

La elección social: racionalidad y decisividad

por

ANDREU MAS COLELL.

University of Minnesota.

Nuestra intención en este artículo es desarrollar algunos aspectos del problema de la elección social en la tradición teórica inaugurada por la obra de K. Arrow¹. Damos primero una introducción general al problema; en la segunda parte enunciamos y demostramos el resultado objeto de este trabajo. Siguen unos comentarios finales.

I

Estamos interesados en el estudio de reglas de elección social de la siguiente forma: la sociedad está formada por un número N de miembros ($i = 1, \dots, N$). Existe un número M de alternativas básicas, de estados sobre los que la sociedad tiene control. Informalmente, una regla de elección social asigna a cada constelación de preferencias individuales sobre las alternativas básicas, una ordenación social de las mismas. Un ejemplo es el método de elección por mayoría simple entre pares de alternativas.

Consideremos este ejemplo más detenidamente. Es bien sabido que la llamada «paradoja del voto» puede presentarse². Si a, b, c son alternativas sociales es posible que a sea «preferida», por mayoría simple,

¹ K. Arrow [2]. Toda la obra merece ser leída. Las páginas 96-100 contienen una síntesis, densa pero completa, del "Teorema de Imposibilidad". Una buena exposición divulgadora es: M. Barbut [3]. Exposiciones de contribuciones posteriores son: Luce y Raiffa [11], (Capítulo 14), A. Sen [18], Y. Murakami [14].

² La "paradoja del voto" fue señalada por primera vez por Condorcet en 1785. Una investigación de los orígenes históricos de la teoría de la elección social está contenida en D. Black [5] (Part II).

a b, b a c y c a a³. La ordenación social entre a, b y c es cíclica. No hay ninguna alternativa que sea preferida a las demás. Esto ocurre aun cuando las preferencias individuales sean totalmente consistentes.

A la vista de esto un problema que se plantea naturalmente es preguntarse por la existencia de alguna regla que conserve todas las buenas propiedades del sistema de voto mayoritario por pares, pero no exhiba resultados paradójicos. En pocas palabras, el «Teorema de Imposibilidad» de Arrow afirma que este objetivo está condenado al fracaso. No existe tal regla.

Evidentemente, como con cualquier proposición demostrable formalmente, el resultado anterior depende de lo que se entienda por «buenas propiedades del sistema de voto mayoritario por pares» y de «comportamiento no paradójico». Se hace necesario un tratamiento más preciso.

Dados un conjunto M de alternativas (x, y, z, ...) designaremos por R cualquier relación reflexiva y completa en M⁴:

i) Si $x \in M$, $x R x$.

Si $x, y \in M$, $x R y$ o $y R x$.

$x R y$ debe leerse: x es, al menos, tan deseada como y. Llamamos P a la relación no reflexiva derivada de R:

$x P y$ si $y R x$ no es cierto.

$x P y$ debe leerse: x es preferida a y. Llamamos I a la relación de equivalencia derivada de R:

$x I y$ si $y R x$ y $x R y$.

$x I y$ debe leerse: x es indiferente a y.

⁴ Por ejemplo, si $N = 3$ y las ordenaciones individuales son las de la tabla:

1	2	3
a	c	b
b	a	c
c	b	a

⁵ Abusando del lenguaje, M y N representan a la vez conjuntos y el número de elementos en ellos. Elementos genéricos de M se designan por las últimas letras del alfabeto (s, t, u, v, ...).

Una relación R será *racional* si es transitiva, es decir:

ii) Si $x, y, z \in M$ y xRy, yRz , entonces xRz .

Observemos que si M es un conjunto finito (como supondremos en lo sucesivo) y R es una relación transitiva, hay un elemento máximo en M respecto a esa relación (un elemento al menos tan deseado como todos los demás).

Diremos que una relación R es decisiva si en cualquier subconjunto M' de M hay un elemento máximo respecto a esa relación. En particular esto implica (pero no es equivalente) que los ciclos de tres elementos de la paradoja del voto no pueden presentarse:

iii) No existen tres elementos $x, y, z \in M$ tales que $xPy,$

yPz, zPx ; es decir:

si xPy, yPz , entonces xRz .

Una relación racional es decisiva, pero una decisiva puede no ser racional; por ejemplo si sólo existen tres alternativas x, y, z la relación definida por (xPy, yPz, xIz) es decisiva pero no racional.

Arrow está interesado en la siguiente formulación de la idea de «no paradojas»: Una «Función de Bienestar Social»⁵ es una función F que asigna a cada N -uplo de relaciones racionales (R_1, \dots, R_N) — las ordenaciones de los N miembros de la sociedad—una relación racional R —la ordenación social. En lo que concierne a las buenas propiedades del sistema de voto por pares, se requiere que F cumpla las siguientes condiciones:

I. (Independencia de Alternativas Irrelevantes)⁶. La ordenación social de las alternativas x e y depende solamente de las ordenaciones individuales de estas alternativas. Es decir, la consideración que una tercera alternativa z merezca a los miembros de la sociedad no influye en la elección entre x e y .

Formalmente: Si (R_1, \dots, R_N) y (R'_1, \dots, R'_N) son dos N -uplos de relaciones individuales y $(xR_1 y$ si y solamente si $xR'_1 y)$, entonces $(xR y$ si y solamente si $xR' y)$.

⁵ Esta es la terminología de Arrow [2]. Es un concepto muy distinto de las Funciones de Bienestar Social de A. Bergson [4] y P. Samuelson [15]. Por supuesto, la igualdad de nombres no es casual. Arrow presenta su resultado como esencialmente, la demostración de la imposibilidad de la función de Bienestar Social de Bergson, lo cual no ha contribuido a aclarar las confusiones.

⁶ Los nombres de las condiciones son los de Arrow.

II. (Condición de Pareto)⁷. Si todos los miembros de la sociedad prefieren x a y , entonces la sociedad prefiere x a y .

III. (Ausencia de dictadura). No hay ningún miembro $i \in N$, tal que para *todo* par de alternativas $x, y \in M$:

Si $x P_i y$, entonces $x P y$.

Para referencia posterior registramos ahora una condición introducida por K. May [13] en su caracterización del método de mayoría simple.

IV. (Respuesta Positiva). Supongamos que dadas las relaciones (R_1, \dots, R_N) resulta $x I y$. Si, *caeteribus paribus*, un miembro $i \in N$ para el que $x I_i y$ pasa a preferir x sobre y , entonces x pasa a ser socialmente preferida a y . Formalmente:

Sean (R_1, \dots, R_N) y (R'_1, \dots, R'_N) dos N -tuplos de relaciones individuales entre $x, y \in M$. Supongamos que:

para todo $j \neq i$, $x R_j y$ si y sólo si $x R'_j y$

para i y $R_i x y$ y $x P'_i y$

o

$y P_i x$ y $x R'_i y$

entonces $x R y$ implica $x P' y$.

El resultado de Arrow puede enunciarse así: *no existe Función de Bienestar Social alguna que cumpla I, II y III.*

Nuestro propósito, aquí, es examinar un aspecto de una de las salidas propuestas al dilema que esta «paradoja del voto generalizada» plantea.

A. K. Sen [17], [18]⁸ ha argumentado que es una exigencia excesiva requerir de las relaciones sociales (en el rango de la función F) que sean racionales. Estaríamos más interesados en la capacidad del proceso político para actuar que en su racionalidad. Sen estudia, por tanto, el caso en que la regla de elección social origina relaciones decisivas y exhibe un ejemplo que cumple todas las condiciones de Arrow⁹. El ejemplo, sin embargo, es muy poco atractivo. La ordenación social es extremadamente poco sensible a los cambios de preferencias individuales. De hecho cada miembro tiene (casi) derecho de veto¹⁰. Natural-

⁷ Nótese que ya desde ahora aceptamos, implícitamente, que la Condición I prevalece. Esto simplifica el enunciado de las condiciones restantes.

⁸ También Schick [16].

⁹ Un ejemplo similar: la sociedad prefiere x a y si y sólo si todos los miembros de la sociedad prefieren x a y .

¹⁰ Véase Sen [18], Guha [9], Gibbard [8], para el estudio de la clase de reglas a la que pertenece la de la nota anterior.

mente, la condición de Respuesta Positiva (IV) no se cumple. Si tomamos a ésta como un índice de sensibilidad, Sen [17] planteó la cuestión de existencia de una función F con relaciones decisivas en su rango cumpliendo I, II, III, IV. La respuesta a esta pregunta es afirmativa [12]¹¹. Nuestro objetivo es probar que la respuesta debe ser negativa si el problema se reformula ligeramente de una forma que creemos natural.

Las Funciones de Bienestar Social de Arrow asignan a N -uplos de relaciones racionales una relación racional. Ello es lo que corresponde al contexto de Economía del Bienestar en que *Social Choice and Individual Values* fue escrito. Sin embargo, si el énfasis se quiere desplazar de racionalidad a decisividad (de normatividad a descripción) no hay demasiadas razones para detenerse a mitad de camino y permitir que la función F genere, simplemente, relaciones decisivas. Debemos exigir también que F «funcione» si las relaciones individuales son, solamente, decisivas¹². Al fin y al cabo la literatura, y la evidencia, sobre la no existencia de preferencias individuales consistentes es abundante (Armstrong [1], Luce [10], K. May [13]).

En la segunda parte demostraremos que no existe tal función si I, II y IV deben cumplirse.

II

Una relación R será llamada cuasitransitiva¹³ si la relación derivada P es transitiva, es decir:

iv) si $x P y$, $y P z$ entonces $x P z$; $x, y, z \in M$.

Una relación cuasitransitiva es decisiva; una relación puede ser decisiva, pero no cuasitransitiva o cuasitransitiva, pero no racional¹⁴.

\mathcal{A} : el conjunto de N -uplos de relaciones cuasitransitivas.

¹¹ Aunque en [12] se recupera el resultado de imposibilidad sustituyendo III por una condición más fuerte.

¹² O, como de hecho haremos, que la relación de preferencia es transitiva, pero la relación de indiferencia puede no serlo. Esta es una hipótesis más fuerte que decisividad (véase la segunda parte).

¹³ Esta es la terminología de Sen [17]. "Negativamente transitiva" sería mejor.

¹⁴ Ejemplos con tres alternativas:

decisiva, pero no cuasitransitiva: $x P y$, $y P z$, $x I z$;
cuasitransitiva, pero no racional: $x I y$, $x I z$, $x P z$.

Q): el conjunto de relaciones cuasitransitivas.

Como de costumbre supondremos que $M > 2$ y $N > 2$.

Proposición: No existe ninguna función $F: \mathcal{A} \rightarrow \mathcal{B}$, que cumpla las condiciones I, II y IV¹⁵ ($N > 2$).

Demostración: La demostraremos, sin pérdida de generalidad, para $N = 3$. Para el caso general reinterpretense todas las etapas de la demostración, suponiendo que a todos los miembros de N —excepto a los dos primeros— se asigna la misma relación que al tercero.

Como de costumbre, el argumento es por contradicción. Tómense tres alternativas arbitrarias $x, y, z \in M$. Por la Condición I podemos suponer, sin pérdida de generalidad, que M consta solamente de estos tres elementos. Supongamos que existe una función F que cumple las condiciones de la proposición. Entonces exhibiremos un N -uplo de relaciones individuales que deben necesariamente originar una relación cíclica (no decisiva) entre las tres alternativas, contradiciendo la definición de F .

La demostración procede en cinco etapas. La Condición I se utiliza de forma implícita constantemente. Compruébese que, en cada paso, las relaciones individuales postuladas son cuasitransitivas.

(A): Existe un par de alternativas $x, y \in M$ y un miembro $i \in N$ tal que si $x P_i y$ e $y P_j x$ para todo $j \neq i$, entonces $x R y$.

Esto ha sido demostrado en [12] (Lema 3). La demostración no necesita ninguna modificación. Supondremos $i = 1$. Obsérvese que (A) y la Condición IV implican:

(A.1) si $x P_1 y$ y $x I_2 y$ entonces $x P y$; independientemente de la relación del tercer miembro de N .

(B) Existe un par de alternativas $u, v \in M$ tal que si $u I_1 v$ para todo $i \in N$, entonces $u I v$ ¹⁶.

Sean a, b, c elementos (distintos) genéricos de M . Supóngase que $a I_1 b$, $b I_1 c$ y $c P_1 a$ para todo $i \in N$. Por la Condición II, $c P a$, y, teniendo en cuenta iii) (pág. 3) se sigue que:

¹⁵ Nótese que la Condición III no es necesaria. Tampoco la condición $N > 3$ usada en el Teorema 3 de [12].

¹⁶ Obsérvese que el enunciado es simétrico en u y v . Esto nos permite tolerar, en lo sucesivo, una cierta ambigüedad en si los pares de alternativas (ver más adelante) deben entenderse como pares ordenados o no, lo cual reduce el número de subcasos a considerar. Advirtamos que la demostración, para ser completamente rigurosa, debería ser más formal. Pero entonces el aparato formal (y las dificultades de lectura) no correspondería a la elementalidad del análisis sustantivo. Este es un defecto común en la literatura sobre este tema. En cualquier caso como consta en el texto, no creemos que haya posibilidad alguna de confusión.

- (1.a) si $u P b$ entonces $c \bar{R} b$.
 (1.b) si $b P c$ entonces $b \bar{R} a$.

Si el enunciado que debemos probar no es cierto, existe un par $u, v \in M$ tal que $u I_i v$ para todo $i \in \mathbb{N}$ genera $u P v$.

Sea w el tercer elemento de M . Tomando $u = a, v = b$ y $w = c$ en (1.a) obtenemos: $w I_i v$ para todo $i \in \mathbb{N}$ y $w R v$. Por la hipótesis de contradicción $w P v$. Tomando $b = w, c = v, a = u$ en (1.b), obtenemos: $w I_i u$ para todo $i \in \mathbb{N}$ y $w R u$. Por la misma razón que antes $w P u$. Aplicando de nuevo (1.a) —con $w = a, u = b$ — obtenemos $v R u$. Contradicción.

Por tanto (B) es cierto.

(C) Para cualquier $i \in \mathbb{N}$ y para cualquier par $s, t \in M$:

Si $s I_i t$, existe algún par de relaciones, \bar{R}_j, \bar{R}_g entre s, t ($g \neq j \neq i$) tal que (R_i, R_j, R_g) origina $s P t$.

Supóngase que no es éste el caso. Hay un par $s, t \in M$ para el que (C) no es cierto. Si (u, v) es el par cuya existencia se ha probado en (B), (u, v) , habida cuenta de la condición IV, satisface (C). Por tanto $(s, t) \neq (u, v)$. Deben distinguirse dos subcasos:

- a. $(s, t) = (u, w)$.
 b. $(s, t) = (v, w)$.

Puesto que todo es análogo, nos limitaremos a obtener una contradicción para el subcaso a. Tomemos $i = 1$ y consideremos la siguiente tabla:

T A B L A 1

1	2	3
$u I_1 w$	$u I_2 w$	$u P_3 w$
$v P_1 w$	$v P_2 w$	$v P_3 w$
$u I_1 v$	$u I_2 v$	$u P_3 v$

Por la hipótesis de contradicción tenemos que si $u I_1 w, u P_2 w$ y $u P_3 w$, entonces $w R u$; de esto y la Condición IV se sigue que de la tabla anterior obtenemos $w P u$. De la Condición II obtenemos $v P w$, y de (B) y la Condición IV, $u P v$. Hemos encontrado un ciclo y, por tanto, una contradicción.

(D) El par (x, y) de (A) puede considerarse distinto al par (u, v) de (B).

Es suficiente demostrar que si (A) se cumple para (x, y) se cumple para (x, z) . (u, v) será distinto a uno de los dos pares y, por tanto, podemos designar en la forma convenida al que más nos convenga.

Consideremos la tabla:

T A B L A 2

1	2	3
$x P_1 y$	$x I_2 y$	$y P_3 x$
$y P_1 z$	$y I_2 z$	$y P_3 z$
$x P_1 z$	$z P_2 x$	$z P_3 x$

De (A.1) se sigue $x P y$ y de (C) — más la Condición IV— $y P z$. Por tanto [(iii), página 3] $x R z$ que es precisamente el enunciado de (A) para (x, z) .

Estamos ya en condiciones de construir un conjunto de tres relaciones que originan un ciclo. En virtud de (D), $(u, v) \neq (x, y)$. Deben distinguirse dos subcasos:

- a. $(u, v) = (x, z)$.
- b. $(u, v) = (y, z)$.

Puesto que todo es análogo, nos limitamos a considerar el subcaso b.

De la tabla:

T A B L A 3

1	2	3
$x P_1 y$	$x I_2 y$	$y P_3 x$
$y I_1 z$	$y I_2 z$	$y P_3 z$
$z I_1 x$	$z P_2 x$	$z P_3 x$

se sigue:

- $x P y$ por (A.1).
- $y P z$ por (B) y la Condición IV.
- $z P x$ por (C) y la Condición IV.

Esto concluye la demostración.

Para $N = 2$ el sistema de voto mayoritario por pares cumple todas las condiciones. La Condición IV es bastante fuerte. Podría debilitarse ¹⁷ de la misma forma que en [12]. Pero la Proposición es ya exponente suficiente del punto en que queremos insistir: que el cambio de «racionalidad» a «decisividad» no resuelve de ninguna forma sustancial los problemas que la paradoja del voto generalizada —el «Teorema de Imposibilidad» de Arrow— plantea.

III

Aun cuando pudiera parecer ocioso vale la pena observar que la primera y más importante consecuencia del «Teorema de Imposibilidad» y resultados relacionados (como la proposición de la segunda parte) es que el planteamiento que los origina no puede servir como base para un estudio positivo del proceso de elección social. Recogemos, para concluir, algunas de las limitaciones que le han sido repetidamente señaladas y que por sí mismas apuntan a las cuestiones a tener en cuenta.

En primer lugar la Condición de Independencia de las Alternativas Irrelevantes es mucho más fuerte de lo que a primera vista parece. De hecho elimina la posibilidad de comportamiento estratégico de los individuales, y hace plausible el estudio del problema como el de reglas que actúan sobre preferencias individuales y no, simplemente, sobre preferencias *reveladas*. En la realidad no siempre la máxima conveniencia de un agente social es actuar conforme a sus verdaderas preferencias. Cada regla (incluida la elección por mayoría entre más de dos alternativas) induce formas de falseamiento o desestimamiento de las mismas.

En segundo lugar, difícilmente puede sostenerse que las reglas de elección social discriminen entre alternativas finales. A lo más, el campo de elección se limita a «programas» completos. Una teoría razonable debe explicar los compromisos y transacciones que esta situación origina.

¹⁷ Pero no suprimirse completamente. Véase el ejemplo de la nota 9.

(D) El par (x, y) de (A) puede considerarse distinto al par (u, v) de (B).

Es suficiente demostrar que si (A) se cumple para (x, y) se cumple para (x, z) . (u, v) será distinto a uno de los dos pares y, por tanto, podemos designar en la forma convenida al que más nos convenga.

Consideremos la tabla:

T A B L A 2

1	2	3
$x P_1 y$	$x I_1 y$	$y P_3 x$
$y P_2 z$	$y I_2 z$	$y P_2 z$
$x P_1 z$	$z P_2 x$	$z P_3 x$

De (A.1) se sigue $x P y$ y de (C) —más la Condición IV— $y P z$. Por tanto [(iii), página 3] $x R z$ que es precisamente el enunciado de (A) para (x, z) .

Estamos ya en condiciones de construir un conjunto de tres relaciones que originan un ciclo. En virtud de (D), $(u, v) \neq (x, y)$. Deben distinguirse dos subcasos:

- a. $(u, v) = (x, z)$.
- b. $(u, v) = (y, z)$.

Puesto que todo es análogo, nos limitamos a considerar el subcaso b.

De la tabla:

T A B L A 3

1	2	3
$x P_1 y$	$x I_1 y$	$y P_3 x$
$y I_1 z$	$y I_1 z$	$y P_2 z$
$z I_1 x$	$z P_2 x$	$z P_3 x$

se sigue:

- $x P y$ por (A.1).
- $y P z$ por (B) y la Condición IV.
- $z P x$ por (C) y la Condición IV.

Esto concluye la demostración.

Para $N = 2$ el sistema de voto mayoritario por pares cumple todas las condiciones. La Condición IV es bastante fuerte. Podría debilitarse¹⁷ de la misma forma que en [12]. Pero la Proposición es ya exponente suficiente del punto en que queremos insistir: que el cambio de «racionalidad» a «decisividad» no resuelve de ninguna forma sustancial los problemas que la paradoja del voto generalizada —el «Teorema de Imposibilidad» de Arrow— plantea.

III

Aun cuando pudiera parecer ocioso vale la pena observar que la primera y más importante consecuencia del «Teorema de Imposibilidad» y resultados relacionados (como la proposición de la segunda parte) es que el planteamiento que los origina no puede servir como base para un estudio positivo del proceso de elección social. Recogemos, para concluir, algunas de las limitaciones que le han sido repetidamente señaladas y que por sí mismas apuntan a las cuestiones a tener en cuenta.

En primer lugar la Condición de Independencia de las Alternativas Irrelevantes es mucho más fuerte de lo que a primera vista parece. De hecho elimina la posibilidad de comportamiento estratégico de los individuos, y hace plausible el estudio del problema como el de reglas que actúan sobre preferencias individuales y no, simplemente, sobre preferencias *reveladas*. En la realidad no siempre la máxima conveniencia de un agente social es actuar conforme a sus verdaderas preferencias. Cada regla (incluida la elección por mayoría entre más de dos alternativas) induce formas de falsamiento o desestimamiento de las mismas.

En segundo lugar, difícilmente puede sostenerse que las reglas de elección social discriminen entre alternativas finales. A lo más, el campo de elección se limita a «programas» completos. Una teoría razonable debe explicar los compromisos y transacciones que esta situación origina.

¹⁷ Pero no suprimirse completamente. Véase el ejemplo de la nota 9.

Véanse: Farquharson [7], Luce y Raiffa [11] (Capítulo 16), Wilson [19], y sus referencias para algunos análisis de estas situaciones desde el punto de vista de la teoría de los juegos¹⁸.

BIBLIOGRAFIA

- [1] ARMSTRONG, W.: "The determinateness of the utility function". *Economic Journal*, 67 (1939).
- [2] ARROW, K.: "Social Choice and Individual Values". Segunda edición, 1963 (primera, 1951). John Wiley and Sons, Inc.
- [3] BARBU I, M.: "Quelques aspects mathématiques de la décision rationnelle" *Les Temps Modernes*, 15 (1959).
- [4] BERGSON, A.: "A Reformulation of Certain Aspects of Welfare Economics" *Quarterly Journal of Economics* (1938).
- [5] BLACK, D.: "The Theory of Committees and Elections", 1958. Cambridge University Press.
- [6] CAMACHO, A.: "Social Decision Functions" (1971) (ciclostilado). Northwestern University.
- [7] FARQUHARSON, R.: "Theory of Voting", 1964. Yale University Press.
- [8] GIBBARD: "Intransitive Social Indifference and the Arrow Dilemma", 1969 (no publicado).
- [9] Guha, A.: "Neutrality, Monotonicity and the right of veto", 1970 a aparecer en *Econometrica*.
- [10] LUCE, D.: "Semiororders and a theory of utility discrimination", *Econometrica*, 24 (1956).
- [11] LUCE Y RAIFFA: "Games and Decisions", New York, John Wiley & Sons, Inc. (1957).
- [12] MAS COLELL, A. y H. SONNENSCHEN: "General Possibility Theorems for Group Decisions". *Review of Economic Studies*. Abril 1972.
- [13] MAY, K.: "A set of independent necessary and sufficient conditions for simple majority voting", *Econometrica*, 1952.
- [14] MURAKAMI, Y.: "Logic and Social Choice". New York, 1968. Dover Publications.
- [15] SAMUELSON, P. A.: *Foundations of Economic Analysis*, 1948, Harvard University Press.
- [16] SCHICK, F.: "Arrow's proof and the logic of preference", *Philosophy of Science*, 36 (1969).
- [17] SIN, A. K.: "Quasi-transitivity, Rational Choice and Collective Decisions". *Review of Economic Studies*, 36 (1969).
- [18] SIN, A. K.: "Collective Choice and Social Welfare", 1970, Holden-Day, San Francisco.
- [19] WILSON, R.: "Stable Coalition Proposals in Majority-Rule Voting", *Journal of Economic Theory*, 3 (1971).

¹⁸ Como es sabido, la Condición I pretende recoger el principio ordinalista tradicional en el análisis económico estático. En una interesante aportación Antonio Camacho [6] argumenta la necesidad de abandonarlo en el análisis teórico de la elección social. (Advertencia importante: no se confunda cardinalismo de las preferencias con comparabilidad interpersonal de las mismas.)