Arrow's impossibility theorem

Hiroaki Sakamoto

July 31, 2015

1 Preference

Consider an economy consisting of $n \geq 2$ individuals. Let $I \subseteq \mathbb{N}$ be the set of all individuals, i.e.,

$$I := \{1, 2, \ldots, n\}. \tag{1}$$

For each nonempty subset $L \subseteq I$, we write $|L| \in \mathbb{N}$ to mean the cardinality of L. Roughly speaking, the cardinality of a set is the number of elements contained in the set. If $L = \{2, 4, 8\}$, for instance, then $|L| = 3$. Denote by X the set of alternatives, which we assume contains more than three elements. Each individual has a preference $\succsim_i$ on X. As usual, we interpret $x \succsim_i x'$ as meaning that for individual $i \in I$, an alternative x is at least as good as another alternative x'. Similarly, we write $x \succ_i x'$ to mean that x is strictly better than x' for $i \in I$. To be more precise, we define $\succ_i$ by

$$x \succ_i x' \iff x \succsim_i x' \text{ and not } x' \succsim_i x \tag{2}$$

for each $x, x' \in X$.

Let us define two conditions we want every preference to satisfy.

> Definition: **Completeness and transitivity**
>
> A preference $\succsim_i$ on X is said to be
>
> (a) *complete* if for any $x, x' \in X$,
>
> $$x \succsim_i x' \text{ or } x' \succsim_i x; \tag{3}$$
>
> (b) *transitive* if for any $x, x', x'' \in X$,
>
> $$x \succsim_i x' \text{ and } x' \succsim_i x'' \implies x \succsim_i x''. \tag{4}$$

We note in particular that if $\succsim_i$ is transitive, then

$$x \succ_i x' \text{ and } x' \succ_i x'' \implies x \succ_i x'' \tag{5}$$

for any $x, x', x'' \in X$. Let $\mathcal{C}$ be the set of all complete preferences on X and $\mathcal{B} \subseteq \mathcal{C}$ be the set of all complete and transitive preferences on X.

A *preference profile* is an n-tuple

$$\rho := (\succsim_1, \succsim_2, \dots, \succsim_n) \in \mathcal{B}^n, \tag{6}$$

which describes the preferences of all individuals in the economy.

2 Aggregation

Let us now define an aggregation rule.

Definition: Aggregation rule

An *aggregation rule* on $\mathcal{B}^n$ is a function $\succsim : \mathcal{B}^n \to \mathcal{C}$, which assigns an aggregate preference on X to every possible preference profile in $\mathcal{B}^n$.

Notice that the aggregate preference is not necessarily transitive. For notational convenience, we denote by $\succsim_\rho \in \mathcal{C}$ the aggregate preference associated with a preference profile $\rho \in \mathcal{B}^n$ under an aggregation rule $\succsim : \mathcal{B}^n \to \mathcal{C}$.

We consider a couple of criteria an aggregation rule might satisfy.

Definition: Unanimity, independence, and dictatorship

An aggregation rule $\succsim$ on $\mathcal{B}^n$ is said to be

(a) *transitive* if $\succsim_\rho$ is transitive for every $\rho \in \mathcal{B}^n$.

(b) *unanimous* if

$$x \succ_i x' \text{ for all } i \in I \implies x \succ_\rho x' \tag{7}$$

for every $\rho = (\succsim_1, \dots, \succsim_n) \in \mathcal{B}^n$ and for every $x, x' \in X$.

(c) *independent* if

$$\begin{aligned} &\succsim_i \text{ and } \succsim_i' \text{ have the same ranking over } \{x, x'\} \text{ for all } i \in I, \\ &\implies \succsim_\rho \text{ and } \succsim_{\rho'} \text{ have the same ranking over } \{x, x'\} \end{aligned} \tag{8}$$

for every $\rho = (\succsim_1, \dots, \succsim_n), \rho' = (\succsim_1', \dots, \succsim_n') \in \mathcal{B}^n$ and for every $x, x' \in X$.

(d) *dictatorial* if there exists $i \in I$ such that

$$x \succ_i x' \implies x \succ_\rho x' \tag{9}$$

for every $\rho = (\succsim_1, \dots, \succsim_n) \in \mathcal{B}^n$ and for every $x, x' \in X$.

(e) *non-dictatorial* if it is not dictatorial.

It should be easy to see what is meant by these conditions, except for condition (c). Condition (c) is often called *independence of irrelevant alternatives*, *pairwise independence*, or *binary independence*. This condition requires that when two alternatives are compared, their relative positions against the other alternatives should not be taken into account.

To facilitate the discussion below, we also introduce the concept of *decisiveness* as follows.

Definition: Decisive and semidecisive group

Consider an aggregation rule $\succsim$ on $\mathcal{B}^n$ and a nonempty subset $L \subseteq I$ of individuals. We say that

(a) L is *decisive under* $\succsim$ if

$$x \succ_i x' \text{ for all } i \in L \implies x \succ_\rho x' \tag{10}$$

for every $\rho = (\succsim_1, \dots, \succsim_n) \in \mathcal{B}^n$ and for every $x, x' \in X$.

(b) L is *decisive for* $x \in X$ *against* $x' \in X$ *under* $\succsim$ if

$$x \succ_i x' \text{ for all } i \in L \implies x \succ_\rho x' \tag{11}$$

for every $\rho = (\succsim_1, \dots, \succsim_n) \in \mathcal{B}^n$.

(c) L is *semidecisive for* $x \in X$ *against* $x' \in X$ *under* $\succsim$ if

$$x \succ_i x' \text{ for all } i \in L \text{ and } x' \succ_j x \text{ for all } j \in I \setminus L \implies x \succ_\rho x' \tag{12}$$

for every $\rho = (\succsim_1, \dots, \succsim_n) \in \mathcal{B}^n$.

With this definition, one could say that an aggregation rule $\succsim: \mathcal{B}^n \to \mathcal{C}$ is unanimous if and only if I is decisive under $\succsim$. Similarly, it should be easy to see that an aggregation rule $\succsim: \mathcal{B}^n \to \mathcal{C}$ is dictatorial if and only if there exists $L \subseteq I$ such that $|L| = 1$ and L is decisive under $\succsim$.

The following lemma is fundamental for our purpose.

Lemma: Field expansion lemma

Suppose that an aggregation rule $\succsim$ on $\mathcal{B}^n$ is transitive, unanimous, and independent. Then, for any nonempty subset $L \subseteq I$, the following are equivalent:

(a) L is decisive under $\succsim$;

(b) for some $x, x' \in X$, L is semidecisive for x against x' under $\succsim$.

Proof. Since (a) obviously implies (b), it suffices to prove the converse. Suppose (b) is true so that there exist $x, x' \in X$ such that L is semidecisive for x against x' under $\succsim$. What we want to show is that L is decisive under $\succsim$. We divide the proof into five steps.

Step 1:

We first prove that for any $x'' \in X \setminus \{x, x'\}$, L is decisive for x against x'' under $\succsim$. Fix $x'' \in X \setminus \{x, x'\}$ arbitrarily and let $\rho = (\succsim_1, \dots, \succsim_n) \in \mathcal{B}^n$ be a preference profile such that

$$x \succ_i x'' \text{ for all } i \in L. \tag{13}$$

We shall show that

$$x \succ_\rho x''. \tag{14}$$

To this end, consider another preference profile $\rho' = (\succsim'_1, \dots, \succsim'_n) \in \mathcal{B}^n$ such that

$$x \succ'_i x' \text{ and } x' \succ'_i x'' \text{ for all } i \in L, \tag{15}$$

$$x' \succ'_j x \text{ and } x' \succ'_j x'' \text{ for all } j \in I \setminus L, \tag{16}$$

and

$$\succsim_j \text{ and } \succsim'_j \text{ have the same ranking over } \{x, x''\} \text{ for all } j \in I \setminus L. \tag{17}$$

Since L is semidecisive for x against x' under $\succsim$, it follows from (15) and (16) that

$$x \succ_{\rho'} x'. \tag{18}$$

Since $\succsim$ is unanimous, on the other hand, combining (15) and (16) yields

$$x' \succ_{\rho'} x''. \tag{19}$$

Hence, by transitivity of $\succsim_{\rho'}$, we have

$$x \succ_{\rho'} x''. \tag{20}$$

Since $\succsim$ is independent, it follows from (13), (15), and (17) that

$$\succsim_\rho \text{ and } \succsim_{\rho'} \text{ have the same ranking over } \{x, x''\}, \tag{21}$$

which, together with (20), implies (14) as desired.

Step 2:

Next we prove that L is decisive for x against x' under $\succsim$ as well. Let $\rho = (\succsim_1, \dots, \succsim_n) \in \mathcal{B}^n$ be a preference profile such that

$$x \succ_i x' \text{ for all } i \in L. \tag{22}$$

We shall show that

$$x \succ_\rho x'. \tag{23}$$

To this end, fix $x'' \in X \setminus \{x, x'\}$ and consider another preference profile $\rho' = (\succsim'_1, \dots, \succsim'_n) \in \mathcal{B}^n$ such that

$$x \succ'_i x'' \text{ and } x'' \succ'_i x' \text{ for all } i \in L, \tag{24}$$

$$x'' \succ'_j x' \text{ for all } j \in I \setminus L, \tag{25}$$

and

$$\succsim_j \text{ and } \succsim'_j \text{ have the same ranking over } \{x, x'\} \text{ for all } j \in I \setminus L. \tag{26}$$

Notice that from Step 1 above, we know that L is decisive for x against x'' under $\succsim$. Hence, (24) directly implies

$$x \succ_{\rho'} x''. \tag{27}$$

Since $\succsim$ is unanimous, on the other hand, combining (24) and (25) yields

$$x'' \succ_{\rho'} x'. \tag{28}$$

Then, by transitivity of $\succsim_{\rho'}$, we have

$$x \succ_{\rho'} x'. \tag{29}$$

Since $\succsim$ is independent, it follows from (22), (24), and (26) that

$$\succsim_{\rho} \text{ and } \succsim_{\rho'} \text{ have the same ranking over } \{x, x'\}, \tag{30}$$

which, together with (29), implies (23).

Step 3:

As a third step, we prove that for any $x'' \in X \setminus \{x, x'\}$, L is decisive for x' against x'' under $\succsim$. Fix $x'' \in X \setminus \{x, x'\}$ arbitrarily and let $\rho = (\succsim_1, \ldots, \succsim_n) \in \mathcal{B}^n$ be a preference profile such that

$$x' \succ_i x'' \text{ for all } i \in L. \tag{31}$$

We shall show that

$$x' \succ_{\rho} x''. \tag{32}$$

To this end, consider another preference profile $\rho' = (\succsim'_1, \ldots, \succsim'_n) \in \mathcal{B}^n$ such that

$$x' \succ'_i x \text{ and } x \succ'_i x'' \text{ for all } i \in L, \tag{33}$$

$$x' \succ'_j x \text{ for all } j \in I \setminus L, \tag{34}$$

and

$$\succsim_j \text{ and } \succsim'_j \text{ have the same ranking over } \{x', x''\} \text{ for all } j \in I \setminus L. \tag{35}$$

Notice that from Step 1 above, we know that L is decisive for x against x'' under $\succsim$. Hence, (33) directly implies

$$x \succ_{\rho'} x''. \tag{36}$$

Since $\succsim$ is unanimous, on the other hand, combining (33) and (34) yields

$$x' \succ_{\rho'} x. \tag{37}$$

Then, by transitivity of $\succsim_{\rho'}$, we have

$$x' \succ_{\rho'} x''. \tag{38}$$

Since $\succsim$ is independent, it follows from (31), (33), and (35) that

$$\succsim_{\rho} \text{ and } \succsim_{\rho'} \text{ have the same ranking over } \{x', x''\}, \tag{39}$$

which, together with (38), implies (32) as desired.

Step 4:

We then prove that L is decisive for x' against x under $\succsim$ as well. Let $\rho = (\succsim_1, \ldots, \succsim_n) \in \mathcal{B}^n$ be a preference profile such that

$$x' \succ_i x \text{ for all } i \in L. \tag{40}$$

We shall show that

$$x' \succ_\rho x. \tag{41}$$

To this end, fix $x'' \in X \setminus \{x, x'\}$ and consider another preference profile $\rho' = (\succsim'_1, \ldots, \succsim'_n) \in \mathcal{B}^n$ such that

$$x' \succ'_i x'' \text{ and } x'' \succ'_i x \text{ for all } i \in L, \tag{42}$$

$$x'' \succ'_j x \text{ for all } j \in I \setminus L, \tag{43}$$

and

$$\succsim_j \text{ and } \succsim'_j \text{ have the same ranking over } \{x, x'\} \text{ for all } j \in I \setminus L. \tag{44}$$

Notice that from Step 3 above, we know that L is decisive for x' against x'' under $\succsim$. Hence, (42) directly implies

$$x' \succ_{\rho'} x''. \tag{45}$$

Since $\succsim$ is unanimous, on the other hand, combining (42) and (43) yields

$$x'' \succ_{\rho'} x. \tag{46}$$

Then, by transitivity of $\succsim_{\rho'}$, we have

$$x' \succ_{\rho'} x. \tag{47}$$

Since $\succsim$ is independent, it follows from (40), (42), and (44) that

$$\succsim_\rho \text{ and } \succsim_{\rho'} \text{ have the same ranking over } \{x, x'\}, \tag{48}$$

which, together with (47), implies (41) as required.

Step 5:

Finally, we prove that L is decisive under $\succsim$. In other words, we shall show that for every $\tilde{x}, \tilde{x}' \in X$, L is decisive for $\tilde{x}$ against $\tilde{x}'$ under $\succsim$. There are two cases to consider: $\tilde{x} \in \{x, x'\}$ and $\tilde{x} \notin \{x, x'\}$. The first case has already been dealt with in the preceding steps. Let us focus on the second case and fix $\tilde{x} \in X \setminus \{x, x'\}$ arbitrarily. Then, from Step 1 above, we know that L is decisive for x against $\tilde{x}$ under $\succsim$, which in turn implies that L is semidecisive for x against $\tilde{x}$ under $\succsim$. Hence, from Steps 3 and 4, we conclude that L is decisive for $\tilde{x}$ against $\tilde{x}'$ under $\succsim$ for any $\tilde{x}' \in X$. Since $\tilde{x}$ is chosen arbitrarily, this completes the proof. ■

3 Theorem

The following theorem, originally proven by Arrow (1951), shows that it is impossible to design an aggregation rule which simultaneously satisfies unanimity, independence, and non-dictatorship.

> Theorem: **Arrow's impossibility theorem**
>
> If an aggregation rule on $\mathcal{B}^n$ is transitive, unanimous, and independent, then it must be dictatorial.

Proof. Suppose that an aggregation rule $\succsim : \mathcal{B}^n \to \mathcal{C}$ is transitive, unanimous, and independent. We shall show that $\succsim$ must be dictatorial. In the light of lemma above, it suffices to show that there exists a subset $L \subseteq I$ with $|L| = 1$ such that for some $x, x' \in X$, L is semidecisive for x against x' under $\succsim$.

First notice that for each $\tilde{x}, \tilde{x}' \in X$, the set

$$\{|L| \in \mathbb{N} \mid L \text{ is semidecisive for } \tilde{x} \text{ against } \tilde{x}' \text{ under } \succsim\} \tag{49}$$

is a nonempty subset of $\mathbb{N}$ because, by unanimity, I is decisive under $\succsim$ and thus $n = |I|$ is contained in this set. Hence, we can define a function $\lambda : X \times X \to \mathbb{N}$ by

$$\lambda(\tilde{x}, \tilde{x}') := \min\{|L| \in \mathbb{N} \mid L \text{ is semidecisive for } \tilde{x} \text{ against } \tilde{x}' \text{ under } \succsim\}. \tag{50}$$

Define $\lambda^* \in \mathbb{N}$ by

$$\lambda^* := \min\{\lambda(\tilde{x}, \tilde{x}') \in \mathbb{N} \mid \tilde{x}, \tilde{x}' \in X\}. \tag{51}$$

Denote by $L^* \subseteq I$ the nonempty subset of I such that $|L^*| = \lambda^*$ and let $x, x' \in X$ be a pair of alternatives such that L^* is semidecisive for x against x' under $\succsim$. We will be done if we show $\lambda^* = 1$.

Suppose, by way of contradiction, that $\lambda^* = |L^*| \geq 2$. Fix $i \in L^*$ and choose $x'' \in X \setminus \{x, x'\}$ arbitrarily. Let $\rho = (\succsim_1, \ldots, \succsim_n) \in \mathcal{B}^n$ be a preference profile such that

$$x \succ_i x' \text{ and } x \succ_i x'' \tag{52}$$

and

$$x \succ_j x' \text{ and } x'' \succ_j x' \text{ for all } j \in L^* \setminus \{i\}. \tag{53}$$

Since $\succsim$ is transitive, unanimous, and independent, and since L^* is semidecisive for x against x' under $\succsim$, the field expansion lemma shows that L^* is decisive under $\succsim$. Hence, (52) and (53) immediately imply

$$x \succ_\rho x'. \tag{54}$$

Also, it follows from (52) and (53) that

$$x' \succsim_\rho x'' \tag{55}$$

because otherwise $L\setminus\{i\}$ would be decisive (and hence semidecisive) for x'' against x' under $\succsim$ and $|L\setminus\{i\}| < \lambda^*$, a contradiction[1]. Similarly, it follows from (52) and (53) that

$$x'' \succsim_\rho x \tag{56}$$

because otherwise $\{i\}$ would be decisive (and hence semidicisive) for x against x'' under $\succsim$ and $|\{i\}| = 1 < \lambda^*$, another contradiction[2]. Combining (54), (55), and (56), however, violates the transitivity of $\succsim_\rho$. Therefore, it must be the case that $\lambda^* < 2$ and we conclude that $\lambda^* = 1$. ■

This proof closely follows the one provided by Austen-Smith and Banks (2000).

References

Arrow, K.J. (1951), *Social choice and individual values*, John Wiley & Sons.

Austen-Smith, D. and J.S. Banks (2000), *Positive political theory I: collective preference*, University of Michigan Press.

[1]Notice that we did not impose any restriction on ρ about how x' and x'' are ranked by the individuals in $I\setminus(L^*\setminus\{i\})$.

[2]Notice that we did not impose any restriction on ρ about how x and x'' are ranked by the individuals in $I\setminus\{i\}$.