

La teoría de la elección social:
métodos de votación no manipulables

Jordi Massó

Departamento de Economía e Història Econòmica y CODE
Universitat Autònoma de Barcelona
08193, Bellaterra (Barcelona)
e-mail: jmasso@volcano.uab.es

Enero de 1996

El objetivo de esta nota es ilustrar, a través de un ejemplo, el tipo de investigación realizada por un grupo de economistas, que utilizando las matemáticas como lenguaje, se ha ido consolidando en el Departamento de Economía e Historia Económica de la Universitat Autònoma de Barcelona, alrededor de los temas de la teoría de la elección social, la teoría de los juegos y sus aplicaciones a la teoría económica. El ejemplo que he escogido es el de los métodos de votación no manipulables. A pesar de los abundantes resultados negativos de la teoría de la elección social, los procedimientos de votación son ampliamente utilizados como métodos de asignación y de toma de decisiones. Parece, pues, especialmente productivo colocarse en ámbitos suficientemente restringidos para que en ellos se produzcan resultados de posibilidad, y donde puedan plantearse, por tanto, cuestiones de diseño y selección de métodos alternativos. Veámoslo.

La teoría de la elección social trata sobre la toma de decisiones colectivas a partir de las preferencias de los individuos que conforman una sociedad. Consideremos un conjunto de alternativas sociales y una sociedad cuyos individuos tienen preferencias sobre dicho conjunto. Representemos estas preferencias por relaciones binarias sobre el conjunto de alternativas; tengamos presente que los individuos pueden tener opiniones distintas sobre las alternativas sociales. La teoría de la elección social estudia el proceso de agregación de las preferencias individuales en una preferencia social. Las decisiones colectivas, entonces, se tomarán a partir de la relación binaria social que se ha obtenido al agregar las

preferencias individuales. Más formalmente, dado un conjunto de alternativas sociales, una función de bienestar social asigna a cada estado de opinión una relación binaria. No es difícil darse cuenta de que, en general, existen infinidad de procesos de agregación; algunos parecen interesantes y otros lo parecen muy poco (por ejemplo, la función constante es muy poco interesante ya que la preferencia social no depende de las preferencias individuales).

Arrow [1] propone algunas propiedades que, en su opinión, los procesos de agregación social deberían cumplir:

(i) Función de bienestar social: si las preferencias individuales son preórdenes totales, entonces la preferencia social debería ser también un preorden total.

(ii) Dominio universal: cualquier preferencia individual es legítima.

(iii) Principio de Pareto: Si hay unanimidad en considerar una alternativa mejor que otra, entonces la preferencia social debería colocar a la alternativa mejor por delante de la peor.

(iv) Independencia de alternativas irrelevantes: La ordenación social entre dos alternativas sólo depende de las ordenaciones individuales de las dos alternativas y no de la forma en que éstas ordenan a otras.

El famoso teorema de imposibilidad de Arrow, Arrow [1], nos dice que las cuatro condiciones anteriores sólo son compatibles con procesos de agregación dictatoriales; a saber, si una función de bienestar social satisface las propiedades de dominio universal, pareto e independencia de alternativas irrelevantes, debe de ser dictatorial, es decir, las ordenaciones sociales son las ordenaciones de uno de los individuos fijado previamente.

Existe una amplia literatura generada por el resultado de imposibilidad de Arrow. Parte de esta literatura ha tratado de proponer conjuntos de propiedades alternativas al propuesto por Arrow, manteniendo el mismo marco conceptual determinado por el objeto "función de bienestar social", es decir, la propiedad (i). Otra parte de la literatura trata de salirse del contexto de Arrow. Nuestros intereses se sitúan en esta segunda parte. Una gran proporción de estas dos vertientes de la literatura tiene como objetivo la obtención de resultados positivos; esto es, la existencia de funciones de bienestar social no dictatoriales.

En muchos casos quizás, la dificultad en el proceso de agregación radica en determinar qué alternativas sitúa la preferencia social en tercer y cuarto lugar, por ejemplo. Se

pretende obtener una ordenación social para luego escoger una, y sólo una, de las alternativas. Un objetivo natural, entonces, no es plantearse obtener como resultado del proceso de agregación una ordenación completa de las alternativas, sino determinar directamente cuál de ellas será la elegida. Más formalmente, dado un conjunto de alternativas sociales, una función de decisión social asigna a cada estado de opinión una y sólo una de las alternativas sociales. La mayoría de los métodos de votación constituyen ejemplos de funciones de decisión social; en muchos casos se vota no tanto para determinar la ordenación entre las alternativas sino para elegir (quizás temporalmente) una y sólo una de ellas. Tampoco es difícil darse cuenta de que, en general, existen infinidad de funciones de decisión social; algunas de ellas pueden también parecer interesantes y otras pueden parecerlo muy poco (por ejemplo, la función constante también es una función de decisión social; la manipulación de los procesos de votación para escoger a un candidato previamente determinado es un ejemplo de función constante de decisión).

Un segundo aspecto no contemplado por Arrow (él mismo explícitamente lo reconoce), es el aspecto estratégico del problema. Si para tomar decisiones colectivas tenemos que basarnos en las opiniones individuales, deberemos de preguntar por ellas. Pero entonces, cada individuo, al darse cuenta de que las decisiones colectivas se basan en parte en sus opiniones particulares, podría intentar deformar su opinión (y de hecho estaría dispuesto a hacerlo) si con ello obtuviera decisiones sociales más favorables desde su verdadero punto de vista. Supongamos que, conociendo la función de decisión social que se va a utilizar, cada individuo fuera capaz de determinar que, independientemente de las preferencias de los demás, declarar preferencias falsas sólo puede perjudicarlo. Es decir, estamos suponiendo que declarar la verdad es una estrategia dominante para todos los individuos. Una función de decisión social con esta propiedad se llama no manipulable. Esta clase de funciones es interesante porque nos ahorra considerar, en el propio diseño de la función particular, los aspectos estratégicos del proceso de decisión social. La teoría de la implementación trata explícitamente sobre estos aspectos estratégicos del proceso de decisión, utilizando nociones más débiles de equilibrio.

El teorema de Gibbard y Satterthwaite, demostrado independientemente por Gibbard [2] y Satterthwaite [3], da una respuesta negativa a la posibilidad de considerar simultáneamente, y de forma interesante, los dos aspectos anteriormente mencionados. Más precisamente, una función de decisión social cuyo

rango tenga al menos tres alternativas es no manipulable si y sólo si es dictatorial (elige la mejor alternativa de un individuo preterminado).

Una hipótesis esencial en el teorema anterior es que, dado el conjunto de alternativas, los individuos pueden ordenarlo como les plazca; es decir, todas las posibles preferencias sobre el conjunto de alternativas son admisibles. De todos modos, pueden haber situaciones en las que, debido a que el conjunto de alternativas sociales tiene una estructura determinada, no tiene sentido esta hipótesis de dominio universal. Pensemos por ejemplo en la ordenación izquierda-derecha cuando las alternativas son partidos políticos, o en la ordenación de distintos proyectos legales sobre el aborto en función de los supuestos en los que el aborto es tratado como una práctica legal, o en la ordenación de las distintas temperaturas que puede tener una sala, etc. Cada una de estas ordenaciones (exclusivamente asociadas con las propiedades objetivas de las alternativas), podría ser unánimemente aceptada por todos los individuos. Supongamos, por lo tanto, que el conjunto de alternativas puede ser ordenado en una dimensión. Al mismo tiempo, podríamos pensar que los individuos no ordenarían de cualquier manera estas alternativas. La restricción de dominio más conocida es la de las preferencias unimodales (o de único "pico"). Una preferencia es unimodal si una vez fijada la mejor alternativa, al alejarnos en cualquiera de las dos direcciones encontramos alternativas menos preferidas. Las funciones de decisión no manipulables en el ámbito de las preferencias unimodales han sido ampliamente estudiadas en la literatura; véase, por ejemplo, Black [4] y Moulin [5]. Moulin demuestra que situados en este ámbito de dominio restringido (preferencias unimodales en una dimensión), existen funciones de decisión social no manipulables y de notable interés. Él mismo caracteriza a este conjunto como la familia de funciones que pueden ser representadas como un sistema del votante mediano.

Hay varias maneras alternativas de describir un sistema del votante mediano. La más intuitiva es la basada en los sistemas de coaliciones por la izquierda (o simétricamente por la derecha). Para simplificar supongamos que el conjunto de alternativas es finito y que lo tenemos ordenado de izquierda a derecha. Para cada subconjunto de votantes y para cada alternativa, un sistema de coaliciones por la izquierda nos indica si este conjunto de votantes (coalición) pertenece o no al sistema en esta alternativa concreta. Si pertenece, querrá decir que, dada una votación, la alternativa escogida por la función no podrá ser mayor (en el sentido de la ordenación inicial), siempre que las votaciones de

todos los miembros de la coalición sea menor o igual que la alternativa considerada. Para que esta interpretación sea consistente, se deben imponer algunas condiciones naturales a los conjuntos de coaliciones. En primer lugar, si una coalición está en el sistema para una alternativa determinada cualquier conjunto que contenga a la coalición también está en el sistema. En segundo lugar, si una coalición está en el sistema para una alternativa determinada también lo está para todas las alternativas mayores en la ordenación inicial. En tercer lugar, todas las coaliciones pertenecen al sistema de la alternativa mayor. Cualquier sistema de coaliciones por la izquierda sobre un conjunto de alternativas ordenado genera una función de decisión social a través del siguiente procedimiento: los votantes declaran sus alternativas preferidas, la alternativa escogida es la única con la propiedad de que el conjunto de individuos que votaron a la misma alternativa o por alternativas inferiores es una coalición del sistema para la alternativa en cuestión, mientras que el conjunto de individuos que votaron por alternativas estrictamente inferiores no está en el sistema. Informalmente, podemos pensar que la función escoge la alternativa más alta posible, dadas las habilidades de las coaliciones de "tirar" del resultado lo más abajo posible. Estas habilidades vienen expresadas por el sistema de coaliciones por la izquierda. A una función de elección social que escoja según este procedimiento se le llama un sistema del votante mediano.

Varios autores han extendido el concepto de unimodalidad a varias dimensiones con el objeto de estudiar funciones de decisión social en este ámbito más general. Por ejemplo Border y Jordan [6], Barberà, Sonnenschein y Zhou [7], Serizawa [8], Barberà, Gul y Stacchetti [9], etc. El interés de esta línea de investigación radica en que una gran variedad de problemas de decisión social tienen la propiedad de que la descripción completa de una alternativa social requiere varias dimensiones. Por ejemplo, una empresa que quiere contratar a nuevos trabajadores debe de escoger un conjunto de empleados de entre todos los posibles subconjuntos de solicitantes. En este caso, el conjunto de alternativas se puede identificar con los vértices del hipercubo, ya que cada subconjunto puede ser descrito por su función característica: uno si el elemento pertenece al subconjunto y cero si no pertenece. Otro ejemplo consiste en el problema de decisión de una institución pública que debe de escoger entre distintas opciones alternativas de varios proyectos públicos. En este caso podemos identificar el conjunto de alternativas sociales con una retícula en un espacio euclídeo. Nótese que en ambos casos hemos identificado en conjunto de alternativas como un producto

cartesiano de un espacio euclídeo.

La generalización natural de un sistema del votante mediano a contextos multidimensionales es el llamado sistema generalizado del votante mediano que consiste en escoger tantos sistemas de coaliciones por la izquierda como dimensiones, uno para cada una de ellas. Cada votante declarará su alternativa preferida que será proyectada en cada dimensión. Entonces, se aplicarán los correspondientes sistemas del votante mediano para determinar cada una de las componentes de la decisión colectiva. Estos mecanismos están basados en un proceso de toma de decisión enormemente descentralizada: se selecciona un valor para cada una de las características que describen a una alternativa, y entonces el resultado de la decisión colectiva es la alternativa definida conjuntamente por estos valores seleccionados independientemente.

Los artículos de Barberà, Sonnenschein y Zhou [7], para el caso del hipercubo, y Barberà, Gul y Stacchetti [9], para el caso general, caracterizan los sistemas generalizados del votante mediano como las únicas funciones de decisión social no manipulables en estos contextos multidimensionales. Sus resultados utilizan la hipótesis adicional de que la función de decisión social es exhaustiva.

En nuestro trabajo (Barberà, Massó y Neme [10]) introducimos explícitamente el hecho de que en muchas decisiones sociales las restricciones económicas provocan que el conjunto factible de alternativas no sea un producto cartesiano. Por ejemplo, conjuntos generados por restricciones presupuestarias caen plenamente dentro del ámbito de aplicación de nuestros resultados, pero en cambio son excluidos por la hipótesis de exhaustividad de la función de decisión.

Lo que es específico de nuestro enfoque es la incorporación explícita de diferentes clases de restricciones en la toma de decisiones colectivas. Nos concentramos en situaciones donde los individuos tienen preferencias unimodales (generalizadas a varias dimensiones) e investigamos en este contexto el conjunto de funciones de decisión social no manipulables. En primer lugar demostramos que son sistemas de votación (es decir, las funciones no manipulables sólo pueden depender de las mejores alternativas declaradas por los individuos), descomponibles por dimensiones y que pueden expresarse como sistemas generalizados del votante mediano. Obtenemos también una caracterización de las funciones que preservan factibilidad y damos un resultado de maximalidad de la restricción de dominio. Finalmente, vemos que el teorema de

Gibbard y Satterthwaite, el resultado más general en el área, es un Corolario inmediato de nuestra caracterización.

Resumiendo, nuestro proyecto consiste en explorar las posibilidades abiertas por una forma concreta de extender la noción de unimodalidad a conjuntos multidimensionales de alternativas. Esto es especialmente productivo porque dicha restricción permite también describir contextos en los que pueden evitarse manipulaciones estratégicas de las reglas de decisión colectivas incluso cuando se tenga que considerar restricciones de factibilidad.

Referencias

[1] K. Arrow, *Social Choice and Individual Values*, Wiley, New York, (1963).

[2] A. Gibbard, "Manipulation of Voting Schemes: A General Result", *Econometrica* 41, pp. 587-601, (1973).

[3] M.A. Satterthwaite, "Strategy-proofness and Arrow's Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions", *Journal of Economic Theory* 10, pp. 187-217, (1975).

[4] D. Black, "On the Rationality of Public Decision Making", *Journal of Political Economy* 56, pp. 23-34, (1948).

[5] H. Moulin, "On Strategy Proofness and Single-Peakedness", *Public Choice* 35, pp. 437-455, (1980).

[6] K. Border y J. Jordan, "Straightforward Elections, Unanimity and Phantom Agents", *Review of Economic Studies* 50, pp. 153-170, (1983).

[7] S. Barberà, H. Sonnenschein y L. Zhou, "Voting by Committees", *Econometrica* 59, pp. 595-609, (1991).

[8] S. Serizawa, "Cross-Shaped Preferences and Voting by Committees", de próxima aparición en *Journal of Economic Theory*.

[9] S. Barberà, F. Gul y E. Stacchetti, "Generalized Median Voter Schemes and Committees", *Journal of Economic Theory* 61, pp. 262-289, (1993).

[10] S. Barberà, J. Massó y A. Neme, "Voting under Constraints", mimeo, Universitat Autònoma de Barcelona, (1995).