13 Social Choice

The fundamental question in social choice theory is the following: Given a collection of agents with preferences over different *alternatives (*allocations, outcomes), how should society evaluate these alternatives. Put differently, are there any "fair" ways of aggregating individual preferences into a "welfare function"? The most important result is the famous (im)possibility theorem due to Arrow, which says that it is impossible to find a "universal aggregator" that works for all problems of social choice and only uses the *ordinal* structure of preferences.

Let:

- X be a set of "social alternatives".

- R be a *preference ordering* or *weak ordering* on X. This means that R describes some "sensible ranking" of the alternatives in X. By "sensible" we mean that R satisfies a couple of axioms which together imply that R agrees with our common sense usage of the word ranking. These axioms are:

  Axiom 1 (Completeness) *For all* $x, y \in X$ *either* xRy *or* yRx

  This simply means that all alternatives can be compared. Note that (since we are considering *weak* orderings) *both* xRy and yRx may be true.

  Axiom 2 (Transitivity) *For all* $x, y, z \in X$. *If* xRy *and* yRz, *then* xRz.

- From a weak preference ordering, a strict preference and an indifference relation can be derived if desired. I.e., from R we can define P as the relation where xPy if xRy and not yRx. This is the strict preference order associated with R. Similarily, we can define $\sim$ as the relation $x \sim y$ if xRy and yRx to get the associated indifference relation.

- $I = \{1, ..I\}$ is the set of agents of society

- $\mathcal{R} = \{R | R \text{ is a weak ordering on } X\}$

We are now in the position to formally describe a social welfare function. The role of this object is to assign a *social ordering* R to each configuration $(R_1, .., R_I)$ of individual ordering.

Remark 1 (very very important) *The problem of social choice as posed by Arrow is to find a procedure that works for all conceivable preference orderings. Hence, while in welfare maximization problems we usually think of the preferences as given, we will now consider a procedure that should work for all "economies" or "problems". All bite comes from this search for a general procedure that "solves all problems".*

Definition 1 *A social welfare function is is a function* $F : \mathcal{R}^I \to \mathcal{R}$.

Interpretation The way to think about it is as a process that assigns "social preferences" $F(R_1, ..., R_I)$ to any collection of individual preferences $(R_1, ..., R_I)$.

Remark 2 *Since the image of* F *is in* $\mathcal{R}$ *we have built in the requirements that a social welfare function must always generate transitive and complete "social preferences".*

13.1 Arrows Desiderata

If we don't have any requirements for how the "social preferences" depend on individual preferences we can obviously trivially generate welfare functions. However, the way to think about how Arrow set up the problem is that he wanted to ask whether there are ways to "aggregate preferences" using ordinal information of preferences only. He therefore proceeded by postulating some reasonable axioms on how the social preferences should relate with individual preferences.

Definition 2 $i \in N$ *is a dictator for* F *if for all* $x, y \in X$ *and all* $(R_1, ..., R_I) \in \mathcal{R}^I$:

$$xP_iy \Rightarrow xPy$$

where P_i, P *are strict preference orderings that can be defined from the weak preference ordering* $F(R_1, .., R_I)$.

Definition 3 F *satisfies non-dictatorship if* F *has no dictator.*

Clearly, "aggregation" by just making the social preferences coincide with the individual preferences of some particular agent is not that interesting (or appealing) procedure.

Definition 4 F *satisfies unanimity if for all* $x, y \in X$ *and any* $(R_1, ..., R_I) \in \mathcal{R}^I$ *such that* xP_iy *for all* $i \in N$, *then* xPy

This is the Pareto principle.

Definition 5 *F satisfies independence of irrelevant alternatives if $(R_1, .., R_I)$ and $(R'_1, ..., R'_I)$ is such that they agree on the choice between a pair x, y ($xR_iy \Leftrightarrow xR'_iy$ for all $i \in N$), then $R = F(R_1, ..., R_I)$ and $R' = F(R'_1, ..., R'_I)$ also agree on the choice between x and y (i.e. $xF(R_1, ..., R_I)\, y \Leftrightarrow xF(R'_1, ..., R'_I)\, y$*

We will follow Arrow and accept non-dictatorship (ND), unanimity (U) and independence of irrelevant alternatives (IIA) as reasonable minimal criteria for a social welfare function.

13.2 Arrows Theorem

Theorem 1 *If $|X| \geq 3$ there exists no social welfare function satisfying ND, U and IIA*

Proof. For contradiction we assume that there is some $F : \mathcal{R}^I \to \mathcal{R}$ satisfying ND, U and IIA. Now fix $x, y \in X$ and $S \subset I$.

Definition 6 *$S \subset I$ is said to be decisive for the pair (x, y) if xPy for all $\boldsymbol{R}= (R_1, .., R_I)$ such that xP_iy for all agents $i \in S$ and yP_ix for all $i \notin S$*

We will use notation xD_Sy to say that S is decisive for the comparison between x and y if all agents in S prefers x to y. To simplify the proof we will assume that there are at least 5 elements in X. This is purely for convenience.

Lemma 1 *Let x, y be distinct alternatives. Then:*

1. *$xD_Sy \Rightarrow xD_Sz$ for any $z \neq x$*
2. *$xD_Sy \Rightarrow zD_Sy$ for any $z \neq y$*

Proof. (Part 1) The key to all these arguments is that IIA implies that we only need to consider the alternatives under consideration. Assume that xD_Sy and consider the strict preference orderings

P_S	P_N
x	y
y	z
z	x

and assume that every $i \in S$ has preferences P_S and that every $i \notin S$ has preferences P_N. Now:

i) yP_iz for all $i \in I \Rightarrow$(unanimity axiom) yPz

ii) xD_Sy (and the two preference profiles being so that xP_iy for $i \in S$ and yP_ix for $i \notin S$)$\Rightarrow xPy$

iii) xPy and yPz (transitivity of P)$\Rightarrow xPz$

iv) xP_iz for all $i \in S$ and zP_ix for all $i \notin S$ and $xPz \Rightarrow$(definition of decisive group) xD_Sz.

(Part 2) Now consider

P_S	P_N
z	y
x	z
y	x

and assume that every $i \in S$ has preferences P_S and that every $i \notin S$ has preferences P_N. Then:

i) zP_ix for all $i \in I \Rightarrow$(unanimity axiom) xPz

ii) xD_Sy (and the two preference profiles being so that xP_iy for $i \in S$ and yP_ix for $i \notin S$)$\Rightarrow xPy$

iii) zPx and xPy (transitivity of P)$\Rightarrow zPy$

iv) zP_iy for all $i \in S$ and yP_iz for all $i \notin S$ and $zPy \Rightarrow$(definition of decisive group) zD_Sy. ■

Lemma 2 *Let x, y, z be distinct alternatives. Then:*

1. *$xD_Sy \Rightarrow zD_Sw$ for any $w \in X \setminus \{z\}$*
2. *$xD_Sy \Rightarrow wD_Sz$ for any $w \in X \setminus \{z\}$*

Proof. (1) By Part 2 of Lemma 1 $xD_Sy \Rightarrow zD_Sy$. Applying Part 1 (to the pair $\{z, y\}$) it follows that zD_Sw for any $w \neq z$. (2) By Part 1 of Lemma 1 $xD_Sy \Rightarrow xD_Sz$. Applying Part 2 (to the pair $\{z, x\}$) it follows that wD_Sz for any $w \neq z$. ■

Lemma 3 *Suppose there exists $\{x, y\}$ and $S \subset I$ such that xD_Sy. Then zD_Sw for all $z, w \in X$.*

Proof. Take any $\{z, w\} \in X$. If $z \notin \{x, y\}$ then the previous Lemma gives the result directly. If $z = x$ then Lemma 1 gives the result. Finally, if $z = y$, then we know that there is some alternative $v \in X$ distinct from $\{x, y\}$. Applying the Lemma above (to x, y, v) we conclude that $xD_Sy \Rightarrow vD_Sw$ for any $w \neq v$. Applying it again (to (v, w, y) it follows that $vD_Sw \Rightarrow yD_Sw$. Since the case we consider is when $z = y$ we have thus shown that zD_Sw. The cases we have considered are exhaustive, so the result follows. ■

What we have concluded is thus that if a group S have the same preferences for some particular ranking and the rest of the agents have the opposing preferences, then the group will "always have its will" as long as the group shares the same rank and all others have the opposite relative ranking.

Define:

$$\begin{aligned} W &= \{S \subset I | xD_S y \text{ for some } (x,y) \in X \times X\} \\ \mathcal{L} &= 2^I \backslash W \end{aligned}$$

I.e. think of W as the set of coalitions that would "win" if they shared the same preferences and all others had the opposite view. It now seems rather plausible that S would still "win" even if we let the opposite group have arbitrary preferences.

Lemma 4 *For all $x, y \in X$, if $S \in W$ and $xP_i y$ for all $i \in S$, then xPy*

The difference between this claim and the earlier is that there are no restrictions on preferences outside S (as in definition of "Decisive group")

Proof. Consider a preference profile where the strict preference ordering implied for all agents in S are as follows

$$\begin{array}{c} P_S \\ x \\ t \\ y \end{array}$$

where t is an arbitrary alternative different from x and y. For agents outside S, let the preference orderings be arbitrary. However, using IIA, the ranking in society between x and y is unaffected by moving t to the top for every $i \notin S$, so assume that t is the most preferred alternative for all $i \notin S$.Then:

i) $S \in W \Rightarrow$exists x, y such that $xD_S y$. By Lemma 3 this implies $xD_S t$, which since $tP_i x$ for all $i \notin S$ and $xP_i t$ for all $i \in S$ implies xPt.

ii)$tP_i y$ for all $i \in I$, so tPy by unanimity.

iii)xPt and $tPy \Rightarrow xPy$ (transitivity). ■

Lemma 5 *$F, G \in \mathcal{L}$ and $F \cap G = \phi \Rightarrow F \cup G \in \mathcal{L}$*

Proof. Consider preference profile $(R_1, ..., R_I)$ such that the induced strict preferences over three alternatives x, y, z are given by P_F for all $i \in F$, P_G for all $i \in G$ and P_N for all $i \in I \backslash F \cup G$,where

P_F	P_G	P_N
x	z	y
y	x	z
z	y	x

Now:

i) $F \in L \Rightarrow zRx$ (i.e. not xPz). For contradiction, suppose xPz. Then F would be a decisive group between x and z (all others rank z above x), so $F \in W$, which is a contradiction.

ii) $G \in L \Rightarrow yRz$, since G would otherwise be decisive.

iii) yRz and $zRx \Rightarrow$(transitivity) yRx. Hence $S \cup G \in L$. ■

To finish the proof we just need to observe that non-dictatorship$\Rightarrow$ $\{i\} \in L$ for all i and by Lemma 5 this in turn implies that $I \in L$, which contradicts unanimity. ■

13.3 What does Arrows theorem tell us?

The force of the result is that it says that any conceivable procedure to aggregate preferences will violate at least one of the assumptions, so the search for the universal aggregation procedure is a dead end, unless we are willing to commit to the sin of making interpersonal comparisons of utilities.

In principle, one could argue that the postulates of what is a reasonable social welfare function are too strict, but waving the unanimity requirement opens the possibility for a welfare function that picks an arbitrary alternative, irrespective of individual preferences. Similarly, waving non-dictatorship doesn't seem very interesting. The only remaining postulate to relax is thus the independence axiom.

If independence is waved there are lots of social welfare functions that fulfill the remaining axioms. For example, Borda rules would work just fine. The obvious objection is of course that the social ranking would be affected by changes in the ranges of alternatives (so that the relative rank of UW and Temple could be affected if Notre Dame drops out from the ranking due to violations of NCAA rules).

It was also noted by Vickrey that there is a close relation between the Independence requirement and possibilities for truthful revelation of preferences. Vickreys first insight is that if the independence axiom and a "non-perversity" postulate would be satisfied, then there would be no gains from misreporting preferences. This non-perversity postulate was actually one of the explicit postulates in Arrow's original proof of the (im)possibility theorem (Arrow used slightly weaker forms of unanimity and non-dictatorship), but can be derived from the independence and unanimity axioms in the form of these notes. The requirement is that

Definition 7 *For any $i \in I$, let R_i and R_i' be two rankings such that xP_iy and $yP_i'x$. Let $\boldsymbol{R}_{-i} = \{R_1, ..., R_{i-1}, R_{i+1}, .., R_n\}$ be some arbitrary collection of preorderings for the other agents. The social welfare function satisfies non-perversity if there exists no pair $x, y \in X$, no agent $i \in I$, no collection of preorderings for the other agents $\boldsymbol{R}_{-i} = \{R_1, ..., R_{-1}, R_{+1}, .., R_n\} \in \mathcal{R}^{n-1}$ and no pair of preorderings R_i and R_i' satisfying xP_iy and $yP_i'x$ such that yPx and xPy (where P is the strict preference operator associated with $F(R_i, \boldsymbol{R}_{-i})$ and P' with $F(R_i', \boldsymbol{R}_{-i})$).*

In words, non-perversity means that a change in preferences for an individual shall never change the social preferences in the opposite order.

Now, suppose that non-perversity is satisfied and we are considering the relative merits of alternatives x and y. Exactly like in majority voting between two alternatives there is then absolutely no point for a person who likes x to claim that her actual preferences are the opposite (we are now supposing that the alternative ranked the highest by society is actually implemented so that people care about the aggregation procedure). It is a dominant strategy to tell the truth since no matter what the others are doing: non-perversity implies that an individual who likes x to y and as a result of reporting this society nevertheless implements y, then the choice can't be x if the individual reports y. By the independence requirement, changes in reported preferences concerning other alternatives can have no effect on the relative ranking between x and y, so no matter what the other agents are doing there are no incentives to misreport preferences.

Vickrey conjectured that the converse is also true: if the independence criterion is not satisfied, then the social welfare function is not immune against strategic misrepresentations of preferences. This conjecture is almost the statement of the Gibbard-Sattertwaite theorem, which was proved some ten year later.

13.4 Solutions?

One may argue that the problem of social choice as formulated by Arrow (and Bergson) is a rather ambitious one, maybe too ambitious for most purposes. Maybe society doesn't need a social preordering at all. For example, the Pareto criterion defines a non-dictatorial social decision rule that satisfies unanimity and independence if the rule is completed with indifference between all alternatives that are not Pareto rankable. Clearly, in some circumstances this produces too much indifference and the "indifference relation" or "indecision relation" would lack the transitivity properties of a standard indifference relation: x may Pareto dominate z even if x does not Pareto dominate y and y does not Pareto dominate z. However, the Pareto criterion has proved very useful both in Public and many other fields.

Another much used "solution" is to put restrictions on individual preferences. Arrows' theorem relies crucially on the requirement that the welfare function must work for all possible collections of individual orderings that are conceivable. In some cases this may be an unnecessarily stringent assumption. For example, sometimes it may be natural to the problem that alternatives will be ranked in some systematic fashion.

The by far most popular restriction on individual preferences in the literature is that of "single-peaked" preferences that we have already discussed in connection with our informal discussion on voting over the level of a public good. If we start with an abstract (finite) set of alternatives X the definition is restated as:

Definition 8 *The set of admissable individual rankings is single-peaked if the alternatives i X can be ordered along a one dimensional continuum in such a way such that if $x_i \in X$ is i's most preferred alternative, then alternatives rank successively lower in either direction along the continuum.*

Obviously, in a society with a single agent we can always order the alternatives so that they are single-peaked (for example call agent i's most preferred alternative 1, second $2, ...$) and the bite in the restriction comes from the requirement that the ordering along the continuum is *the same for all agents.*