A UNIFYING IMPOSSIBILITY THEOREM

PRISCILLA T. Y. MAN AND SHINO TAKAYAMA

Abstract. This paper considers social choice correspondences assigning a choice set to each non-empty subset of social alternatives. We impose three requirements on these correspondences: unanimity, independence of preferences over infeasible alternatives and choice consistency with respect to choices out of all possible alternatives. With more than three social alternatives and the universal preference domain, any social choice correspondence that satisfies our requirements is serially dictatorial. A number of known impossibility theorems — including Arrow's Impossibility Theorem, the Muller-Satterthwaite Theorem and the impossibility theorem under strategic candidacy — follow as corollaries. Our new proof highlights the common logical underpinnings behind these theorems.

1. Introduction

This paper considers social choice correspondences defined on all subsets of social alternatives. Three axioms are imposed:

- **Strong Unanimity:** Only the unique weakly Pareto dominant alternative within the subset is chosen whenever there is one.
- **Independence of Infeasible Alternatives:** Choices from subsets depend only on the preferences over the subsets.
- **Independence of Losing Alternatives:** An alternative is chosen from a subset if and only if it is chosen out of the set of all social alternatives whenever such a choice remains available.

With more than three social alternatives and the universal preference domain, any social choice correspondence that satisfies our three axioms is serially dictatorial. Since our axioms

Date: January 2, 2012.
Key words and phrases. Impossibility Theorem, Independence of Infeasible Alternatives, Independence of Losing Alternatives, Social Choice.
This paper was previously circulated under the title "Three Impossibility Theorems, Two Independence Axioms and One Unified Proof". We thank Hugo Sonnenschein for suggesting a new title. We would also like to thank Salvador Barberà, Matthew Jackson, Andy McLennan, Phil Reny, Arunava Sen, William Thomson, John Weymark, participants at the "Votes and Incentives: the Design of Institutions" Workshop at Universitat Autònoma de Barcelona and four anonymous referees for their helpful comments. Takayama gratefully acknowledges the financial support from the Australian Research Council Grant (DP1093105) and the Early Career Researcher Grant at the University of Queensland. Man is grateful for the financial support of the University of Queensland Grant 2010002370, as well as the hospitality of Takayama and Andy McLennan during the Brisbane flood in 2011, when this paper is written.

imply Arrow's Choice Axiom (an alternative is chosen out of a subset if and only if it is chosen out of a superset whenever choices from the superset are available), our theorem can be viewed as a choice-theoretic version of Arrow's theorem with the Pareto and rationality axioms weakened as much as possible.

Weakening Arrow's axioms is more than an aesthetic exercise. It clarifies the basic axioms on social choice that imply dictatorship. Arrow's Choice Axiom has been interpreted as a rationality axiom (Arrow, 1959; Hansson, 1968). Our result indicates, however, that this rationality condition follows from three fundamental requirements: unanimity, independence of preferences over infeasible alternatives and choice consistency with respect to choices out of all possible alternatives. Our formulation also provides the logical link between Arrow's Independence of Irrelevant Alternatives and Arrow's Choice Axiom — two related but different concepts that have caused some confusions[1].

More importantly, it is easy to check that axioms in many impossibility theorems imply our three weakened axioms (with an appropriate extension to our unrestricted domain). Thus a number of impossibility theorems — including Arrow's Impossibility Theorem (Arrow, 1963), the Gibbard-Satterthwaite Theorem (Gibbard, 1973; Satterthwaite, 1975), the Muller-Satterthwaite Theorem (Muller and Satterthwaite, 1977), the Jackson-Srivastava characterization of game theoretic solutions that implement only dictatorial social choice functions (Jackson and Srivastava, 1996), the Grether-Plott Theorem (Grether and Plott, 1982) and the Dutta-Jackson-Le Breton impossibility theorem under strategic candidacy (Dutta, Jackson, and Le Breton, 2001) — are corollaries of our main theorem. While it is known that one can convert the questions in many of these theorems into a social welfare function problem and apply Arrow's Theorem, our unified proof indicates that there are common and intuitive principles — unanimity and two independence axioms — behind these theorems.

The weakened axioms also point us to a new line of proof. Example 2.5 shows that our three axioms do not immediately imply Monotonicity (a chosen alternative remains chosen whenever its relative ranking has improved), which is a key property in proofs of impossibility theorems (e.g.: Barberà (1980, 1983); Benoît (2000); Sen (2001); Geanakoplos (2005)). Rather than following the usual practice of moving an alternative up a chain of preference profiles to find the pivotal individual, our proof exploits a fact about social choice when the union of individuals' set of favorite alternatives has only two elements (Lemma 3.2). This technique employs fewer preference profiles and shed light on the set of preference profiles necessary for proving an impossibility theorem.

[1]The original definition of Independence of Irrelevant Alternatives (Arrow, 1963, p. 27) corresponds to our Independence of Infeasible Alternatives. However, Arrow's motivating examples — death of a candidate (p. 26) and rank-order voting (p. 27) — are violations of our Independence of Losing Alternatives, hence Arrow's Choice Axiom. See Denicolò (2000) for a discussion.

Needless to say, this is not the first paper to provide an alternative proof of (a version of) Arrow's Theorem[2], nor is it the first attempt to offer a unified proof of several impossibility theorems. Reny (2001) provides parallel proofs of Arrow's Impossibility Theorem and the Gibbard-Satterthwaite Theorem, upon which Eliaz (2004) builds a single proof of the two theorems. Using a social preference framework, Barberà (2003) highlights the common properties shared by the social aggregation rules underlying several impossibility theorems. All these proofs require some version of Monotonicity, which is not directly implied by our axioms. Recently, Vohra (2011) proves Arrow's Impossibility Theorem, the Gibbard-Satterthwaite Theorem and the strategic candidacy impossibility theorem using integer programming techniques. Unlike us, he uses Arrow's Theorem to prove the other two theorems.

2. Model

Let X be the set of all possible social alternatives, $|X| \geq 3$ and finite[3]. Let R be the set of all complete, transitive binary relations on X, that is, the set of all weak preferences over X. Let $P \subset R$ be the set of all strict preferences over X. We abuse notation to use N as both the set of all individuals and its cardinality, which is finite.

A (weak) preference of individual $i \in N$ is denoted by $\succsim_i \in R$. The symbols $\succ_i$ and $\sim_i$ will have their usual derived meanings. Let $\succsim = (\succsim_1, \ldots, \succsim_N) \in R^N$ be a preference profile. Similarly, a strict preference of an individual $i \in N$ and a strict preference profile are denoted as $\succ_i \in P$ and $\succ \in P^N$ respectively.

Two preference profiles $\succsim, \succsim' \in R^N$ *agree on* Y if they induce the same preference ordering on the subset of alternatives $Y \subseteq X$. Given a subset of alternatives $Y \subseteq X$ and a preference profile $\succsim \in R^N$, we say $\succsim^Y \in R^N$ *takes* Y *to the top from* $\succsim$ if $\succsim$ and $\succsim^Y$ agree on Y and $y \succ_i^Y z$ for all i whenever $y \in Y$ and $z \notin Y$.

A *social choice correspondence* is a mapping $f : (2^X \setminus \emptyset) \times R^N \to 2^X \setminus \emptyset$ such that $f(Y, \succsim) \subseteq Y$ for all $Y \subseteq X$ and all $\succsim \in R^N$. We impose the following axioms on a social choice correspondence:

Definitions 2.1 (Main Axioms)**.** A social choice correspondence f is

Strongly Unanimous (SU) if $f(Y, \succsim) = \{y\}$ whenever y is uniquely weakly Pareto dominant in $Y \subseteq X$ (i.e., for all $y' \in Y \setminus \{y\}$, $y \succsim_i y'$ for all $i \in N$ with at least one individual having a strict preference);

[2]See Campbell and Kelly (2002) for an overview.

[3]Finiteness of X ensures the existence of best alternatives on any feasible set according to any preferences. This assumption can be relaxed at the cost of distracting technical qualifications, which we would like to avoid.

Independent of Infeasible Alternatives (IIF)[4] if $f(Y, \succsim) = f\left(Y, \succsim'\right)$ whenever $\succsim$ and $\succsim'$ agree on $Y \subseteq X$;

Independent of Losing Alternatives (ILA) if $f(Y, \succsim) = f(X, \succsim) \cap Y$ whenever the intersection is non-empty.

Strong Unanimity by itself is weaker than Strong Pareto Optimality (any weakly Pareto dominated alternative cannot be chosen). Together with ILA, though, it does imply that weakly Pareto dominated alternatives cannot be chosen out of the set of all alternatives:

Claim 2.2. If a social choice correspondence f satisfies SU and ILA, then $x \notin f(X, \succsim)$ whenever x is weakly Pareto dominated at $\succsim$ (i.e., there exists a $y \in X$ such that $y \succsim_i x$ for all $i \in N$ with at least one strict preference).

Proof. Suppose y weakly Pareto dominates x at $\succsim$. SU implies $f(\{x, y\}, \succsim) = \{y\}$. Thus $x \notin f(X, \succsim)$ or else ILA will be violated. □

Of more interests is the relationship between our axioms and two other axioms commonly employed in social choice theory, defined as follows:

Definitions 2.3. A social choice correspondence f satisfies

Arrow's Choice Axiom (ACA)[5] if $f(Z, \succsim) = f(Y, \succsim) \cap Z$ for any $Y \supseteq Z$ whenever the intersection is non-empty;

Monotonicity if $Y \subseteq f(X, \succsim') \subseteq f(X, \succsim)$ whenever $Y \subseteq f(X, \succsim)$ and $y \succsim_i x$ implies $y \succsim'_i x$ (with $y \succ_i x$ implies $y \succ'_i x$) for all $y \in Y$, all $x \in X$ and all $i \in N$.[6]

Claim 2.4. A social choice correspondence satisfies ACA if it satisfies SU, IIF and ILA.

Proof. Let $Z \subseteq Y \subseteq X$ and suppose $f(Y, \succsim) \cap Z \neq \emptyset$. Obtain $\succsim^Y$ by taking Y to the top of $\succsim$. Claim 2.2, ILA and IIF imply $f(X, \succsim^Y) = f(Y, \succsim^Y) = f(Y, \succsim)$, which has a non-empty intersection with Z. ILA requires $f(Z, \succsim^Y) = f(X, \succsim^Y) \cap Z$. Using $f(X, \succsim^Y) = f(Y, \succsim)$ and $f(Z, \succsim) = f(Z, \succsim^Y)$ (IIF, since $Z \subseteq Y$), we get $f(Z, \succsim) = f(Y, \succsim) \cap Z$. □

Denicolò (2000) shows that a Pareto optimal social choice function satisfies both IIF and ACA if and only if it satisfies Hansson's Independence (Hansson, 1969). Claim 2.4 improves upon his statement by giving the conditions that bridge the gap between IIF and ACA.

[4]This condition is best known as "Independence of Irrelevant Alternatives" (Arrow, 1963). "Independence of Infeasible Alternatvies" (Ehlers and Weymark, 2003; Le Breton and Weymark, 2011) is more precise on the "irrelevance" of alternatives outside Y. To avoid confusion, we follow Le Breton and Weymark (2011) in using the acronym IIF and reserve IIA for Independence of Irrelevant Alternatives. Another option is "Independence of Nonfeasible Alternatives" (Karni and Schmeidler, 1976).

[5]Arrow's Choice Axiom has also been known as Independence of Nonoptimal Alternatives (Karni and Schmeidler, 1976) and Strong Stability (Campbell, 1979).

[6]This definition coincides with Strong Positive Association (Muller and Satterthwaite, 1977) when f is a function.

$\succ$			$\succ'$		
1	2	3	1	2	3
x	y	z	x	y	z
y	z	y	y	z	x
z	x	x	z	x	y

(a) Preference Profiles

Y	$f(Y,\succ)$	$f(Y,\succ')$
X	x	y
$\{x,y\}$	x	y
$\{x,z\}$	x	x
$\{y,z\}$	y	y

(b) Social Choices

FIGURE 1. Preferences and Social Choices for Example 2.5

One can stop here and prove a choice-theoretic version of Arrow's theorem (c.f.: Le Breton and Weymark, 2011, Theorem 19) in the usual manner — convert the problem into a social preference problem and apply Arrow's Impossibility Theorem. This route calls for a proof of Arrow's Impossibility Theorem, which typically invokes Monotonicity. Intriguingly, SU, IIF and ILA do not immediately imply Monotonicity[7], as the next example shows.

Example 2.5. There are 3 alternatives, $X = \{x, y, z\}$ and 3 individuals, $N = \{1, 2, 3\}$. The preference domain admits only two strict preference profiles, $\succ$ and $\succ'$, as depicted in Figure 1(a).[8] Note that $\succ'$ differs from $\succ$ only by moving x above y for individual 3. The social choice function given by Figure 1(b) satisfies SU, IIF and ILA[9] but not monotonicity. To see this, first note that SU has no bite in this example. The two profiles $\succ$ and $\succ'$ agree only on $\{x, z\}$ and $\{y, z\}$. IIF is satisfied since choices from these subsets are equal across preference profiles. ILA is satisfied since $f(X,\succ) = \{x\}$ and $f(Y,\succ) = \{x\}$ for all subsets Y containing x; and similarly $f(X,\succ') = \{y\}$ and $f(Y,\succ') = \{y\}$ for all subsets Y containing y. Yet Monotonicity is violated as $f(X,\succ) = \{x\}$ and the ranking of x improves from $\succ$ to $\succ'$ but $f(X,\succ') = \{y\}$.

Thus, as a direct proof (i.e., one that does not appeal to Arrow's Impossibility Theorem), our proof in Section 3 offers a new approach for proving impossibility theorems. Before doing so, we need to define dictatorship.

For any subset of alternatives $Y \subseteq X$ and any individual preference $\succsim_i \in R$, let

$$T(Y,\succsim_i) = \{y \in Y : y \succsim_i y' \text{ for all } y' \in Y\}$$

be the "top set" — the set of favorite alternatives — within Y according to $\succsim_i$. Let $\pi = (\pi_1, \ldots, \pi_N)$ be a permutation of the set of all individuals. Write also $\pi^k = (\pi_1, \ldots, \pi_k)$

[7]As a trivial corollary of our main theorem, any social choice correspondence that is SU, IIF and ILA is Monotonic. This example indicates such an implication is hard to establish without first proving dictatorship.
[8]Since the relative ranking between x and y changes across the profiles, introducing $\succ^{\{x,y\}}$ by bringing $\{x, y\}$ to the top does not help.
[9]Since there are only three alternatives, ILA is equivalent to ACA. Thus this example remains even if we assume ACA.

for any $k \leq N$. Given a preference profile $\succsim$ and π^k, write $\succsim^{\pi^k} = (\succsim_{\pi_1}, \ldots, \succsim_{\pi_k})$ as the preferences of individuals in π^k. For all subsets Y, all preference profiles $\succsim$ and all permutations of individuals π, define the kth iteration of the top set operator such that

$$T^0(Y, \succsim^{\pi^0}, \pi^0) = Y$$
$$T^k(Y, \succsim^{\pi^k}, \pi^k) = T(T^{k-1}(Y, \succsim^{\pi^{k-1}}, \pi^{k-1}), \succsim_{\pi_k}).$$

Definitions 2.6 (Dictatorship)**.** A social choice correspondence f is

Dictatorial if there exists an individual $i \in N$ such that $f(Y, \succsim) \subseteq T(Y, \succsim_i)$ for all $Y \subseteq X$ and all $\succsim \in R^N$;

Serially Dictatorial if there exists a permutation of individuals π and a tie-breaking preference $\rho \in R$ such that

$$f(Y, \succsim) = T(T^N(Y, \succsim, \pi), \rho) \quad \text{for all } Y \subseteq X, \text{ all } \succsim \in R^N.$$

Two features of our definition of serial dictatorship should be noted. First, the permutation of individuals π and the tie-breaking preference ρ are fixed for all preference profiles. This rules out serial dictatorship in which the order of later dictators depends on the preferences of earlier dictators. Second, a tie-breaking preference ρ is applied at the end. If f is required to be a function, we can choose some $\rho \in P$.[10] On the other extreme, we can set ρ as the preference that is indifferent between all alternatives if no ties are to be broken.

3. Main Theorem

Theorem 3.1. *Any social choice correspondence that is Strongly Unanimous, Independent of Infeasible Alternatives and Independent of Losing Alternatives is Serially Dictatorial.*

The rest of this section contains the proof. We will first give a useful lemma. It will then be used in Section 3.1 and 3.2, which construct the permutation π and the tie-breaking preference ρ, respectively. Our argument also indicates that if the preference domain is restricted to P^N, any social choice correspondence that satisfies SU, IIF and ILA is dictatorial.

Lemma 3.2. *Let f be a SU, IIF and ILA social choice correspondence and $\succsim$ be a preference profile. Let $S \subseteq N$ be the set of individuals who are not indifferent between all alternatives at $\succsim$. Then*

$$f(X, \succsim) \subseteq \bigcup_{i \in S} T(X, \succsim_i)$$

whenever $\left|\bigcup_{i \in S} T(X, \succsim_i)\right| \leq 2$.

[10] This will not violate Strong Unanimity since SU applies only when there is a *unique* weakly Pareto dominant alternative.

Type x	Type y	Type xy
x	y	$x\,y$
·	·	·
y	·	·
·	x	·
·	·	·

(a) Profile $\succsim$

Type x	Type y	Type xy
x	y	$x\,y$
·	x	·
y	·	·
·	·	·
·	·	·

(b) Profile $\succsim'$

Type x	Type y	Type xy
x	y	$x\,y$
y	·	·
·	·	·
·	x	·
·	·	·

(c) Profile $\succsim''$

FIGURE 2. Preference Profiles for Proof of Lemma 3.2

Proof. The case of $\left|\bigcup_{i\in S} T(X,\succsim_i)\right| = 1$ follows from SU. So let $\bigcup_{i\in S} T(X,\succsim_i) = \{x,y\}$ and suppose by contradiction that $z \in f(X,\succsim)$ for some $z \neq x, y$.

Since $\bigcup_{i\in S} T(X,\succsim_i) = \{x,y\}$, there are three types of individuals in S: those whose favorite is x (Type x), y (Type y) and those whose favorites are x and y (Type xy) (Figure 2(a)). Construct a new preference profile $\succsim'$ by moving, for all Type y individuals, the ranking of x up to just below y, keeping all else unchanged (Figure 2(b)). Similarly, construct $\succsim''$ by moving y to just below x for all Type x individuals (Figure 2(c)). Observe that

(1) $\succsim$, $\succsim'$ and $\succsim''$ agree on $\{x,y\}$;
(2) $\succsim$ and $\succsim'$ agree on $\{y,z\}$; and
(3) $\succsim$ and $\succsim''$ agree on $\{x,z\}$.

Since $z \in f(X,\succsim)$, ILA requires $z \in f(\{y,z\},\succsim)$. By observation (2) and IIF, $z \in f\left(\{y,z\},\succsim'\right)$. Meanwhile, all alternatives other than x and y are weakly Pareto dominated by x at $\succsim'$. Claim 2.2 says none of them can be chosen out of X at $\succsim'$. Thus $y \notin f(X,\succsim')$, otherwise $z \in f\left(\{y,z\},\succsim'\right) \neq \{y\} = f\left(X,\succsim'\right) \cap \{y,z\}$, which violates ILA. Therefore $f\left(X,\succsim'\right) = \{x\}$. By a similar argument using the subset $\{x,z\}$ and observation (3), $f\left(X,\succsim''\right) = \{y\}$.

Now ILA requires $f\left(\{x,y\},\succsim'\right) = \{x\}$ and $f\left(\{x,y\},\succsim''\right) = \{y\}$. This contradicts IIF in light of observation (1). □

3.1. **Serial Dictators.** Given a SU, IIF and ILA social choice correspondence f, we construct in this subsection the permutation of individuals π such that for all $k \geq 0$,

$$f(Y,\succsim) \subseteq T^k(Y,\succsim^{\pi^k},\pi^k) \quad \text{for all } Y \subseteq X, \text{ all } \succsim \in R^N. \tag{1}$$

The case for $k = 0$ follows by definition. Now suppose π^{k-1} is defined and Equation (1) holds for $k-1$. We construct π_k that satisfies Equation (1) for k in 3 steps: Step 1 identifies a group of individuals containing the desired π_k. Step 2 shows that whenever this group of individuals have the same preferences over $Y \subseteq X$, the social choice out of Y is always a subset of their favorites in $T^{k-1}(Y,\succsim^{\pi^{k-1}},\pi^{k-1})$. Step 3 shrinks this group to a singleton,

Type x	Type y	Type z
x	y	z
y	z	x
z	x	y
$\vdots$	$\vdots$	$\vdots$

FIGURE 3. Condorcet Cycle

Group S	Not S
x	z
y	$\cdot$
z	$\cdot$
$\vdots$	$\vdots$

(a) Profile $\succsim^1$

Group S	Not S
y	z
x	$\cdot$
z	$\cdot$
$\vdots$	$\vdots$

(b) Profile $\succsim^2$

FIGURE 4. Preference Profiles for Step 2.1

giving us π_k. Step 4 is given when the above 3-step proof is infeasible (this happens if $N < 3$ or $k > N - 2$).

Step 1. If $N < 3$ or $k > N - 2$ proceed directly to Step 4. Otherwise, construct a preference profile $\succsim^*$ where (1) $\pi_1, \dots \pi_{k-1}$ are indifferent between all alternatives; (2) the remaining individuals have one of the Condorcet preferences in Figure 3; and (3) all individual preferences in the Condorcet cycle are assigned to at least one individual.

Since all alternatives other than x, y and z are weakly Pareto dominated, by Claim 2.2, $f(X, \succsim^*) \subseteq \{x, y, z\}$. Without loss assume $x \in f(X, \succsim^*)$. Let $S \subseteq N \setminus \{\pi_1, \dots, \pi_{k-1}\}$ be the set of individuals who have the Type x preference in the Condorcet cycle.

Step 2. We show whenever individuals in S have the same preference over $Y \subseteq X$, the social choice from Y is a subset of their favorites in $Y_{k-1} \equiv T^{k-1}(Y, \succsim^{\pi^{k-1}}, \pi^{k-1})$. This is done by three smaller steps: Step 2.1 proves this for $Y = \{x, y\}$ (the top 2 alternatives in the Type x Condorcet preference); Step 2.2 for any two-element subset Y; and Step 2.3 for any subset $Y \subseteq X$.

Step 2.1. We show whenever individuals in S have the same preferences over $\{x, y\}$, the social choice from $\{x, y\}$ is a subset of their favorites in $\{x, y\}_{k-1} \equiv T^{k-1}(\{x, y\}, \succsim^{\pi^{k-1}}, \pi^{k-1})$. This is trivial if $\{x, y\}_{k-1}$ is a singleton or if $x \sim_i y$ for all $i \in S$.

Construct a preference profile $\succsim^1$ where (1) $\pi_1, \dots, \pi_{k-1}$ are indifferent between all alternatives; (2) all individuals in S have the Type x Condorcet preference; and (3) all other individuals' unique favorite is z (Figure 4(a)). Notice that $\succsim^1$ and $\succsim^*$ agree on $\{x, z\}$. Since

Group S	Not S
w	y
z	$\cdot$
x	$\cdot$
y	$\cdot$
$\vdots$	$\vdots$

FIGURE 5. Preference Profile $\succsim^3$ for Step 2.2

x is chosen out of X at $\succsim^*$, IIF and ILA imply $x \in f(\{x,z\},\succsim^1) = f(\{x,z\},\succsim^*)$. Meanwhile, Lemma 3.2 requires $f(X,\succsim^1) \subseteq \{x,z\}$. By ILA $x \in f(X,\succsim^1)$. Since $y \notin f(X,\succsim^1)$ (Lemma 3.2), applying ILA once more gives $f(\{x,y\},\succsim^1) = \{x\}$. Notice that $\succsim^1$ puts no restriction on the relative ranking between x and y for individuals not in $\{\pi_1,\ldots,\pi_{k-1}\}\cup S$. IIF therefore implies $f(\{x,y\},\succsim) = \{x\}$ for all $\succsim$ such that $\pi_1,\ldots,\pi_{k-1}$ are indifferent between x and y and all individuals in S strictly prefer x to y.

Next construct a preference profile $\succsim^2$ from $\succsim^1$ by switching the positions of x and y in $\succsim^1_S$ and keeping everything else unchanged (Figure 4(b)). Since $\succsim^1$ and $\succsim^2$ agree on $\{x,z\}$, IIF requires $x \in f(\{x,z\},\succsim^2)$. Yet Lemma 3.2 requires $f(X,\succsim^2) \subseteq \{y,z\}$, so $z \notin f(X,\succsim^2)$ or else ILA would be violated. Hence $f(X,\succsim^2) = \{y\}$. Applying ILA once more gives $f(\{x,y\},\succsim^2) = \{y\}$. Since $\succsim^2$ puts no restriction on the relative ranking between x and y for individuals not in $\{\pi_1,\ldots,\pi_{k-1}\}\cup S$, IIF implies $f(\{x,y\},\succsim) = \{y\}$ for all $\succsim$ such that $\pi_1,\ldots,\pi_{k-1}$ are indifferent between x and y and all individuals in S strictly prefer y to x.

Step 2.2. We show whenever all individuals in S have the same preferences over a two-element subset $\{w,z\} \subseteq X$, the social choice from $\{w,z\}$ is a subset of their favorites in $\{w,z\}_{k-1} \equiv T^{k-1}(\{w,z\},\succsim^{\pi^{k-1}},\pi^{k-1})$. The statement is trivial if $\{w,z\}_{k-1}$ is a singleton or if $w \sim_i z$ for all $i \in S$. There is also nothing to prove if $\{w,z\} = \{x,y\}$. So without loss assume $y \notin \{w,z\}$.

Create a preference profile $\succsim^3$ where (1) $\pi_1,\ldots,\pi_{k-1}$ are indifferent between all alternatives; (2) all individuals in S have w as their unique favorite and strictly prefer x to y; and (3) all other individuals' unique favorite is y (Figure 5). By Lemma 3.2, $f(X,\succsim^3) \subseteq \{w,y\}$. However, all individuals in S strictly prefer x to y. Step 2.1 implies $f(\{x,y\},\succsim^3) = \{x\}$. ILA then requires $f(X,\succsim^3) = \{w\}$. Applying ILA once more gives $f(\{w,z\},\succsim^3) = \{w\}$. Since $\succsim^3$ puts no restriction on the relative ranking between w and z for individuals not in $\{\pi_1,\ldots,\pi_{k-1}\}\cup S$, IIF implies $f(\{w,z\},\succsim) = \{w\}$ for all $\succsim$ such that $\pi_1,\ldots,\pi_{k-1}$ are indifferent between w and z and all individuals in S strictly prefer w to z.

Switching the names of w and z gives $f(\{w,z\},\succsim) = \{z\}$ for all $\succsim$ such that $\pi_1,\ldots,\pi_{k-1}$ are indifferent between w and z and all individuals in S strictly prefer z to w.

Step 2.3. We show whenever all individuals in S have the same preferences over $Y \subseteq X$, the social choice from Y is a subset of their favorites in $Y_{k-1} \equiv T^{k-1}(Y, \succsim^{\pi^{k-1}}, \pi^{k-1})$.

Given a preference profile $\succsim$ at which all individuals in S have the same preferences over Y, obtain $\succsim^Y$ by taking Y to the top from $\succsim$ (see Section 2). If $k > 1$, the induction hypothesis (Equation (1)) ensures $f(X, \succsim^Y) \subseteq Y_{k-1}$. Otherwise, $f(X, \succsim^Y) \subseteq Y = Y_0$ since all alternatives not in Y are Pareto dominated (Claim 2.2). Moreover, $\pi_1, \ldots, \pi_{k-1}$ are indifferent between all alternatives in Y_{k-1}.

If y is a favorite for group S in Y_{k-1} and $y' \in Y_{k-1}$ is not, Step 2.2 requires $f(\{y, y'\}, \succsim^Y) = \{y\}$. ILA implies $y' \notin f(X, \succsim^Y)$. Thus $f(X, \succsim^Y)$ is a subset of group S's favorites in Y_{k-1}. Applying ILA once more gives $f(Y, \succsim^Y) \subseteq T(Y_{k-1}, \succsim_S^Y)$. Since $f(Y, \succsim) = f(Y, \succsim^Y)$ (IIF) and $T(Y_{k-1}, \succsim_S^Y) = T(Y_{k-1}, \succsim_S)$, $f(Y, \succsim) \subseteq T(Y_{k-1}, \succsim_S)$.

Remark on Step 2 Step 2 implies that the choice out of X at any Condorcet profile is a singleton. For if not there will be at least two disjoint subsets of individuals that can get their favorites out of Y_{k-1} whenever preferences over Y within each group are the same. Contradiction arises when the preferences of these groups conflict with each other.

Step 3 If S is a singleton, letting $\pi_k = S$ completes our induction step. Otherwise, construct a preference profile $\succsim^{**}$ where (1) $\pi_1, \ldots, \pi_{k-1}$ are indifferent between all alternatives; (2) all individuals in S have either the Type x or Type y Condorcet preference (with both types assigned to at least one individual); and (3) all remaining individuals get the Type z Condorcet preference.

Notice that only Type z individuals rank z above y in the Condorcet cycle (see Figure 3). Since $y \succ_i^{**} z$ for all $i \in S$, Step 2 implies $z \notin f(\{y, z\}, \succsim^{**})$. By ILA and Claim 2.2, $f(X, \succsim^{**}) \subseteq \{x, y\}$. Also, by the Remark on Step 2, $f(X, \succsim^{**})$ must be a singleton.

So let $S_2 \subset S$ be the set of individuals whose favorite at $\succsim^{**}$ is chosen out of X. Repeat Step 2 applied to S_2. Proceeding this way gives us a strictly decreasing sequence of subsets of individuals $S_n \subset \cdots \subset S_2 \subset S$ such that each S_n group gets their favorites out of Y_{k-1} whenever they have the same preference over Y. Since N is finite, S_n must be a singleton at some finite n. Setting $\pi_k = S_n$ completes our induction proof.

Step 4. Step 1 is infeasible when there are two or fewer individuals not assigned to the permutation π (this happens when $N < 3$ or $k > N - 2$). This step takes care of such cases.

When there are only two individuals left, construct the Condorcet profile $\succsim^*$ as in Step 1 without using the Type z preference. Lemma 3.2 requires $f(X, \succsim^*) \subseteq \{x, y\}$. Proceed with the same argument as above.

When there is only one individual left, construct π by appending the last individual to π^{N-1}. Given $Y \subseteq X$ and $\succsim \in R^N$, construct $\succsim^Y$ by taking Y to the top from $\succsim$. The induction hypothesis ensures $f(X, \succsim^Y) \subseteq Y_{N-1}$. Now if y is a favorite for π_N in Y_{N-1} and y' is not, y weakly Pareto dominates y' (since $\pi_1, \ldots, \pi_{N-1}$ are indifferent between all alternatives in Y_{N-1}). By Claim 2.2, $f(X, \succsim^Y) \subseteq T(Y_{N-1}, \succsim^Y_{\pi_N})$. Applying ILA once more we get $f(Y, \succsim^Y) \subseteq T(Y_{N-1}, \succsim^Y_{\pi_N})$. Since $f(Y, \succsim) = f(Y, \succsim^Y)$ (IIF) and $T(Y_{N-1}, \succsim^Y_{\pi_N}) = T(Y_{N-1}, \succsim_{\pi_N})$, $f(Y, \succsim) \subseteq T(Y_{N-1}, \succsim_{\pi_N})$.

Remark. The argument in the $k = 1$ step is unaffected if the preference domain is restricted to P^N. Since the existence of π_1 implies dictatorship, we have the following theorem:

Theorem 3.3. *Any social choice correspondence $f : (2^X \setminus \emptyset) \times P^N \to (2^X \setminus \emptyset)$ that is Strongly Unanimous, Independent of Infeasible Alternatives and Independent of Losing Alternatives is Dictatorial.*

3.2. **Tie-Breaking Preference.** It remains to find the tie-breaking preference $\rho \in R$ such that

$$f(Y, \succsim) = T(T^N(Y, \succsim, \pi), \rho) \quad \text{for all } Y \subseteq X, \text{ all } \succsim \in R^N.$$

So let $\sim$ denote the preference profile in which all individuals are indifferent between all alternatives. Given a social choice correspondence f, define a binary relation ρ on X such that for all $x, y \in X$,

$$x \,\rho\, y \quad \text{if and only if} \quad x \in f(\{x, y\}, \sim).$$

Claim 3.4. The binary relation ρ is complete and transitive. That is, $\rho \in R$.

Proof. Completeness: For any $x, y \in X$, $f(\{x, y\}, \sim) \neq \emptyset$.[11] Thus either $x \,\rho\, y$ or $y \,\rho\, x$.

Transitivity: Take $x, y, z \in X$ and suppose $x \,\rho\, y$ and $y \,\rho\, z$. Construct $\succsim^\circ$ by taking $\{x, y, z\}$ to the top from $\sim$. IIF implies x and y are chosen out of $\{x, y\}$ and $\{y, z\}$ at $\succsim^\circ$ respectively. We claim $x \in f(X, \succsim^\circ)$. Suppose not, then since $x \in f(\{x, y\}, \succsim^\circ)$, ILA requires $y \notin f(X, \succsim^\circ)$. Applying ILA once more implies $z \notin f(X, \succsim^\circ)$. But this contradicts Claim 2.2 since all alternatives other than x, y and z are strictly Pareto dominated at $\succsim^\circ$. Now by ILA and IIF we have $x \in f(\{x, z\}, \succsim^\circ) = f(\{x, z\}, \sim)$. Therefore $x \,\rho\, z$. ☐

Fix $Y \subseteq X$ and $\succsim \in R^N$. Construct $\succsim^Y$ by taking Y to the top from $\succsim$. Our argument in Section 3.1 ensures $f(X, \succsim^Y) \subseteq Y_N = T^N(Y, \succsim, \pi)$. Notice that $\succsim^Y$ and $\sim$ agree on Y_N since all individuals are indifferent between all alternatives in Y_N.

The proof of $f(Y, \succsim^Y) \subseteq T(Y_N, \rho)$ is essentially the same as the proof in Step 2.3 in Section 3.1 and will therefore be omitted. We show that $f(Y, \succsim^Y) \supseteq T(Y_N, \rho)$. Let $y \in T(Y_N, \rho)$ and $y' \in f(X, \succsim^Y) \subseteq Y_N$. By the definition of ρ, $y \in f(\{y, y'\}, \sim)$. Since both y and y' are in Y_N, $\sim$ and $\succsim^Y$ agree on $\{y, y'\}$. IIF requires $y \in f(\{y, y'\}, \succsim^Y) = f(\{y, y'\}, \sim)$. Meanwhile, as $y' \in f(X, \succsim^Y) \cap \{y, y'\}$, ILA requires $y \in f(X, \succsim^Y)$. Applying ILA once more gives $y \in f(Y, \succsim^Y)$. Hence $T(Y_N, \rho) \subseteq f(Y, \succsim^Y)$.

Combining the results we get $f(Y, \succsim^Y) = T(Y_N, \rho)$. By IIF, $f(Y, \succsim) = f(Y, \succsim^Y)$. Hence $f(Y, \succsim) = T(Y_N, \rho) = T(T^N(Y, \succsim), \rho)$. This completes the proof of Theorem 3.1.

[11] $f(\{x, x\}, \sim) = f(\{x\}, \sim) = \{x\}$ for all $x \in X$.

4. Arrow's Impossibility Theorem

In Arrow (1963), a *social welfare function* is a mapping $F : R^N \to R$. Given a preference profile $\succsim \in R^N$, a permutation of individuals $\pi = (\pi_1, \ldots, \pi_N)$ and a tie-breaking preference $\rho \in R$, write $\succsim_{\pi_{N+1}} = \rho$ and define the lexicographic ordering $L(\succsim, \pi, \rho) \in R$ such that $x\, L(\succsim, \pi, \rho)\, y$ if and only if whenever $y \succ_{\pi_k} x$, there exists an $l < k$ such that $x \succ_{\pi_l} y$.

Definitions 4.1. A social welfare function $F : R^N \to R$ is

Strongly Pareto if x is strictly preferred to y according to $F(\succsim)$ whenever x weakly Pareto dominates y under $\succsim$;

Independent of Irrelevant Alternatives if $F(\succsim)$ and $F(\succsim')$ agree on $\{x, y\}$ whenever $\succsim$ and $\succsim'$ agree on the same set;

Serially Dictatorial if there exists a permutation of individuals π and a tie-breaking preference $\rho \in R$ such that $F(\succsim) = L(\succsim, \pi, \rho)$ for all $\succsim \in R^N$.

We now translate Arrow's setting into ours. Given a social welfare function F, define an induced social choice correspondence as follows: For all subsets of alternatives $Y \subseteq X$ and all preference profiles $\succsim \in R^N$,

$$f(Y, \succsim) = T(Y, F(\succsim)). \tag{2}$$

The following proposition relates properties of the social welfare function and those of the induced social choice correspondence.

Proposition 4.2. *If the social welfare function F is Strongly Pareto and Independent of Irrelevant Alternatives, then the social choice correspondence f defined in Equation* (2) *satisfies SU, IIF and ILA.*

Proof. SU: Let y be uniquely weakly Pareto dominant in $Y \subseteq X$ at $\succsim \in R^N$. Since F is strongly Pareto, y is strictly preferred to all other $y' \in Y$ according to $F(\succsim)$. Hence $T(Y, F(\succsim)) = \{y\}$. By Equation (2), $f(Y, \succsim) = \{y\}$.

IIF: Suppose $\succsim$ and $\succsim'$ agree on $Y \subseteq X$. Independence of Irrelevant Alternatives requires $F(\succsim)$ and $F(\succsim')$ to agree on all pairs $y, y' \in Y$. Hence $T(Y, F(\succsim)) = T(Y, F(\succsim'))$. By Equation (2), $f(Y, \succsim) = f(Y, \succsim')$.

ILA: Suppose $f(X, \succsim) \cap Y \neq \emptyset$. If $x \in f(X, \succsim) \cap Y$ and $y \in Y$ is not, then x is strictly preferred to y according to $F(\succsim)$. Thus $x \in T(Y, F(\succsim))$ and y is not. By Equation (2), $f(Y, \succsim) = f(X, \succsim) \cap Y$. □

By Theorem 3.1, f is serially dictatorial. Thus there exists a permutation of individuals π and a tie-breaking preference ρ such that $f(\{x, y\}, \succsim) = T(T^N(\{x, y\}, \succsim, \pi), \rho)$ for all pairs of $x, y \in X$. By Equation (2), $x \in f(\{x, y\}, \succsim)$ is equivalent to $x\, F(\succsim)\, y$. Meanwhile,

$x \in T(T^N(\{x,y\}, \succsim, \pi), \rho)$ is equivalent to $x \; L(\succsim, \pi, \rho) \; y$. Arrow's Impossibility Theorem[12] follows immediately:

Corollary 4.3 (Arrow's Impossibility Theorem)**.** *Any social welfare function $F : R^N \to R$ that is Strongly Pareto and Independent of Irrelevant Alternatives is Serially Dictatorial.*

5. Social Choice and Implementation

To avoid confusion with our social choice correspondence on the unrestricted domain, we call a social choice correspondence defined only on the set of all social alternatives an *overall social choice correspondence,* which is a mapping $f^* : R^N \to 2^X \setminus \emptyset$.

Definitions 5.1. An overall social choice correspondence $f^* : R^N \to 2^X \setminus \emptyset$ is

Strongly Pareto if $f^*(\succsim) = \{x\}$ whenever x is uniquely weakly Pareto dominant at $\succsim$;

Monotonic if $Y \subseteq f^*(\succsim') \subseteq f^*(\succsim)$ whenever $Y \subseteq f^*(\succsim)$ and $y \succsim_i x$ implies $y \succsim'_i x$ (with $y \succ_i x$ implies $y \succ'_i x$) for all $y \in Y$, all $x \in X$ and all $i \in N$;

Dictatorial if there exists an individual $i \in N$ such that $f^*(\succsim) \subseteq T(X, \succsim_i)$ for all $\succsim \in R^N$;

Serially Dictatorial if there exists a permutation of individuals π and a tie-breaking preference $\rho \in R$ such that $f^*(\succsim) = T(T^N(X, \succsim, \pi), \rho)$ for all $\succsim \in R^N$.

In this section, we first show that the Muller-Satterthwaite Theorem (Muller and Satterthwaite, 1977) can be derived from Theorem 3.1. Next we turn to implementation and discuss the Gibbard-Satterthwaite Theorem (Gibbard, 1973; Satterthwaite, 1975) and the Jackson-Srivastava Characterization (Jackson and Srivastava, 1996).

5.1. **The Muller-Satterthwatie Theorem.** Given an overall social choice correspondence $f^* : R^N \to 2^X \setminus \emptyset$, extend it to our unrestricted domain by defining: for all $Y \subseteq X$ and all $\succsim \in R^N$

$$f(Y, \succsim) = f^*(\succsim^Y), \tag{3}$$

where $\succsim^Y$ is a preference profile taking Y to the top from $\succsim$.[13] If f^* is a function, so is f.

We need to show that f defined by Equation (3) is a valid social choice correspondence, that is, $f(Y, \succsim) \subseteq Y$ for all Y and all $\succsim$. This is accomplished by the following lemma.

Lemma 5.2. *If an overall social choice correspondence f^* is Strongly Pareto and Monotonic, then $x \notin f^*(\succsim)$ whenever x is weakly Pareto dominated at $\succsim$.*

[12] This is a serial dictatorship version. See Luce and Raiffa (1957, Section 14.5) and Gevers (1979, Theorem 3).

[13] There are multiple $\succsim^Y$ that take Y to the top from $\succsim$. One can pick any one of them for each pair of $\succsim$ and Y to define an extension. All our results are unaffected by the choice of $\succsim^Y$, hence the choice of the particular extension used. The same remark applies to Equation (5) in Section 6.

Proof. Suppose by contradiction that y weakly Pareto dominates x at $\succsim$ but $x \in f^*(\succsim)$. Construct $\succsim'$ by taking $\{x, y\}$ to the top from $\succsim$. Monotonicity requires $x \in f^*(\succsim')$ but Strong Pareto optimality requires $f^*(\succsim') = \{y\}$. □

The next proposition relates axioms on f^* with those on f.

Proposition 5.3. *If an overall social choice correspondence f^* is Strongly Pareto and Monotonic, then the social choice correspondence f defined in Equation* (3) *is SU, IIF and ILA.*

Proof. SU: Follows from Lemma 5.2.

IIF: Suppose $\succsim$ and $\succsim'$ agree on $Y \subseteq X$. Then $\succsim^Y$ and $\succsim'^Y$ differ only by the ranking among alternatives not in Y, which are all ranked below $f^*(\succsim^Y) \subseteq Y$ in both $\succsim^Y$ and $\succsim'^Y$. Monotonicity requires $f^*(\succsim^Y) = f^*(\succsim'^Y)$. By Equation (3), $f(Y, \succsim) = f(Y, \succsim')$.

ILA: Suppose $Z = f^*(\succsim) \cap Y = f(X, \succsim) \cap Y$ is non-empty. Notice that all individuals (strictly) prefer each $z \in Z$ to each $x \in X$ at $\succsim^Y$ if they (strictly) prefer z to x at $\succsim$. By Monotonicity $Z \subseteq f^*(\succsim^Y) \subseteq f^*(\succsim)$. Meanwhile, Lemma 5.2 dictates $f^*(\succsim^Y) \subseteq Y$. Since $f^*(\succsim^Y) \subseteq f^*(\succsim)$ (above, by Monotonicity), $f^*(\succsim^Y) \subseteq f^*(\succsim) \cap Y = Z$. Hence $f^*(\succsim^Y) = Z$. By Equation (3) and the definition of Z, $f(Y, \succsim) = f(X, \succsim) \cap Y$. □

None of the arguments in Lemma 5.2 and Proposition 5.3 is affected if the preference domain is restricted to P^N. Hence two generalized versions of the Muller-Satterthwaite Theorem follow immediately from Theorems 3.1 and 3.3:

Corollary 5.4 (Muller-Satterthwaite Theorem)**.** *Any overall social choice correspondence $f^* : R^N \to 2^X \setminus \emptyset$ (respectively, $f^* : P^N \to 2^X \setminus \emptyset$) that is Strongly Pareto and Monotonic is Serially Dictatorial (Dictatorial).*

5.2. **The Gibbard-Satterthwaite Theorem.** We restrict our attention to overall social choice *functions* for the Gibbard-Satterthwaite Theorem since introducing strategy-proofness for set-valued mappings requires an extension of preferences on X to preferences on 2^X, which is too far a digression[14]. We also restrict the preference domain to P^N, that is, only strict preferences are admitted. The discussion on the latter restriction is deferred to the end of this subsection.

Definitions 5.5. An overall social choice function $f^* : P^N \to X$ is
onto if for every $x \in X$ there is a $\succ \in P^N$ such that $f^*(\succ) = x$;
Strategy-proof if $f^*(\succ) \succsim_i f^*(\succ'_i, \succ_{-i})$ for all $\succ$, all i and all $\succ'_i$.

Define also Strong Pareto optimality and Monotonicity for overall social choice functions defined on P^N in the same manner as in Definitions 5.1.

[14]See Gärdenfors (1979) for a survey on early work on this topic. Examples on different approaches to this problem include Duggan and Schwartz (2000); Barberà, Dutta, and Sen (2001); Ching and Zhou (2002).

1	2	3
$x\,y\,z$	$y\,z$	z
	x	y
		x

(a) Profile $\succsim$

1	2	3
$x\,y\,z$	x	z
	y	y
	z	x

(b) Profile $\succsim'$

1	2	3
$x\,y$	x	z
z	y	y
	z	x

(c) Profile $\succsim''$

FIGURE 6. Preference Profiles for Example 5.8

Lemma 5.6.

(1) *Any Strategy-proof overall social choice function $f^* : P^N \to X$ is Monotonic.*
(2) *Any onto and Monotonic overall social choice function $f^* : P^N \to X$ is Strongly Pareto.*

Proof. See Muller and Satterthwaite (1977); Reny (2001, pp. 104-105). □

The Gibbard-Satterthwaite Theorem now follows from Corollary 5.4:

Corollary 5.7 (Gibbard-Satterthwaite Theorem)**.** *Any onto*[15] *and Strategy-proof overall social choice function $f^* : P^N \to X$ is dictatorial.*

The restriction to strict preferences in this subsection is not innocuous, as illustrated by the next example[16]:

Example 5.8. There are 3 alternatives, $X = \{x, y, z\}$, and 3 individuals, $N = \{1, 2, 3\}$. The overall social choice function $f^* : R^N \to X$ takes the following form: Individual 1 is the first dictator. If Individual 1 has multiple favorites and z is one of them, then Individual 2 will be the second dictator and any remaining tie will be broken by alphabetical order. If instead z is not among Individual 1's favorite, Individual 3 will be the second dictator and any remaining tie will be broken by alphabetical order.

It can be easily verified that f^* is onto and Strategy-proof. However, it is not Strongly Pareto: At profile $\succsim$ (Figure 6(a)), z is the uniquely weakly Pareto dominant alternative, yet $f^*(\succsim) = y$. Moreover, if $f : (2^X \setminus \emptyset) \times R^N \to X$ extends f^*, f violates either IIF or ILA. If it satisfies ILA, then $f(\{x, y\}, \succsim') = x$ and $f(\{x, y\}, \succsim'') = y$, which violates IIF since $\succsim$ and $\succsim'$ agree on $\{x, y\}$.[17]

[15] One can replace the onto assumption with the assumption that the range of f^* contains at least 3 elements by redefining X to be the range of f^* (see Barberà and Peleg, 1990). The same remark applies to the strict preference version of Corollary 5.11 in the Section 5.3.
[16] We thank Salvador Barberà for this example.
[17] For similar reasons f^* also violates Monotonicity.

5.3. **The Jackson-Srivastava Characterization.** Again we restrict our attention to overall social choice functions as it is unclear what it means for a mechanism to implement a set-valued social choice mapping.

A *mechanism* $M = (A, g)$ consists of an action profile space $A = \prod_{i \in N} A_i$ and an outcome function $g : A \to X$. Let $\mathcal{M}$ be the set of all mechanisms. An *equilibrium concept* is a mapping $E : \mathcal{M} \times R^N \to 2^A$. The *equilibrium outcome correspondence* associated with E is given by

$$O_E(M, \succsim) = \{x \in X : \exists\, a \in E(M, \succsim) \text{ s.t. } g(a) = x\}.$$

An overall social choice function $f^* : R^N \to X$ is *implemented* via equilibrium concept E and mechanism M if $O_E(M, \succsim) = f^*(\succsim)$ for all $\succsim \in R^N$.

Definitions 5.9 (Jackson-Srivastava)**.** Let $M = (A, g)$ be a mechanism. Take $a \in A$, $\succsim \in R^N$ and two groups of individuals $S, S' \subseteq N$. The action profile $a' = (a'_S, a_{-S})$ is an (S, S')-*improvement from a at $\succsim$* if $g(a') \succsim_i g(a)$ for all $i \in S'$ with at least one strict preference. A pair of groups (S, S') is *responsive* with respect to mechanism M under equilibrium concept E if $a \notin E(M, \succsim)$ whenever there exists an (S, S')-improvement from a at $\succsim$.

An equilibrium concept E satisfies *direct breaking with respect to M* if, whenever $O_E\left(M, (\succsim'_i, \succsim_{-i})\right) \neq O_E(M, \succsim)$, for each $a \in E(M, \succsim)$ there exists a responsive pair of groups (S, S') under E and an (S, S')-improvement from a at $(\succsim'_i, \succsim_{-i})$.

When the preference domain is P^N, iterative elimination of strictly dominated strategies, Nash equilibrium and Strong equilibrium satisfy direct breaking with respect to all mechanisms, while undominated strategies satisfies direct breaking with respect to all bounded mechanisms[18] (Jackson and Srivastava, 1996). The proofs for the first three concepts (iterative elimination of strictly dominated strategies, Nash equilibrium and Strong equilibrium) extend easily when weak preferences are allowed.

The next lemma adapts Jackson and Srivastava's result on Monotonicity and implementation to the weak preference domain.

Lemma 5.10. *Suppose a mechanism M implements an overall social choice function $f^* : R^N \to X$ via equilibrium concept E. If E satisfies direct breaking with respect to M, then f^* is Monotonic.*

Proof. Let $f^*(\succsim) = x$ and consider $\succsim'$ such that $x \succsim_i y$ implies $x \succsim'_i y$ (with $x \succ_i y$ implying $x \succ'_i y$) for all $y \in X$ and all $i \in N$. Since M implies f^* via E, $O_E(M, \succsim) = x$. We claim that $O_E(M, (\succsim'_i, \succsim_{-i})) = x$ for any $i \in N$. Suppose not and let $a \in E(M, \succsim)$. Since E satisfies direct breaking, there exists a responsive pair of groups (S, S') with respect to

[18]A mechanism is bounded if, at each preference profile, each dominated action is dominated by an undominated action.

M under E and an (S, S')-improvement $a' = (a'_S, a_{-S})$ from a at $(\succsim'_i, \succsim_{-i})$. By definition, $g(a') \succsim_j g(a) = x$ for all $j \in S'$ with at least one strict preference. If $i \notin S'$, a' is an (S, S')-improvement from a at $\succsim$. If $i \in S'$, since $g(a') \succsim'_i x$ implies $g(a') \succsim_i x$ (with strict preference implying strict preference)[19], a' is also an (S, S')-improvement from a at $\succsim$. This contradicts (S, S') being responsive. Repeat the same argument with another individual $j \neq i$ starting at $(\succsim'_i, \succsim_{-i})$. Proceeding this way we reach $\succsim'$ and the social choice remains x. □

The following two versions of the Jackson-Srivastava characterization on equilibrium concepts that lead to impossibility theorems[20] are now straight-forward:

Corollary 5.11 (Jackson and Srivastava (1996))**.** *Suppose a mechanism M implements a Strongly Pareto (respectively, onto) social choice function $f^* : R^N \to X$ ($f^* : P^N \to X$) via equilibrium concept E. Then f^* is serially dictatorial (dictatorial) if and only if E satisfies direct breaking with respect to M.*

Proof. If: Follows from Lemma 5.10 (and Lemma 5.6 in the case of strict preferences) and Corollary 5.4.

Only if: Let $\pi = (1, \ldots, N)$ be the sequence of serial dictators and ρ be the tie-breaking preference. Add a dummy individual $N+1$ whose only preference is ρ. Suppose $O_E(M, \succsim) = x$ and $O_E(M, (\succsim'_i, \succsim_{-i})) = y \neq x$. Let $k \geq i$ be the individual who strictly prefers y to x at the new preference profile. Now any $a' \in E(M, (\succsim'_i, \succsim_{-i}))$ is an (N, π^k)-improvement from any $a \in E(M, \succsim)$ at $(\succsim'_i, \succsim_{-i})$. Serial dictatorship implies that (N, π^k) is responsive with respect to M under E. Therefore E satisfies direct breaking with respect to M. □

The strict preference version is the original theorem in Jackson and Srivastava (1996). The weak preference version is a new extension of their theorem. It is essential to state this extension using Strong Pareto optimality instead of onto, as onto and Monotonicity guarantee only weak but not strong Pareto optimality when the preference domain is R^N.

Due to the remarks after Definition 5.9, Corollary 5.11 implies that any Strongly Pareto overall social choice function that is implementable via iterative elimination of strictly dominated strategies, Nash equilibrium or Strong equilibrium is Serially Dictatorial. When only strict preferences are admitted, any onto overall social choice function that is implementable via undominated strategies in a bounded mechanism is Dictatorial.

6. Strategic Candidacy

Strategic candidacy concerns the effect of a unilateral withdrawal of candidacy on the election outcome. Hence the social choice is defined on subsets of social alternatives with

[19]If $g(a') \succ_i x$, then $x \not\succsim'_i g(a')$. By the hypothesis on $\succsim'$, $x \not\succsim_i g(a')$, which is equivalent to $g(a') \succ_i x$. If $g(a') \sim_i x$, then $x \not\succ'_i g(a')$. By the hypothesis on $\succsim'$, $x \not\succ_i g(a')$, which is equivalent to $g(a') \succsim_i x$.
[20]We thank Matthew Jackson for pointing us to this theorem.

at least $|X|-1$ elements. More generally, one can consider social choices defined on some $\mathcal{X} \subseteq 2^X$. Say $\mathcal{X}$ satisfies *k-set feasibility* if for all subsets $Y \subseteq X$, $Y \in \mathcal{X}$ whenever $|Y| \geq k$. A *voting procedure* is a correspondence $\hat{f} : \mathcal{X} \times R^N \to 2^X \setminus \emptyset$ such that $\hat{f}(Y, \succsim) \subseteq Y$ for all $Y \in \mathcal{X}$ and all $\succsim \in R^N$.[21]

Definitions 6.1. A voting procedure $\hat{f} : \mathcal{X} \times R^N \to 2^X \setminus \emptyset$ satisfies

k-set Feasibility if $\mathcal{X}$ is k-set feasible;

Strong Unanimity if $\hat{f}(Y, \succsim) = \{y\}$ whenever y is uniquely weakly Pareto dominant in $Y \in \mathcal{X}$;

Independence of Irrelevant Alternatives (IIA) if $\hat{f}(Y, \succsim) = \hat{f}(Y, \succsim')$ whenever $\succsim$ and $\succsim'$ agree on $Y \in \mathcal{X}$;

Arrow's Choice Axiom (ACA) if $\hat{f}(Y, \succsim) = \hat{f}(Z, \succsim) \cap Y$ whenever $Y, Z \in \mathcal{X}$, $Y \subseteq Z$ and the intersection is non-empty;

Strong Candidate Stability (SCS) if $\hat{f}(Y, \succsim) = \hat{f}(X, \succsim) \cap Y$ whenever $Y \in \mathcal{X}$, $|Y| = |X|-1$ and the intersection is non-empty;

Dictatorship if there exists an individual i such that $\hat{f}(Y, \succsim) \subseteq T(Y, \succsim_i)$ for all $Y \in \mathcal{X}$ and all $\succsim \in R^N$;

Serial Dictatorship if there exists a permutation of individuals π and a tie-breaking preference $\rho \in R$ such that $\hat{f}(Y, \succsim) = T(T^N(Y, \succsim, \pi), \rho)$ for all $Y \in \mathcal{X}$ and all $\succsim \in R^N$.

Obviously, if $\hat{f}$ is k-set feasible for some $k < |X|$ and satisfies Arrow's choice axiom, $\hat{f}$ is strongly candidate stable.

Lemma 6.2. *If the voting procedure $\hat{f}$ is $(|X|-1)$-set feasible, strongly unanimous, IIA and SCS, then $f(X, \succsim) \subseteq Y$ whenever each alternative $y \in Y$ weakly Pareto dominates each $x \notin Y$ at $\succsim$.*

Proof. See Eraslan and McLennan (2004, Lemma 1, pp. 41-42). □

Lemma 6.3. *Let $\hat{f}$ be a $(|X|-1)$-set feasible, strongly unanimous, IIA and SCS voting procedure. Then $\hat{f}(X, \succsim^Y) = \hat{f}(X, \succsim) \cap Y$ whenever $\succsim^Y$ takes Y to the top from $\succsim$ and the intersection is non-empty.*

Proof. Fix Y, $\succsim$ and $\succsim^Y$ such that $\hat{f}(X, \succsim) \cap Y \neq \emptyset$. Let $Z = X \setminus Y$. Enumerate the elements of Z as $z_1, \dots, z_K$. Define $Z_0 = \emptyset$ and $Z_k = \{z_1, \dots, z_k\}$. Construct $\succsim^0 = \succsim$ and for all $k > 0$ a preference profile $\succsim^k$ such that: (1) $\succsim^k$ and $\succsim^{k-1}$ agree on $X \setminus \{z_k\}$; (2) $\succsim^k$ and $\succsim^Y$ agree on Z_k; and (3) all $z \in Z_k$ are strictly Pareto dominated by all $x \notin Z_k$. By construction, $\succsim^K = \succsim^Y$.

[21]We do not allow for candidate voters (social alternatives which are also individuals).

We claim that for all $k \geq 0$,

$$\hat{f}(X, \succsim^k) = \hat{f}(X, \succsim) \setminus Z_k. \tag{4}$$

The case of $k = 0$ is trivial. Suppose Equation (4) holds for $k-1$. Since, by assumption, $\hat{f}(X, \succsim) \cap Y \neq \emptyset$, the induction hypothesis implies $\hat{f}(X, \succsim^{k-1}) \neq \{z_k\}$. Thus

$$\begin{aligned}
\hat{f}(X, \succsim^k) &= \hat{f}(X, \succsim^k) \cap (X \setminus \{z_k\}) && \text{(Lemma 6.2)} \\
&= \hat{f}(X \setminus \{z_k\}, \succsim^k) && \text{(SCS)} \\
&= \hat{f}(X \setminus \{z_k\}, \succsim^{k-1}) && \text{(IIA)} \\
&= \hat{f}(X, \succsim^{k-1}) \cap (X \setminus \{z_k\}) && \text{(SCS)} \\
&= \left(\hat{f}(X, \succsim) \setminus Z_{k-1}\right) \setminus \{z_k\} && \text{(Induction hypothesis)} \\
&= \hat{f}(X, \succsim) \setminus Z_k.
\end{aligned}$$

Therefore $\hat{f}(X, \succsim^Y) = \hat{f}(X, \succsim) \setminus Z = \hat{f}(X, \succsim) \cap Y$. □

Given a voting procedure $\hat{f}$ that satisfies k-set feasibility for some $k < |X|$, Strong Unanimity, IIA and ACA, extend it to our unrestricted domain by defining: for all $Y \subseteq X$ and all $\succsim \in R^N$,

$$f(Y, \succsim) = \hat{f}(X, \succsim^Y), \tag{5}$$

where $\succsim^Y$ is a preference profile that takes Y to the top from $\succsim$. Lemma 6.2 guarantees $f(Y, \succsim) \subseteq Y$ so f is a valid social choice correspondence. Lemma 6.3 and ACA ensure $\hat{f}(X, \succsim^Y) = \hat{f}(X, \succsim) \cap Y = \hat{f}(Y, \succsim)$ for all $Y \in \mathcal{X}$, so f is indeed an extension of $\hat{f}$.

Proposition 6.4. *If a voting procedure $\hat{f}$ is $(|X|-1)$-set feasible, strongly unanimous, IIA and SCS, then the social choice correspondence f defined in Equation (5) satisfies SU, IIF and ILA.*

Proof. SU: If y is uniquely weakly Pareto dominant in $Y \subseteq X$ at $\succsim$, then y is uniquely weakly Pareto dominant in X at any $\succsim^Y$ taking Y to the top from $\succsim$. Strong Unanimity of $\hat{f}$ and Equation (5) ensure $f(Y, \succsim) = \hat{f}(X, \succsim^Y) = \{y\}$.

IIF: If $\succsim$ and $\succsim'$ agree on $Y \subseteq X$, then $\succsim^Y$ takes Y to the top from $\succsim'^Y$ as well. Lemma 6.2 guarantees that both $\hat{f}(X, \succsim^Y)$ and $\hat{f}(X, \succsim'^Y)$ are subsets of Y. Applying Lemma 6.3 we obtain $\hat{f}(X, \succsim^Y) = \hat{f}(X, \succsim'^Y) \cap Y = \hat{f}(X, \succsim'^Y)$. By Equation (5), $f(Y, \succsim) = f(Y, \succsim')$.

ILA: Follows from Lemma 6.3. □

Since the above argument is unaffected by restricting the preference domain to P^N, two versions of the Grether-Plott Theorem[22] are immediate from Theorems 3.1 and 3.3:

[22] We thank John Weymark for pointing us to this theorem.

Corollary 6.5 (Grether and Plott (1982)). *If a voting procedure* $\hat{f} : \mathcal{X} \times R^N \to 2^X \setminus \emptyset$ *(respectively,* $\hat{f} : \mathcal{X} \times P^N \to 2^X \setminus \emptyset$*) satisfies* k*-set feasibility for some* $k < |X|$*, Strong Unanimity, Independence of Irrelevant Alternatives and Arrow's Choice Axiom, it is Serially Dictatorial (Dictatorial).*

The impossibility theorem under strategic candidacy follows as a special case:

Corollary 6.6 (Dutta, Jackson, and Le Breton (2001)[23]). *If a voting procedure* $\hat{f} : \mathcal{X} \times R^N \to 2^X \setminus \emptyset$ *(respectively,* $\hat{f} : \mathcal{X} \times P^N \to 2^X \setminus \emptyset$*) is* $(|X| - 1)$*-set feasible, Strongly Unanimous, Independent of Irrelevant Alternatives and Strongly Candidate Stable, it is Serially Dictatorial (Dictatorial).*

7. Conclusion

This paper proposes a unifying impossibility theorem. Unanimity and our two independence conditions underlie the axioms of a number of classical impossibility theorems. Thus even if one finds the axioms of these impossibility theorems disputable, our theorem indicates that any alternative set of axioms that implies ours leads also to dictatorship.

Several extensions are possible. For instance, one can modify our definitions to prove an impossibility theorem under the set of all continuous preferences over a compact metric space of social alternatives (c.f.: Barberà and Peleg, 1990). Another possibility is to allow randomized social choices (c.f.: Benoît, 2002). Finally, one may use the ultrafilter method of Kirman and Sondermann (1972) to obtain dictatorship when there are infinitely many individuals.

References

Arrow, K. J. (1959): "Rational Choice Functions and Orderings," *Economica*, 26(102), 121–127.

——— (1963): *Social Choice and Individual Values*. Yale Univeristy Press, New Haven, 2nd edn.

Barberà, S. (1980): "Pivotal Voters: A New Proof of Arrow's Theorem," *Economics Letters*, 6(1), 13–16.

——— (1983): "Strategy-proofness and Pivotal Voters: A Direct Proof of the Gibbard-Satterthwaite Theorem," *International Economic Review*, 24(2), 413–417.

——— (2003): "A Theorem on Preference Aggregation," Working Paper no. 166, Barcelona Economics Working Paper Series.

Barberà, S., B. Dutta, and A. Sen (2001): "Strategy-proof Social Choice Correspondences," *Journal of Economic Theory*, 101(2), 374–394.

[23] See also Ehlers and Weymark (2003); Eraslan and McLennan (2004); Rodríguez-Álvarez (2006).

Barberà, S., and B. Peleg (1990): "Strategy-proof Voting Schemes with Continuous Preferences," *Social Choice and Welfare*, 7(1), 31–38.

Benoît, J.-P. (2000): "The Gibbard-Satterthwaite Theorem: A Simple Proof," *Economics Letters*, 69(3), 319–322.

——— (2002): "Strategic Manipulation in Voting Games when Lotteries and Ties are Permitted," *Journal of Economic Theory*, 102(2), 421–436.

Campbell, D. E. (1979): "Manipulation of Social Choice Rules by Strategic Nomination of Candidates," *Theory and Decision*, 10(1-4), 247–263.

Campbell, D. E., and J. S. Kelly (2002): "Impossibility Theorems in the Arrovian Framework," in *Handbook of Social Choice and Welfare*, ed. by K. J. Arrow, A. K. Sen, and K. Suzumura, vol. 1, chap. 1, pp. 35–94. North-Holland, Amsterdam.

Ching, S., and L. Zhou (2002): "Multi-valued Strategy-proof Social Choice Rules," *Social Choice and Welfare*, 19(3), 569–580.

Denicolò, V. (2000): "Independence of Irrelevant Alternatives and Consistency of Choice," *Economic Theory*, 15(1), 221–226.

Duggan, J., and T. Schwartz (2000): "Strategic Manipulability without Resoluteness or Shared Beliefs: Gibbard-Satterthwaite Generalized," *Social Choice and Welfare*, 17(1), 85–93.

Dutta, B., M. O. Jackson, and M. Le Breton (2001): "Strategic Candidacy and Voting Procedures," *Econometrica*, 69(4), 1013–1037.

Ehlers, L., and J. A. Weymark (2003): "Candidate Stability and Nonbinary Social Choice," *Economic Theory*, 22(2), 233–243.

Eliaz, K. (2004): "Social Aggregators," *Social Choice and Welfare*, 22(2), 317–330.

Eraslan, H., and A. McLennan (2004): "Strategic Candidacy for Multivalued Voting Procedures," *Journal of Economic Theory*, 117(1), 29–54.

Gärdenfors, P. (1979): "On Definitions of Manipulation of Social Choice Functions," in *Aggregation and Revelation of Preferences*, ed. by J.-J. Laffont. North-Holland, Amsterdam.

Geanakoplos, J. (2005): "Three Brief Proofs of Arrow's Impossibility Theorem," *Economic Theory*, 26(1), 211–215.

Gevers, L. (1979): "On Interpresonal Comparability and Social Welfare Orderings," *Econometrica*, 47(1), 75–89.

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

Grether, D. M., and C. R. Plott (1982): "Nonbinary Social Choice: An Impossibility Theorem," *Review of Economic Studies*, 49(1), 143–149.

Hansson, B. (1968): "Choice Structures and Preference Relations," *Synthese*, 18(4), 443–458.

——— (1969): "Voting and Group Decision Functions," *Synthese*, 20(4), 526–537.

Jackson, M. O., and S. Srivastava (1996): "A Characterization of Game-Theoretic Solutions which Lead to Impossibility Theorems," *Review of Economic Studies*, 63(1), 23–38.

Karni, E., and D. Schmeidler (1976): "Independence of Nonfeasible Alternatives, and Independence of Nonoptimal Alternatives," *Journal of Economic Theory*, 12(3), 488–493.

Kirman, A. P., and D. Sondermann (1972): "Arrow's Theorem, Many Agents and Invisible Dictators," *Journal of Economic Theory*, 5(2), 267–277.

Le Breton, M., and J. A. Weymark (2011): "Arrovian Social Choice Theory on Economic Domains," in *Handbook of Social Choice and Welfare*, ed. by K. J. Arrow, A. K. Sen, and K. Suzumura, vol. 2, chap. 17, pp. 191–299. North-Holland, Amsterdam.

Luce, R. D., and H. Raiffa (1957): *Games and Decisions*. Wiley, New York.

Muller, E., and M. A. Satterthwaite (1977): "The Equivalence of Strong Positive Association and Strategy-proofness," *Journal of Economic Theory*, 14(2), 412–418.

Reny, P. J. (2001): "Arrow's Theorem and the Gibbard-Satterthwaite Theorem: A Unified Approach," *Economics Letters*, 70(1), 99–105.

Rodríguez-Álvarez, C. (2006): "Candidate Stability and Voting Correspondences," *Social Choice and Welfare*, 27(3), 545–570.

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(2), 187–217.

Sen, A. (2001): "Another Direct Proof of the Gibbard-Satterthwaite Theorem," *Economics Letters*, 70(3), 381–385.

Vohra, R. V. (2011): *Mechanism Design: A Linear Programming Approach*, Econometric Society Monographs Series. Cambridge University Press, New York.

School of Economics, University of Queensland
E-mail address, P. Man: `t.man@uq.edu.au`

E-mail address, S. Takayama: `s.takayama@economics.uq.edu.au`