A24. Economía del bienestar (III). Las funciones de bienestar social. Teoría de la elección colectiva. El teorema de imposibilidad de Arrow y desarrollos posteriores

1. La función de bienestar social

Ordenar los estados sociales de forma completa y transitiva obliga a sopesar lo que ganan unos frente a lo que pierden otros, de modo que la economía del bienestar no puede eludir las comparaciones interpersonales de utilidad.

1.1. Criterio original (Bergson, 1938; Samuelson, 1947)

Origen: Bergson (1938) introduce la **función de bienestar social** (FBS) para hacer explícitas las disyuntivas entre individuos, y Samuelson (1947) la incorpora al núcleo de la economía del bienestar, sin comprometerse con una forma concreta.

Idea: la FBS representa un código de valores colectivos y permite elegir, entre los infinitos óptimos de Pareto, el que maximiza el bienestar social.

1.1.1. Definición y propiedades

Definición: es una función definida sobre las utilidades de los H individuos que refleja las preferencias sociales:

$$W = W(\mathbf{u}), \qquad \mathbf{u} = (u^1, \dots, u^H)$$

Propiedades:

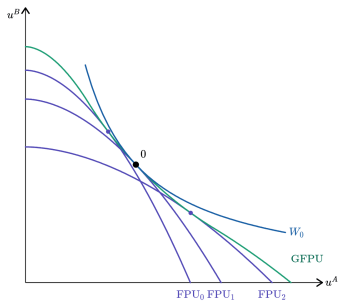
1. *Continuidad y diferenciabilidad*: garantizan un máximo caracterizable por CPO, lo que no ocurre con la *maximin*, no diferenciable en el vértice.
2. *Transformaciones*: exige utilidades comparables entre personas; la utilitarista, cardinales con unidades comparables, y la *maximin*, solo niveles comparables. La FBS resultante es ordinal. Cualquier transformación monótona creciente de W representa, por tanto, los mismos valores colectivos.
3. *Inclinación*: es creciente en cada utilidad, $\partial W/\partial u^h > 0$ (benevolencia), y las mejoras paretianas aumentan el bienestar social. La *maximin* solo es no decreciente, porque una mejora paretiana que no llegue al peor situado deja igual el bienestar social.
4. *Forma*: si la aversión a la desigualdad es deseable, es cuasicóncava, con curvas de indiferencia social convexas. A diferencia de la concavidad, la cuasiconcavidad se conserva ante transformaciones monótonas.
5. *Simetría*: si se desea el anonimato, intercambiar las utilidades de dos individuos no altera el bienestar. Con la cuasiconcavidad, una distribución más igualitaria con la misma suma nunca es peor.

Hereda, además, las propiedades paretianas: independencia del proceso, individualismo, no paternalismo y benevolencia.

1.1.2. Razonamiento

Curvas de indiferencia social (CIS): son las curvas de nivel de la FBS en el plano de utilidades, decrecientes, y su curvatura refleja la aversión a la desigualdad. En la función original son débilmente convexas y sin forma uniforme. La tasa a la que la sociedad sacrifica el bienestar de unos por el de otros decrece conforme la distribución se hace más desigual a favor de estos, pero no a un ritmo fijo.

Punto de felicidad: el *optimum optimorum* es la tangencia entre la CIS más alejada y la gran frontera de posibilidades de utilidad (GFPU). La eficiencia paretiana es necesaria, pero no suficiente, porque un punto ineficiente, dentro de la frontera, puede dar más bienestar que uno eficiente. El punto de felicidad determina la asignación Pareto-óptima que lo genera.



Gran frontera de posibilidades de utilidad y punto de felicidad.

1.1.3. Conclusiones

La FBS resuelve la indeterminación con un juicio de valor explícito y transparente y fundamenta la redistribución. Pero admite cualquier juicio sin decir cuál adoptar, porque no hay método para construirla y su existencia se postula axiomáticamente. El óptimo refleja así los valores de quien la especifica.

1.2. Formas de la FBS: criterios ético-lógicos

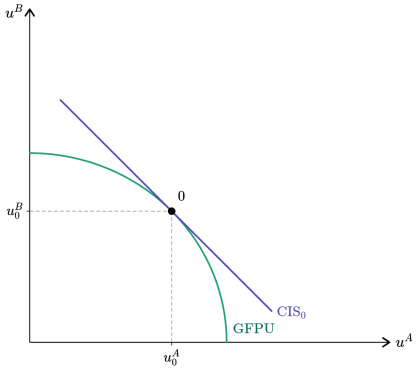
Las formas se ordenan por su aversión a la desigualdad: nula en el utilitarismo, intermedia en Bernoulli-Nash e infinita en el *maximin*.

1.2.1. Utilitarista (Bentham; Harsanyi, 1955)

Idea: el bienestar social es la suma de las utilidades, y solo importa el total, no su reparto.

Velo de la ignorancia: Harsanyi lo justifica porque un individuo que no sabe qué lugar ocupará y maximiza la utilidad esperada de von Neumann-Morgenstern, con posiciones equiprobables, elige la mayor utilidad media.

Planteamiento: con $W = \sum_h u^h(x^h)$, al repartir un bien el óptimo iguala las utilidades marginales, $u'^A = u'^B$. Si difirieran, trasladar una unidad a quien la tiene mayor elevaría la suma. Es un máximo si las utilidades son cóncavas. Las CIS son rectas de pendiente -1, porque las utilidades no se ponderan, y el óptimo es su tangencia con la GFPU.



Criterio utilitarista: CIS rectas de pendiente -1 tangentes a la GFPU en el punto 0.

Respeta el individualismo y el anonimato, pero es indiferente a la distribución: aumenta la utilidad de quien tiene mayor utilidad marginal a costa de quien la tiene menor y es compatible con grandes desigualdades.

1.2.2. Intermedio (Bernoulli; Nash, 1950)

Idea: el bienestar social es el producto de las utilidades y valora su nivel y su igualdad. Sin parámetros, su aversión a la desigualdad es unitaria en la familia de Atkinson.

Planteamiento: con $W = \prod_h u^h(x^h)$, equivalente a $\sum_h \ln u^h$, el óptimo iguala las utilidades marginales divididas por la utilidad, $u'^A/u^A = u'^B/u^B$: el peso social de cada uno es inverso a su utilidad, y la utilidad marginal de quien está peor pesa más. Con utilidades positivas y cóncavas es un máximo. Las CIS son convexas y el óptimo es su tangencia con la GFPU.

Enlaza con la solución de negociación de Nash, pero no permite graduar la aversión y exige un cero bien definido en cada utilidad. Si se desplaza el origen de alguna, el orden social cambia.

1.2.3. *Maximin* (Rawls, 1971)

Idea: en *A Theory of Justice*, el bienestar social es el del más desfavorecido. Tras el velo de la ignorancia, sin atribuir probabilidades, los individuos razonan según el *maximin* y acuerdan dos principios:

- el *principio de libertad*: el conjunto más amplio de libertades básicas compatible con las de los demás;
- el *segundo principio*: las desigualdades solo son legítimas si benefician al más desfavorecido, el *principio de diferencia*, con igualdad equitativa de oportunidades;
- y una *prioridad lexicográfica* del primero sobre el segundo.

Planteamiento: la FBS es de tipo Leontief, $W = \min_h\{u^h(x^h)\}$. No es diferenciable en el vértice y el óptimo exige $u^A = u^B$, la igualdad absoluta de utilidades. No procede de una tangencia, y su suficiencia se sigue de la cuasiconcavidad de W. Las CIS tienen forma de L, y el óptimo está donde la GFPU corta la bisectriz.

Inspira el igualitarismo liberal, pero equivale a una aversión infinita al riesgo, pues ninguna mejora de la mayoría compensa un empeoramiento mínimo del peor situado.

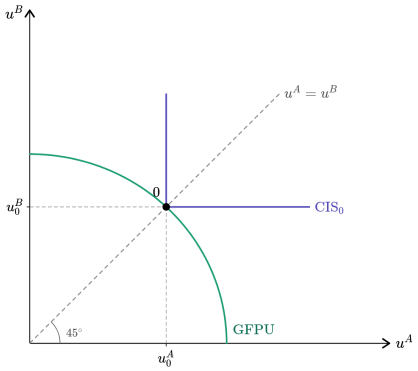
1.2.4. Familia de Atkinson (1970)

Idea: toda medida de desigualdad lleva implícita una FBS. La **renta equivalente igualmente distribuida**, y_e, es la que, repartida por igual, da el mismo bienestar que la distribución real. No supera la media, y el índice mide la pérdida relativa:

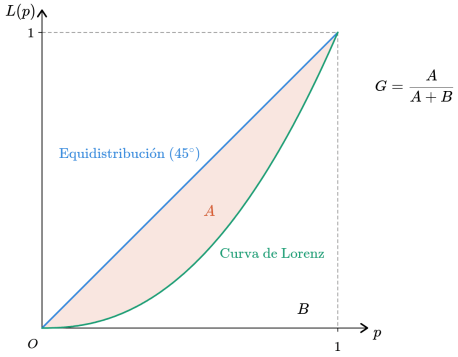
$$A_\varepsilon = 1 - \frac{y_e}{\bar{y}}, \qquad y_e = \left[\frac{1}{H}\sum_h (y^h)^{1-\varepsilon}\right]^{1/(1-\varepsilon)}$$

donde $\bar{y}$ es la renta media.

El parámetro $\varepsilon \geq 0$ es la aversión a la desigualdad: con $\varepsilon = 0$ se obtiene el utilitarismo; con $\varepsilon = 1$, la media geométrica de Bernoulli-Nash, y con $\varepsilon \to \infty$, el *maximin*. Cuanto mayor es ε, mayor es la pérdida que



Criterio *maximin* (Rawls): CIS en ángulo recto y óptimo donde la GFPU corta la bisectriz.



Curva de Lorenz e índice de Gini.

se atribuye a una misma desigualdad. A diferencia del Gini, permite elegir cuánto pesa la desigualdad y es muy usado en la evaluación de las políticas redistributivas. Pero hace explícito el juicio sin resolverlo, porque ε se fija fuera del modelo y, si las curvas de Lorenz se cruzan, la ordenación puede invertirse según el parámetro. Supone, además, una FBS aditiva y simétrica en la que la utilidad se identifica con la renta.

1.3. Más allá del bienestarismo

1.3.1. Capacidades básicas (Sen, 1980)

Idea: Sen sostiene que no importan los bienes primarios, como en Rawls, ni la satisfacción, como en el utilitarismo, sino las **capacidades**: lo que las personas pueden *hacer y ser*, como estar bien nutrido, sano o educado. Dos personas con los mismos bienes pueden tener capacidades muy distintas, por diferencias físicas, sociales o de entorno, y dos con la misma satisfacción pueden vivir en condiciones muy distintas, porque las preferencias se adaptan a las circunstancias. La FBS debe definirse, por tanto, sobre las capacidades. Fundamenta el índice de desarrollo humano, pero las capacidades no se observan directamente, lo que obliga a aproximarlas con indicadores de funcionamientos.

1.3.2. Sin envidia (Foley, 1967; Varian, 1974)

Idea: una asignación es **sin envidia** si nadie prefiere la cesta de otro, y *justa* si además es eficiente. No exige comparar utilidades. Con dotaciones iguales, el equilibrio competitivo es justo. Nadie envidia a nadie, porque todos eligieron con la misma renta, y es eficiente por el primer teorema. Muestra así que equidad y eficiencia no son incompatibles y da contenido concreto al segundo teorema. Con producción y habilidades distintas, puede no alcanzarse, porque el más productivo puede envidiar el ocio del menos productivo.

1.4. Medición de la desigualdad y la pobreza

1.4.1. Desigualdad unidimensional

Curva de Lorenz (1905): representa la proporción acumulada de renta frente a la de población, ordenada de menor a mayor renta. La igualdad perfecta es la bisectriz. Si dos curvas no se cruzan y la renta media es igual, la más cercana a la bisectriz es preferida por cualquier FBS simétrica y cóncava. Si se cruzan, depende de la FBS. Con las rentas ordenadas de menor a mayor, cada punto de la curva es:

$$L\left(\frac{k}{H}\right) = \frac{\sum_{h=1}^{k} y^h}{\sum_{h=1}^{H} y^h}, \qquad k = 0, 1, \dots, H$$

Índice de Gini (1912): es el doble del área entre la curva y la bisectriz, o la mitad de la diferencia media entre rentas en proporción a la media. Vale 0 con igualdad perfecta:

$$G = \frac{1}{2H^2\bar{y}} \sum_{h=1}^{H} \sum_{h'=1}^{H} \left| y^h - y^{h'} \right|, \qquad G \in [0, 1]$$

Vale $1 - 1/H$ si una persona lo recibe todo. Es el más usado, pero es más sensible al centro que a las colas, y su FBS implícita pondera cada renta según su posición, sin parámetro que la sociedad pueda elegir.

Índice de Theil (1967): basado en la entropía, compara la renta de cada individuo con la media y vale entre 0 y $\ln H$:

$$T = \frac{1}{H} \sum_{h=1}^{H} \frac{y^h}{\bar{y}} \ln \frac{y^h}{\bar{y}}, \qquad T \in [0, \ln H]$$

Es **descomponible** (la desigualdad total se separa en la de dentro y la de entre grupos) y permite atribuirla a regiones, sexos o niveles educativos:

$$T = \sum_{j=1}^{J} s_j T_j + \sum_{j=1}^{J} s_j \ln \frac{\bar{y}_j}{\bar{y}}$$

donde s_j es la parte de la renta total que recibe el grupo j, $\bar{y}_j$ su renta media y T_j su índice.

1.4.2. Pobreza (Sen, 1976)

Idea: medir la pobreza exige fijar una línea z, absoluta o relativa, como el 60 % de la mediana en la UE; contar a los pobres, la tasa de incidencia H_p, y medir cuánto les falta. La tasa de incidencia ignora la intensidad y la desigualdad entre los pobres. Así, una transferencia de un pobre a otro menos pobre que lo saque de la pobreza la reduce, aunque empeore al más pobre. La *monotonicidad* exige que la pobreza aumente si baja la renta de un pobre, y la *transferencia*, que aumente si se transfiere renta de un pobre a alguien más rico.

Índice de Sen: exige *monotonicidad* y *transferencia*, y pondera la brecha de cada pobre según su puesto:

$$S = H_p \left[I + (1 - I)\, G_p \right]$$

donde I es la intensidad de la pobreza y G_p el Gini de los pobres. Convierte la medición de la pobreza en un problema de bienestar social. No es descomponible por grupos, lo que motivó la familia FGT (Foster, Greer y Thorbecke, 1984), y toma como dada la línea, que es un juicio de valor previo.

1.4.3. Medición multidimensional

Si importan las capacidades, la pobreza se mide en varias dimensiones: el índice de desarrollo humano (PNUD, 1990), el índice de pobreza multidimensional (Alkire y Foster, 2011) y la tasa AROPE de la UE, que combina la pobreza relativa de renta, la privación material y social severa y la muy baja intensidad laboral del hogar. Qué dimensiones incluir y cómo ponderarlas son juicios de valor.

En suma, la FBS da el óptimo social con valores explícitos, y todo índice lleva implícita una valoración. Pero el óptimo es tecnocrático: depende de una forma funcional que ningún método permite construir.

2. Elección colectiva: el teorema de Arrow

La **teoría de la elección colectiva** se pregunta si los valores colectivos pueden obtenerse de las preferencias individuales mediante un procedimiento que las respete.

2.1. La regla de elección social

Distinción: la FBS de Bergson-Samuelson es una valoración, que asigna un número a cada estado dado un perfil de utilidades cardinales y comparables. La **regla de elección social** (RES), o *constitución*, es un procedimiento. Asocia a cada perfil posible de ordenaciones individuales ordinales y no comparables, $(R^1, \dots, R^H)$, una ordenación social completa y transitiva:

$$R = f(R^1, \dots, R^H)$$

La mayoría, la unanimidad o la regla de Borda son ejemplos. La RES es más amplia que la FBS, porque ha de funcionar para cualquier perfil, y no solo para uno dado. El conjunto de estados sociales tiene al menos tres alternativas, porque con dos el problema es trivial. Cada ordenación individual es completa y transitiva, pero dice qué se prefiere y no cuánto.

Arrow busca una RES que respete unas exigencias mínimas de racionalidad y democracia, y concluye que no existe. Las salidas a la imposibilidad consisten en relajarlas.

2.2. Axiomas de Arrow

Para Arrow, una RES es racional y democrática si cumple cinco axiomas.

1. *Universalidad*: se aplica a cualquier perfil de preferencias, de forma análoga a la completitud.
2. *Unanimidad* (Pareto débil): si todos prefieren A a B, la sociedad también.
3. *Ausencia de dictadores*: nadie determina la elección al margen del resto.
4. *Transitividad*: si $A \succ B$ y $B \succ C$, entonces $A \succ C$, lo que evita los ciclos.

5. *Independencia de alternativas irrelevantes* (IAI): para ordenar A y B basta saber cómo ordena cada uno A y B. Excluye la cardinalización y la manipulación añadiendo o quitando alternativas.

2.3. Teorema de la imposibilidad (Arrow, 1951b)

Teorema: en *Social Choice and Individual Values* (1951b), Arrow prueba que, con un número finito de individuos y tres o más alternativas, **ninguna RES satisface a la vez los cinco axiomas**.

Demostración (Sen, 1970a): un grupo es *decisivo* sobre un par si, cuando todos sus miembros prefieren una alternativa, la sociedad la prefiere.

1. *Expansión*: por Pareto, la sociedad entera es decisiva. Por la IAI, la universalidad, Pareto y la transitividad, un conjunto decisivo sobre un par lo es sobre todos.
2. *Contracción*: todo conjunto decisivo con más de un miembro se divide en dos, y con un perfil sobre tres alternativas, por transitividad, una de las partes resulta decisiva.

Repitiendo, se llega a un conjunto de un solo individuo (un dictador). La universalidad permite construir los perfiles necesarios, la IAI impide usar la intensidad para bloquear la expansión y la transitividad obliga a contraer. Con solo ordenaciones, sin intensidades ni comparaciones, ninguna regla es a la vez coherente y respetuosa con cada individuo, y elegir una RES obliga a abandonar al menos un axioma. Las salidas de la sección 3 son otras tantas formas de pagar ese precio.

2.3.1. Paradoja liberal (Sen, 1970b)

Idea: Sen muestra que basta con Pareto débil y un **liberalismo mínimo**, según el cual cada uno de dos individuos es decisivo sobre algún par de su esfera privada. Con dominio universal, las preferencias *entrometidas* sobre la esfera ajena generan ciclos. El ejemplo canónico es el de un mojigato, un lascivo y *El amante de Lady Chatterley*.

El conflicto entre eficiencia y derechos es lógico: se resuelve excluyendo las preferencias entrometidas o tratando los derechos como restricciones sobre el proceso (Nozick, 1974). Y el criterio de Pareto, aparentemente neutral, no lo es, porque su aplicación irrestricta puede violar la esfera privada. La paradoja introduce los derechos en la economía del bienestar. Su relevancia práctica depende, sin embargo, de cuánto pesen las preferencias entrometidas.

2.3.2. Conclusiones

Aspectos positivos: generaliza a cualquier regla la incoherencia colectiva que Condorcet mostró con un ejemplo, y su método axiomático se extendió a la manipulabilidad y a la implementación.

Limitaciones: la IAI excluye la intensidad de las preferencias, que la regla de Borda recoge, y el dominio universal obliga a ordenar perfiles extremos, cuya exclusión permite a la mayoría simple una ordenación transitiva.

2.4. Rechazo de Arrow (Little, 1952; Bergson, 1954)

Idea: Little y Bergson niegan que la economía del bienestar busque una regla mecánica que sume valores. Su objeto son los juicios de bienestar social. Para Little, esa máquina no garantiza expresar lo que la sociedad valora, y para Bergson la economía del bienestar asesora al individuo, y la regla de Arrow, al funcionario que decide en nombre de la colectividad.

La crítica distingue la fase *deliberativa*, donde se forman los juicios y opera la FBS, de la *decisiva*, donde se agregan y opera la RES. Arrow aceptó la distinción, pero replicó que la cuestión central es qué políticas debe adoptar la sociedad. La crítica fija el alcance del teorema, pero no lo desactiva, pues toda sociedad que decide necesita una regla.

2.5. Conclusiones

En suma, los valores que la FBS daba por supuestos no pueden obtenerse de ordenaciones ordinales con una regla que cumpla los cinco axiomas. El teorema funda la teoría de la elección social, y la sección 3 estudia sus salidas.

3. Desarrollos posteriores: las salidas a la imposibilidad

Las reglas de votación, las que usan las sociedades, renuncian a algún axioma. Las de Borda y Condorcet son anteriores a Arrow, pero se leen a su luz. Las demás salidas amplían la información, estudian la manipulación y la implementación o pasan a la negociación.

3.1. Decisividad y universalidad

3.1.1. Mayoría óptima (Buchanan y Tullock, 1962)

Origen: Wicksell (1896) propone la unanimidad como regla fiscal, la única que garantiza que nadie resulte perjudicado. Buchanan y Tullock (1962) se preguntan qué mayoría exigir, y fundan la teoría económica de las constituciones.

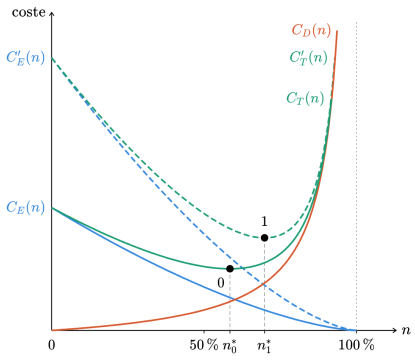
Idea: la **mayoría óptima** minimiza la suma de dos costes opuestos:

$$C_T(n) = C_D(n) + C_E(n)$$

- Los *costes de decisión*, de llegar a un acuerdo, crecen con la mayoría exigida y tienden a infinito con la unanimidad.
- Los *costes externos*, que soportan los perjudicados, decrecen y se anulan con la unanimidad.

La suma tiene forma de U y su mínimo es la mayoría óptima. Las reglas observadas van de la unanimidad, como en la fiscalidad del Consejo de la UE, a la mayoría cualificada, como la doble mayoría del 55 % de los Estados y el 65 % de la población en el Consejo de la UE o tres quintos de los miembros con el 85 % del poder de voto para reformar el convenio del FMI, que da a EE. UU. un veto de hecho, y a las mayorías absoluta y simple, como en el Consejo de Gobierno del BCE. La unanimidad es paretiana y exige pocos juicios de valor, pero favorece el *statu quo* y la *tiranía de la minoría*.

Estática comparativa: la mayoría óptima crece con los costes externos potenciales, como en las reformas constitucionales, y disminuye con los costes de decisión, en decisiones frecuentes y de bajo alcance. En la figura, los costes externos pasan a $C'_E(n)$, los totales a $C'_T(n)$ y la mayoría óptima, de n_0^* a n_1^*.



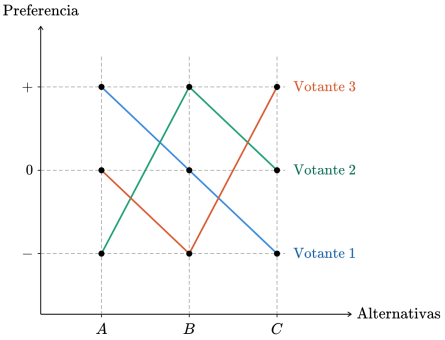
Mayoría óptima (Buchanan y Tullock, 1962): mínimo de los costes totales; crece con los costes externos (trazo discontinuo).

Teorema de May (1952): con dos alternativas, la mayoría simple es la única regla decisiva, anónima, neutral y con respuesta positiva (si una alternativa gana o empata y un votante cambia a su favor, gana). Explica la posición privilegiada de la mayoría simple, y sus problemas aparecen con más alternativas.

La mayoría óptima explica que convivan mayorías distintas en una misma constitución, aunque los costes son difíciles de medir y es más una guía cualitativa que una regla.

3.1.2. Mayoría simple: paradoja de Condorcet (1785)

Idea: la mayoría simple por pares puede dar una ordenación social intransitiva aunque las individuales sean transitivas. Con tres votantes, $A \succ B \succ C$, $B \succ C \succ A$ y $C \succ A \succ B$, A vence a B, B a C y C a A: no hay equilibrio. El problema nace del tercer votante, con dos modas.



Paradoja de Condorcet: el votante 3, con dos máximos, genera el ciclo.

Extensiones: las salidas a la paradoja son preferencias idénticas, que hacen innecesaria la regla; excluir lo ya votado, lo que permite la **manipulación de la agenda**, porque gana lo que entra en la última votación, o exigir preferencias unimodales, renunciando a la universalidad.

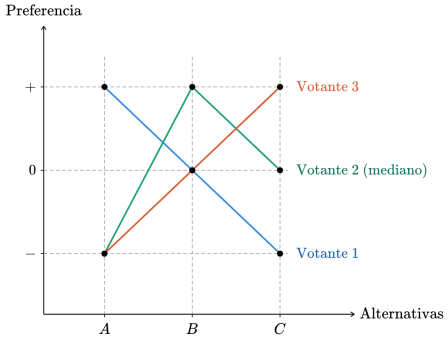
Separa la racionalidad individual de la colectiva y anticipa a Arrow, aunque es un resultado de posibilidad y no de frecuencia: los ciclos solo aparecen con ciertos perfiles y desaparecen cuando las preferencias son unimodales.

3.1.3. Unimodalidad y votante mediano (Black, 1948)

Teorema del votante mediano: si las alternativas se ordenan en una dimensión y las preferencias son **unimodales**, la mayoría elige el óptimo del votante mediano, que deja a cada lado el mismo número de votantes y es un ganador de Condorcet. Se pierde universalidad, pero solo se descartan perfiles extremistas.

Razonamiento: cualquier alternativa situada a la izquierda del óptimo del mediano pierde frente a él con los votos del mediano y de todos los situados a su derecha, y simétricamente las de la derecha. En la figura, B vence a A con los votos de 2 y 3 y a C con los de 1 y 2.

Se considera preferible perder universalidad a perder transitividad. El teorema explica la convergencia de los partidos al centro y fundamenta



Preferencias unimodales: gana B, la alternativa preferida por el votante mediano.

el modelo de Downs. Pero, con bienes públicos y reparto igualitario del coste, la mayoría iguala la valoración del mediano a RMT/H, y solo es eficiente si coincide con la media. En otro caso, hay infraprovisión o sobreprovisión.

3.1.4. Downs-Hotelling (1957)

Idea: Downs (1957) aplica el modelo de localización de Hotelling a la competencia electoral. Con información perfecta, dos partidos que maximizan votos y una sola dimensión ideológica, ambos se desplazan al votante mediano, y este equilibrio de Nash, el *principio de diferenciación mínima*, no tiene por qué ser eficiente. Los votantes, con preferencias unimodales sobre la ideología, votan al partido más cercano. Si un partido propone un programa, el otro puede acercarse más al mediano y acaparar votantes, y el primero reacciona acercándose también, hasta que ambos coinciden con él.

Si las preferencias se desplazan, ambos partidos siguen al mediano, y un cambio en los extremos que no lo altere no modifica el equilibrio. Explica la convergencia al centro en los países bipartidistas. Pero con más partidos aparecen el voto útil, las coaliciones y el extremismo. Con costes de información, los partidos necesitan recursos para sondeos y campañas, lo que da poder a los grupos de presión, y los votantes caen en la abstención racional, porque su voto apenas altera el resultado, la persuasión por carisma y las ilusiones fiscal y de deuda.

3.2. Independencia de alternativas irrelevantes

La mayoría ignora la **intensidad de las preferencias**. Renunciar a la IAI le da entrada.

3.2.1. *Log-rolling* (Buchanan y Tullock, 1962)

Idea: con el intercambio de votos, cada uno apoya la propuesta ajena a cambio de apoyo para la suya, y quien tiene mucho interés en algo lo consigue cediendo en lo que le importa poco. Exige poder pactar antes de votar y que aprobar una propuesta no impida aprobar otra.

Razonamiento: entre tres votantes, el 1 y el 2 se benefician cada uno de un proyecto pagado por los tres y con beneficios totales mayores que su coste. Votados por separado, ambos fracasan. El pacto entre 1 y 2 los aprueba si se votan en un único paquete. Pero falla con muchos votantes, puede explotar a la minoría y llevar a la sobreprovisión. Negociar sale entonces demasiado caro, y la coalición ganadora no está obligada a compensar a los perdedores.

3.2.2. Regla de Borda (1770)

Idea: asigna n puntos a la alternativa preferida, $n-1$ a la segunda, y así sucesivamente. Usa la ordenación completa, evita los ciclos y favorece las alternativas de consenso, pero viola la IAI, porque la puntuación depende de cuántas otras alternativas haya y dónde se sitúen. Genera, además, más empates y permite el voto estratégico (relegar al último puesto a la rival de la opción favorita).

3.3. Base informacional (Sen, 1970a)

Idea: Sen reinterpreta el teorema como un resultado sobre la información, ordinal y no comparable, que usa la regla. Si se admiten utilidades cardinales y comparables, la agregación deja de ser imposible.

Teorema de agregación de Harsanyi (1955): si las preferencias individuales y sociales cumplen la utilidad esperada y la regla respeta Pareto, la FBS es una suma ponderada de utilidades, con pesos iguales si hay anonimato. El utilitarismo se deriva así de la racionalidad bajo incertidumbre, sin postularlo.

D'Aspremont y Gevers (1977): con comparabilidad de niveles, los axiomas de Arrow más anonimato y equidad de Hammond conducen al *leximin*; con comparabilidad de unidades, al utilitarismo; con plena comparabilidad, elegir es un juicio de valor. La comparabilidad de niveles permite decir quién está peor, pero no cuánto, y la de unidades, cuánto gana uno y pierde otro, pero no quién está mejor.

Con información ordinal, restringir el dominio o relajar la transitividad social a la cuasitransitividad permite reglas no dictatoriales, a costa de oligarquías o de indecisión.

Los criterios ético-lógicos aparecen así como consecuencias de la información que se acepta usar. Pero el problema se traslada a justificar las comparaciones interpersonales, sin base empírica aceptada, y los axiomas adicionales, como la equidad de Hammond, son en sí mismos un juicio ético.

3.4. Manipulación e implementación

3.4.1. Manipulación: Gibbard-Satterthwaite (1973)

Teorema: todo mecanismo no dictatorial, con dominio universal y al menos tres alternativas elegibles, es **manipulable**: en algún perfil a alguien le conviene mentir. Equivale al de Arrow, y la salida es restringir el dominio o debilitar el equilibrio. Con preferencias cuasilineales y transferencias, la lógica de Clarke-Groves-Vickrey hace de la verdad la estrategia dominante. El teorema es el origen del diseño de mecanismos y convierte el comportamiento estratégico en un rasgo general que el diseño institucional ha de tener en cuenta, aunque es un resultado de existencia, porque manipular exige conocer las preferencias ajenas.

3.4.2. Implementación (Maskin, 1977)

Idea: Maskin, Nobel en 2007 con Hurwicz y Myerson, plantea la pregunta inversa a la de Arrow. Una regla es implementable en equilibrio de Nash si existe un juego cuyos equilibrios coinciden con lo que prescribe. La condición necesaria es la **monotonicidad de Maskin**. Con tres o más agentes, junto con la ausencia de veto, es suficiente. Si una alternativa elegida no empeora para nadie al cambiar el perfil, ha de seguir eligiéndose: el mecanismo solo reacciona a cambios que los agentes tengan interés en revelar. Fundamenta los mecanismos de revelación y la regulación con información asimétrica, aunque supone que los agentes conocen las preferencias de los demás. Además, sus mecanismos son complejos y poco realistas, y muchas reglas atractivas que eligen una sola alternativa no cumplen la monotonicidad.

3.5. Negociación y juegos cooperativos

3.5.1. Negociación axiomática (Nash, 1950)

Idea: Nash demuestra que cuatro axiomas, eficiencia, simetría, invariancia ante transformaciones afines e IAI, determinan una solución única: maximizar el producto de las ganancias sobre el punto de desacuerdo d:

$$\max_{(u^A,u^B)\in S} (u^A - d^A)(u^B - d^B)$$

donde S es el conjunto de pares de utilidades alcanzables mediante acuerdo y d lo que cada parte obtiene si no lo hay. La compacidad y la convexidad de S y la estricta cuasiconcavidad del producto garantizan que el máximo existe y es único.

Equivale a la FBS de Bernoulli-Nash con origen en el desacuerdo, y, frente a Arrow, obtiene una solución única porque su problema tiene más estructura y renuncia al dominio universal. Es la referencia de la negociación salarial y recibió un fundamento no cooperativo en el modelo de ofertas alternas. Pero no describe el proceso, su axioma de IAI es el más discutido y depende del punto de desacuerdo.

3.5.2. Juegos cooperativos

Núcleo (Gillies, 1959): los repartos que ninguna coalición puede mejorar. Es un concepto de estabilidad, puede ser vacío o contener muchos repartos y, en una economía de intercambio replicada, se reduce a los equilibrios competitivos. **Valor de Shapley** (1953): la media de la contribución marginal de cada jugador en todos los órdenes posibles. Es un concepto de equidad, el único que cumple eficiencia, simetría, tratamiento nulo y aditividad. Existe siempre y es único, pero puede quedar fuera del núcleo. Se aplica al reparto de costes de un bien público o al poder de voto en los organismos internacionales, con el índice de Shapley-Shubik. Ambos conceptos suponen, sin embargo, acuerdos vinculantes dentro de cada coalición y no describen la negociación que conduce a ellos.

3.6. Conclusiones

En suma, cada regla de votación renuncia a algún axioma, y ninguna no dictatorial con dominio universal escapa a la manipulación. La información comparable recupera reglas no dictatoriales, y la negociación obtiene soluciones únicas porque su problema tiene más estructura que el de Arrow. El análisis coste-beneficio y los fallos del sector público se estudian en los temas A20 y A23.