $I := \{i \in N: {}^1P_1 = z\}$, choose $j \in K\backslash I$ with ${}^1P_1 = y$. By a simple induction argument it is enough to show that ${}^1f(P_I, P_j^{\uparrow z}, P_{-(I\cup\{j\})}) = x$ or ${}^1f(P_I^{\uparrow y}, P_j, P_{-(I\cup\{j\})}) = x$.

- Case 1: zP_jx. By (1), ${}^1f(P_I, P_j^{\uparrow z}, P_{-(I\cup\{j\})}) = x$.
- Case 2: xP_jz. Let $J := \{i \in I: xP_iy\}$.
  - Case 2a: $J = \varnothing$. By (1), ${}^1f(P_I^{\uparrow y}, P_j, P_{-(I\cup\{j\})}) = x$.
  - Case 2b: $J \neq \varnothing$. With Q_j defined from P_j by putting z just below x, ${}^1f(Q_j, P_{-j}) = x$ by (1). Let $R_j := Q_j^{xz}$. By (2), ${}^1f(R_j, P_{-j}) \in \{x, z\}$.
    - Case 2b1: ${}^1f(R_j, P_{-j}) = x$. By (1), ${}^1f(P_I, R_j^{\uparrow z}, P_{-(I\cup\{j\})}) = x$. Since $P_j^{\uparrow z} = R_j^{\uparrow z}$, ${}^1f(P_I, P_j^{\uparrow z}, P_{-(I\cup\{j\})}) = x$.
    - Case 2b2: ${}^1f(R_j, P_{-j}) = z$. For $i \in J$, let $Q_i := P_i^{xy}$. By (1), ${}^1f(Q_J, R_j, P_{-(J\cup\{j\})}) = z$. By (2), ${}^1f(Q_J, Q_j, P_{-(J\cup\{j\})}) \in \{x, z\}$. If it is z, by (1), ${}^1f(P_J, Q_j, P_{-(J\cup\{j\})}) = z$, which contradicts ${}^1f(Q_j, P_{-j}) = x$. Thus, it is x. By (1), ${}^1f(Q_J, P_j, P_{-(J\cup\{j\})}) = x$. Also by (1), ${}^1f(Q_J^{\uparrow y}, P_{I\backslash J}^{\uparrow y}, P_j, P_{-(I\cup\{j\})}) = x$. As $P_J^{\uparrow y} = Q_J^{\uparrow y}$, ${}^1f(P_I^{\uparrow y}, P_j, P_{-(I\cup\{j\})}) = x$.■

Lemma GS2. If $n \geq 2 \leq m$ and $g : L^n \to A$ is SP then, for all $P \in L^n$, if $g(P) = x$ and $K := \{i \in N: {}^1P_i \neq x\} \neq \varnothing$ then, for some $v \in \{{}^1P_1, \ldots, {}^1P_n\}\backslash\{x\}$, $g(P_K^{\uparrow v}, P_{-K}) = x$.

Proof. In the proof of Lemma AR2, replace "${}^1f$", "(1)" and "(2)" by, respectively, "g", "(3)" and "(4)".■

Lemma AR3 asserts that the top in the social ranking is one of the tops in the individuals' rankings. This result requires IIA to be complemented by some assumption of the Paretian type, such as (5). The interpretation of Lemma GS3 is analogous.

Lemma AR3. If $n \geq 2 \leq m$ and $f : L^n \to L$ satisfies IIA and (5) below then, for all $P \in L^n$, ${}^1f(P) \in \{{}^1P_1, \ldots, {}^1P_n\}$.

$$\text{For all } P \in L^n \text{ and } x \in A, \text{ if } {}^1P_1 = \ldots = {}^1P_n = x \text{ then } {}^1f(P) = x. \tag{5}$$

Proof. Suppose not: for some $P \in L^n$, ${}^1f(P) = x \notin X := \{{}^1P_1, \ldots, {}^1P_n\}$. By successive application of Lemma AR2, for some $y \in X\backslash\{x\}$, ${}^1f(P_N^{\uparrow y}) = x$, which contradicts (5).■

Lemma GS3. If $n \geq 2 \leq m$ and $g : L^n \to A$ is SP and satisfies (6) below then, for all $P \in L^n$, $g(P) \in \{{}^1P_1, \ldots, {}^1P_n\}$.

$$\text{For all } P \in L^n \text{ and } x \in A, \text{ if } {}^1P_1 = \ldots = {}^1P_n = x \text{ then } g(P) = x. \quad (6)$$

Proof. In the proof of Lemma AR3, replace "${}^1f$", "(5)" and "AR2" by, respectively, "g", "(6)" and "GS2".■

By Lemma AR4, a SWF satisfying IIA and (5) is monotonic: if x is at the social top and raised in some individual ranking, x remains at the social top. Observe that this result requires at least three alternatives. The interpretation of Lemma GS4 is analogous.

Lemma AR4. Let $n \geq 2 < m$, $f : L^n \to A$ satisfy IIA and, for given $P \in L^n$, ${}^1f(P) = x$. If (5) holds and y covers x in P_i then ${}^1f(P_i^{xy}, P_{-i}) = x$.

Proof. By Remark AR6, PAR holds. By (2), ${}^1f(P_i^{xy}, P_{-i}) \in \{x, y\}$. Suppose it is y. With $z \in A \backslash \{x, y\}$, $I := \{j \in N \backslash \{i\}: xP_jy\}$ and $J := \{j \in N \backslash \{i\}: yP_jx\}$, let $R \in L^n$ differ from (P_i^{xy}, P_{-i}) only in that xR_izR_iy, xR_IzR_Iy and yR_JxR_Jz. By PAR and xR_Nz, ${}^1f(R) \neq z$. As ${}^1f(P_i^{xy}, P_{-i}) = y$, by IIA, ${}^1f(R) = y$. Let $S \in L^n$ differ from P only in that zS_iyS_ix, zS_IxS_Iy and yS_JzS_Jx. As ${}^1f(P) = x$, by IIA, ${}^1f(S) \in \{x, z\}$. As zS_Nx, by PAR, ${}^1f(S) \neq x$. Thus, ${}^1f(S) = z$ and IIA fails: $yf(R)z$, $zf(S)y$ but, for all $j \in N$, $R_j\big|_{\{y,z\}} = S_j\big|_{\{y,z\}}$.■

Lemma GS4. Let $n \geq 2 \leq m$, $g : L^n \to A$ be SP and, for given $P \in L^n$, $g(P) = x$. If y covers x in P_i then $g(P_i^{xy}, P_{-i}) = x$.

Proof. With $Q_i := P_i^{xy}$, assume $g(Q_i, P_{-i}) = z \in A \backslash \{x\}$. If zP_ix then $g(Q_i, P_{-i})P_ig(P_i, P_{-i})$, so g is not SP. If xP_iz then xQ_iz; that is, $g(P_i, P_{-i})Q_ig(Q_i, P_{-i})$ and g is not SP.■

By Lemma AR5: (i) $({}^1P_1, \ldots, {}^1P_n) = ({}^1Q_1, \ldots, {}^1Q_n)$ implies ${}^1f(P) = {}^1f(Q)$ and, therefore, f^* is a CF; and (ii) ${}^1f(P) = x$ makes $\{i \in N: {}^1P_i = x\}$ decisive for x in f^*. Thus, a SWF satisfying IIA and (5), and hence PAR, induces a CF satisfying A2. Lemma GS5 is an analogous result for SP SCFs satisfying (6) (see Remark GS6).

Lemma AR5. Let $n \geq 2 < m$ and $f : L^n \to L$ satisfy IIA and (5). Then, for all $P \in L^n$, ${}^1f(P) = x$ implies that $\{i \in N: {}^1P_i = x\}$ is decisive for x in f.

Proof. Let ${}^1f(P) = x$ and $I := \{i \in N: {}^1P_i = x\}$. By a simple induction argument it suffices to choose $j \in N$ and show that ${}^1f(Q_j, P_{-j}) = x$, with Q_j obtained from P_j by exchanging two contiguous alternatives v and z such that $x \notin \{v, z\}$ if $j \in I$. So let $j \in N$ and vP_jz.

- Case 1: $x \notin \{v, z\}$. By (1), ${}^1f(P) = x$ yields ${}^1f(Q_j, P_{-j}) = x$.
- Case 2: $z = x$. Lemma AR4.

• Case 3: $v = x$. By (2) and ${}^1f(P) = x$, ${}^1f(Q_j, P_{-j}) \in \{x, z\}$. With ${}^1P_j = y \in A\backslash\{x, z\}$, it rests to derive a contradiction from ${}^1f(Q_j, P_{-j}) = z$. By Lemma AR3, $K := \{i \in N: {}^1P_k = z\} \neq \emptyset$. Let $J := \{i \in K: xP_iy\}$. For $k \in J$, let $Q_k := P_k^{xy}$.

• Case 3a: $J \neq \emptyset$. By (1), ${}^1f(Q_j, P_{-j}) = z$ implies ${}^1f(Q_j, Q_J, P_{-(J\cup\{j\})}) = z$. By (2), ${}^1f(P_j, Q_J, P_{-(J\cup\{j\})}) \in \{x, z\}$. If z, by (1), ${}^1f(P_j, P_J, P_{-(J\cup\{j\})}) = z$, contradicting ${}^1f(P) = x$. Since it is x, for $k \in J$, define $R_k := Q_k^{yz}$ and, for $k \in K\backslash J$, define $R_k := P_k^{yz}$. By (1) and ${}^1f(P_j, Q_J, P_{-(J\cup\{j\})}) = x$, ${}^1f(P_j, R_K, P_{-(K\cup\{j\})}) = x$. This and (2) yield ${}^1f(Q_j, R_K, P_{-(K\cup\{j\})}) \in \{x, z\}$. It is not z by Lemma AR3. Hence, it is x. By (1), ${}^1f(Q_j, Q_J, P_{K\backslash J}, P_{-(K\cup\{j\})}) = x$. By Lemma AR4, ${}^1f(Q_j, P_J, P_{K\backslash J}, P_{-(K\cup\{j\})}) = x$, contradicting ${}^1f(Q_j, P_{-j}) = z$.

• Case 3b: $J = \emptyset$. By Lemma AR4, (1) and ${}^1f(P) = x$, ${}^1f(Q_K, P_{-K}) = x$. By (1), ${}^1f(Q_j, P_{-j}) = z$ implies ${}^1f(Q_j, Q_K, P_{-(K\cup\{j\})}) = z$, which is an instance of case 3a.■

Lemma GS5. Let $n \geq 2 < m$ and $g : L^n \to A$ be SP and satisfy (6). Then, for all $P \in L^n$, $g(P) = x$ implies that $\{i \in N: {}^1P_i = x\}$ is decisive for x in g.

Proof. In the proof of Lemma AR5, replace "${}^1f$", "(1)", "(2)", "AR3" and "AR4" by, respectively, "g", "(3)", "(4)", "GS3" and "GS4".■

Remark AR6. If $f : L^n \to L$ satisfies IIA then f satisfies PAR if, and only if, (5) holds.

Remark GS6. By Lemma GS4, if $g : L^n \to A$ is SP then g is onto if, and only if, (6) holds.

Arrow's theorem: If $n \geq 2 < m$ then every social welfare function $f : L^n \to L$ satisfying PAR and IIA is dictatorial.

Proof. By Remark AR6, PAR implies (5). By Lemma AR5, f^* is a function satisfying A2. By Lemma AR3, f^* satisfies A1. By Proposition 3.2, some $\{i\}$ is decisive in f^*. By Lemma AR5, $\{i\}$ decisive in f.■

The Gibbard-Satterthwaite theorem: If $n \geq 2 < m$ then every onto and strategy-proof social choice function $g : L^n \to A$ is dictatorial.

Proof. In the proof of Arrow's theorem, replace "AR6", "PAR", "(5)", "AR5", "f" and "AR3" by, respectively, "GS6", "ontoness", "(6)", "GS5", "g" and "GS3".■

5. Comments

It is worth recapitulating the connections and similitudes between Arrow's and the Gibbard-Satterthwaite theorem that emerge from the results in Section 4.

To begin with, both theorems can be reduced to a common impossibility result for CFs: Proposition 3.2. On the one hand, this suggests that both theorems could be viewed as unsuccessful attempts to escape from that impossibility result. On the other, it indicates that it is as if a SWF sastisfying the conditions in Arrow's theorem generated the social top, and the way a SCF sastisfying the conditions in the Gibbard-Satterthwaite theorem selected the social choice, by resorting to an extremely simple rule: a CF (which in addition happens to be dictatorial).

In each theorem, the two assumptions are necessary for the CF reduction to be possible. In fact, IIA would be objectionable if a SWF satisfying IIA made f^* a function, as this would mean that the social top is determined disregarding how individuals rank below the top. The same objection could be adduced against strategy-proofness if a SP SCF made g^* a function. The following examples prove that, for $n \geq 2 < m$, none of these objections can be raised: the inversely dictatorial SWF (there is $i \in N$ such that, for all $P \in L^n$, ${}^1f(P) = {}^mP_i$) satisfies IIA but f^* is not a function; and the SCF g such that, for some $x \in A$, some $i \in N$ and all $P \in L^n$, $g(P) = {}^1P_i$ if ${}^1P_i \neq x$ and $g(P) = {}^2P_i$ if ${}^1P_i = x$, is SP but g^* is not a function.

This notwithstanding, (essentially) the same assumption ((5) for SWFs and (6) for SCFs) added to IIA and SP yields Arrow's and the Gibbard-Satterthwaite theorem, respectively. The weakness of that assumption suggests that IIA and SP put SWFs and SCFs too close to dictatorship (Wilson's (1972) theorem is already evidence of this observation for SWFs; see Saari (1998) for an illuminating analysis of IIA).

But perhaps the most interesting connection between IIA and SP that the proofs in Section 4 reveal comes from Lemmas AR2 and GS2. Actually, these results sustain the following common criticism to IIA and SP: for $n \geq 3 \leq m$, both IIA and SP are inconsistent with the mild majority principle MP according to which, if a strict majority of individuals (but not all of them) rank some alternative x at the top, then x must be at the top of the social ranking for SWFs and must be the social choice for SCFs. To see this, let x, y and z be different alternatives and $\{I, J, K\}$ a partition of N such that I and J have: (i) $n/3$ members if n is a multiple of 3; (ii) $(n + 1)/3$ members if $n + 1$ is a multiple of 3; and (ii) $(n + 2)/3 - 1$ members if $n + 2$ is a multiple of 3. Consider $P \in L^n$ such that, for all $i \in I$, ${}^1P_i = x$; for all $i \in J$, ${}^1P_i = y$; and for all $i \in K$, ${}^1P_i = z$. Assume IIA. If ${}^1f(P) = w \notin \{x, y, z\}$, by Lemma AR2, there are $v \in \{x, y, z\}$, $t \in \{x, y, z\}\backslash\{v\}$ and $i \in N$ such that ${}^1P_i = t$ and ${}^1f(P_i, P_{-i}{}^{\uparrow v}) = w$, which contradicts MP. If, on the other hand, ${}^1f(P) \in \{x, y, z\}$ then, assuming without loss of generality that ${}^1f(P) = x$, by Lemma AR2, ${}^1f(P_I, P_J, P_K{}^{\uparrow y}) = x$, contradicting MP. The same result applies to SP SCFs by invoking Lemma GS2 instead of Lemma AR2.

References

Arrow, K.: *Social Choice and Individual Values*, 2nd ed. New York: Wiley 1963

Barberà, S.: A theorem on preference aggregation, mimeo, *Universitat Autònoma de Barcelona*, Spain (2001)

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

Gibbard, A.: Manipulation of voting schemes: A general result. *Econometrica* 41, 587–601 (1973)

Reny, P.J.: Arrow’s theorem and the Gibbard-Satterthwaite theorem: A unified approach. *Economics Letters* 70, 99–105 (2001)

Saari, D.: Connecting and resolving Sen’s and Arrow’s theorems. *Social Choice and Welfare* 15, 239–261 (1998)

Satterthwaite, M.: Strategy-proofness and Arrow’s conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory* 10, 187–217 (1975)

Wilson, R.: Social choice without the Pareto principle. *Journal of Economic Theory* 5, 14–20 (1972)