CHAPTER 2

Pivotal voters: A simple proof of Arrow's theorem

SALVADOR BARBERA*
Universidad de Bilbao

This chapter presents a new proof of Arrow's "General Possibility Theorem". Since there is no need to elaborate on the importance and the widespread impact of the theorem, we limit our comments here to an outline of the main features of the present approach to its proof.

Arrow's theorem deals with social welfare functions, i.e. rules which assign one "social" preference ordering of the alternatives facing society to each n-tuple (profile) of logically possible "individual" preference orderings of the same alternatives. One way to describe a social welfare function is by appropriately specifying a rich enough structure of coalitions of individuals which are "decisive", in the sense that unanimity of strong preferences within the group regarding a given pair of alternatives is sufficient to determine society's preference vis-à-vis those alternatives under the given social welfare function. Most proofs of Arrow's theorem consist in analyzing the structure of the family of decisive coalitions associated with a Paretian social welfare function independent of irrelevant alternatives, and show that these families must contain one decisive coalition consisting of one individual only – a dictator.

Our approach does not focus so much on the "social" image of given profiles, but on the way "social" preferences change in response to changes in the preferences of individuals. One individual is *pivotal* for a pair of alternatives at a preference profile if he could change the social ordering of these alternatives by just changing his preferences. Once we know what the "social" ordering will be for

* I am grateful to a referee for his help in improving the exposition.

Social Choice and Welfare, edited by P.K. Pattanaik and M. Salles

some given profile (and this is what the Paretian requirement does for us), describing a social welfare function is describing who can change the outcome of the function, when, and by what actions – it is a matter of distributing among individuals the ability to act as pivots. Our proof of Arrow's theorem consists in showing that, under Paretian and independent of irrelevant alternatives social welfare functions, pivotality must be essentially concentrated in the hands of a single individual – the dictator. The essential role of the independence condition, even in the absence of any Paretian requirement, is emphasized by a lemma.

It is hoped that this alternative way of looking at a classical problem might be complementary to others in reaching a better understanding of how to design mechanisms for collective decision-making. In particular, the idea of pivotality has proved basic in the literature on the strategy-proof allocation of public goods, and it might prove a useful bridge between this parallel flow of work and that of social choice theory.

Individuals

Let $I = \{1, 2, ..., n\}$, an initial segment of the integers. Elements of I are called the *individuals*.

Alternatives

Let A be a set (finite or infinite). Elements of A are denoted by $x, y, z, ...$, and are called the *alternatives*.

Preference relations

Let $\mathscr{R}$ be the set of complete, transitive, reflexive binary relations on A. Elements of $\mathscr{R}$ are denoted by $R, R', R_i, R_j, ...$, and are called *preference orderings*. Given $R \in \mathscr{R}$, the *strict preference relation P associated with R* is defined so that $(\forall x, y \in A)\ [xPy \longleftrightarrow (xRy \text{ and } \sim yRx)]$.

Let $\overline{\mathscr{R}} \subset \mathscr{R}$ be the set of preference orderings which are antisymmetric. Elements of $\overline{\mathscr{R}}$ are called *strong preference orderings*. They stand for preferences where no alternative is indifferent to any other. We say that $x, y \in A$ *are contiguous in* $R \in \overline{\mathscr{R}}$ iff $(\forall z \notin \langle x, y \rangle)\ [zRx \longleftrightarrow zRy]$.

Preference profiles

Let $\mathscr{R}^n$ stand for the n-fold Cartesian product of $\mathscr{R}$. Elements of $\mathscr{R}^n$ are denoted by $\boldsymbol{R}, \boldsymbol{R}'$, ..., and are called *preference profiles*. The set of *strong preference profiles* $\overline{\mathscr{R}}^n$ is defined likewise.

When there is no ambiguity, R_i and P_i stand for the ith element of $\boldsymbol{R}$ and for its associated strict relation (respectively), R_j' and P_j' for the jth element of $\boldsymbol{R}'$ and its strict relation, and so on.

Given $R \in \mathscr{R}^n$ and $R' \in \mathscr{R}$, $R/_iR'$ denotes the profile R', where $R'_i = R'$ and $(\forall j \neq i)\, R_j = R'_j$.

Social welfare functions

A social welfare function (SWF) is a function $w: \mathscr{R}^n \longleftrightarrow \mathscr{R}$.[1] Whenever there is no ambiguity, we denote $w(R)$ by R, $w(R')$ by R', etc. Similarly, $P, P', \ldots$, will then denote the strict preference orderings associated with $w(R), w(R'), \ldots$. A SWF w is *Paretian* iff $(\forall R \in \mathscr{R}^n)\, (\forall x, y \in A)\, [(\forall i \in I)\, xP_iy \rightarrow xPy]$. A SWF w is *independent of irrelevant alternatives* (IIA) iff $(\forall R, R' \in \mathscr{R}^n)\, (\forall x, y \in A)\, [(\forall i \in I)\, (xR_iy \longleftrightarrow xR'_iy)$ and $(yR_ix \longleftrightarrow yR'_ix) \rightarrow (xRy \longleftrightarrow xR'y)]$.[2]

A SWF w is *dictatorial* iff $\exists d \in I$ such that $(\forall x, y \in A)\, (\forall R \in \mathscr{R}^n)\, [xP_dy \rightarrow xPy]$. Otherwise, w is *nondictatorial*.

Arrow's theorem. Let $A > 2$. There exists no social welfare function which is Paretian, independent of irrelevant alternatives, and nondictatorial.

Pivotal individuals

Individual $i \in I$ is *pivotal* at profile $R \in \mathscr{R}^n$ under a social welfare function w iff $\exists R' \in \mathscr{R}$ such that $w(R/_iR') \neq w(R)$. For $x, y \in A$, individual $i \in I$ is *xy-pivotal* at $R \in \mathscr{R}^n$ under SWF w iff $\exists R' \in \mathscr{R}$ such that $\sim [x\, w(R)\, y \longleftrightarrow x\, w(R/_iR')y]$. For $x, y \in A$, individual $i \in I$ is *positively xy-pivotal* at $R \in \mathscr{R}^n$ under SWF w iff it is xy-pivotal and, for all $R' \in \mathscr{R}$ such that xP'_iy, $\sim [y\, w(R/_iR')\, x]$.

It is clear from the definitions that whenever i is xy-pivotal it is also yx-pivotal. The same symmetric implication is not necessarily true for positive pivotality. Also, note that when w is IIA, if i is xy-pivotal (resp. positively xy-pivotal) at R, it is also xy-pivotal (resp. positively xy-pivotal) at any R' such that $(\forall j \neq i, j \in I)$ $[xR_jy \longleftrightarrow xR'_jy]$.

A lemma on IIA social welfare functions. Under an IIA social welfare function, there can be no strong preference profile at which two individuals are pivotal for two different nondisjoint pairs of alternatives.

[1] Note that Arrow's condition of universal domain, whereby an image in $\mathscr{R}$ has to be attributed to each preference profile, is included in our definition of a SWF.

[2] The condition above is sometimes called binarity. It is *equivalent*, when applied to social welfare functions, to Arrow's original condition of independence of irrelevant alternatives.

Proof. Let w be an IIA social welfare function, and suppose there were such a strong preference profile, individuals, and alternatives. Then, without loss of generality,[3] $\exists R \in \bar{\mathscr{R}}, R_i', R_j'' \in \bar{\mathscr{R}}, i, j \in I$ and $x, y, z \in A$ such that [where $R' = w(R/_iR_i'), R'' = w(R/_jR_j'')$]:

(a) $xRyRz$, $yP'x$, and $zP''y$;

(b) x and y (resp. y and z) are contiguous in R_i and R_i' (resp. in R_j and R_j''); and

(c) $sR_it \leftrightarrow sR_i't$ (resp. $sR_jt \leftrightarrow sR_j''t$) for all pairs of alternatives $\langle s, t\rangle \neq \langle x,y\rangle$ (resp. $\langle s, t\rangle \neq \langle y, z\rangle$).

But then, i and j's pivotality would imply that, where $\hat{R} = w\,[(R/_jR_j'')/_iR_i']$, $z\hat{P}y, y\hat{P}x$ and yet $x\hat{R}z$ – a contradiction to the transitivity of $\hat{R}$. ■

A proof of Arrow's theorem

We start from any *given* Paretian and IIA social welfare function w, and prove that some individual $d \in I$ will be positively pivotal under w for any ordered pair of alternatives at every possible profile – i.e. a dictator.

We first restrict our attention to strong preference profiles, and for all distinct $x, y \in A$, define $I_{xy} = \{i \in I \mid$ for some $R \in \bar{\mathscr{R}}^n$, i is positively xy-pivotal at $R\}$.

(1) By the Pareto rule, $I_{xy} \neq \varnothing$ for all distinct $x, y \in A$. That is, *for every ordered pair of alternatives x, y there will be at least one strong preference profile at which some individual is positively xy-pivotal.*

(2) For any $x, y, z, w \in A$ $(x \neq y, z \neq w)$, and for all $i, j \in I$, if $i \in I_{xy}$ and $j \in I_{zw}$, then $i = j$. That is, *for all strong profiles where some individual d is positively pivotal, this individual must be one and the same.*

To prove this, suppose not; i.e. $\exists x, y, z, w \in A$ $(x \neq y, z \neq w)$ and $i, j \in I$, such that $i \neq j, i \in I_{xy}, j \in I_{zw}$.

(a) If the pairs (x, y) and (z, w) have exactly one element in common, we can construct one profile where i and j are pivotal, in violation of the lemma.

(b) Suppose $x = w$ and $y = z$. Then, for any $v \in A$ $(v \neq x, v \neq y)$, there is a $k \in I_{yv}$ such that either $i \neq k$ or $j \neq k$. This brings us back to situation (a).

(c) If (x, y) and (z, w) have no element in common, there must be a $k \in I_{yz}$ such that either $i \neq k$ or $j \neq k$. This, again, leads to situation (a).

(3) Thus, there exists $d \in I$ such that $I_{xy} = d$ for all distinct $x, y \in A$. This individual d, is positively pivotal for all pairs of alternatives at all strong profiles. To prove it, suppose not – i.e. that for some $R \in \bar{\mathscr{R}}^n$ and some $x, y \in A$, d is not positively xy-pivotal at R. Since, by (1) there are strong profiles where d is posi-

[3] Use is made here of the facts that w is IIA and that we start from a *strong* preference profile. The latter guarantees that none of the pivots is indifferent between the alternatives upon which he can act. The Lemma does not hold true on the larger set of all possible profiles. A counterexample to such a (false) extension is provided by a sequential dictatorship.

tively xy-pivotal, there must exist two strong profiles which only differ in the preferences of one individual $h \neq d$, and such that d is positively xy-pivotal in one of these profiles while not in the other. But then, there must be a strong profile where $h \neq d$ is xy-pivotal, and thus one where h is xy-pivotal and d is yz-pivotal. This contradicts the lemma.

(4) *Individual d is positively pivotal for all pairs of alternatives at all profiles, i.e. a dictator.* To prove it, suppose not. Then, there would exist a profile $\hat{R} \in \mathscr{R}^n$ and alternatives $x, y \in A$, such that $x\hat{P}_d y$ and yet $y\hat{R}x$ (where $\hat{R} = w(\hat{R})$). Consider a strong profile $R \in \overline{\mathscr{R}}^n$ such that $xP_d zP_d y$, for some arbitrary, fixed alternative $z \notin \langle x, y \rangle$, and such that x and y are contiguous at all preferences other than d's at the profile R. According to (3), zPy. Now, define a new profile R' in such a way that $(\forall i \in I)$ $(xR_i'y \longleftrightarrow x\hat{R}_i y)$, and for all $\langle v, w \rangle \neq \langle x, y \rangle$, $(\forall i)$ $(vR_i'w \longleftrightarrow vR_i w)$. Independence of irrelevant alternatives would require that $yR'x, xP'z$, and $zP'y$ – in contradiction with R's transitivity.