3.A.24: ECONOMÍA DEL BIENESTAR (III). LAS FUNCIONES DE BIENESTAR SOCIAL. TEORÍA DE LA ELECCIÓN COLECTIVA. EL TEOREMA DE LA IMPOSIBILIDAD DE ARROW Y DESARROLLOS POSTERIORES.

Con el cambio de temario, a partir de la convocatoria de 2023 este tema pasará a ser:
3.A.24: Economía del bienestar (III). Las funciones de bienestar social. Teoría de la elección colectiva. El teorema de imposibilidad de Arrow y desarrollos posteriores.

De este modo, con lo escrito en este documento este tema estaría **actualizado**.

A.24. Economía del bienestar (III). Las funciones de bienestar social. Teoría de la elección colectiva. El teorema de imposibilidad de Arrow y desarrollos posteriores

Título anterior	A.24. Economía del bienestar (III). Las funciones de bienestar social. Teoría de la elección colectiva. El teorema de imposibilidad de Arrow y desarrollos posteriores
Motivación del cambio	Sin cambios.
Propuesta de contenido /estructura	I. Breve repaso histórico: antecedentes a la Economía del Bienestar moderna I.I. Utilitarismo: cardinalidad y separabilidad aditiva, comparaciones interpersonales, decrecimiento de la utilidad marginal de la renta I.II. Robbins y la Economía paretiana II. Problemática general de la Teoría de la Elección Social: agregación/agrupación de preferencias II.I. Construcción axiomática de un funcional de bienestar social II.II. Teoremas de la Imposibilidad: Arrow y salidas. Relajamiento axiomático, especial referencia a la universalidad (Black 1947) II.III. Las cuestiones asociadas a la independencia de las alternativas irrelevantes: comparaciones interpersonales, no-manipulabilidad de agenda (Gibbard y Maskin) III. Problemática específica: consideraciones éticas y juicios de valor III.I. Herramienta fundamental: las funciones de bienestar social (construcción y valoración) III.II. Justicia bajo incertidumbre: neoutilitarismo y neocontractualismo III.III. Igualdad: funciones de Atkinson e índices de desigualdad basados en el equivalente igualitariamente distribuido

En general, este tema está muy relacionado con algunos temas del ejercicio 4.B. Se recomienda mirar estos temas juntos.

INTRODUCCIÓN

- ***Enganche***:
  - ALFRED MARSHALL, en sus *Principios de Economía* (1890) define la economía como *la ciencia de la vida diaria en lo que respecta a las acciones humanas tomadas para alcanzar un nivel máximo de bienestar*.
    - Esta definición nos muestra cómo uno de los principios subyacentes a la reflexión económica, pero particularmente enfatizado en la teoría neoclásica, es el del **individualismo metodológico**[1]. Se contempla el objeto de la teoría como una *realidad social compuesta de individuos que se interrelacionan en economías descentralizadas*.
  - En su objetivo fundamental de comprender y predecir el funcionamiento de los mercados, la **microeconomía** examina el comportamiento de 2 agentes fundamentales: *consumidores* y *productores*[2].

[1] El *individualismo metodológico* es un método ampliamente utilizado en las ciencias sociales. Sostiene que todos los fenómenos sociales —estructura y cambios— son en principio explicables por elementos individuales, es decir, por las propiedades de los individuos, como pueden ser sus metas, sus creencias y sus acciones. Sus defensores lo ven como una filosofía-método destinada a la explicación y comprensión amplia de la evolución de toda la sociedad como el agregado de las decisiones de los particulares. En principio es un reduccionismo, es decir, una reducción de la explicación de todas las grandes entidades con referencias en las más pequeñas.

[2] No hay que olvidar que la microeconomía contemporánea contempla esta separación estricta entre consumidores y productores como "una hipersimplificación del proceso por el que los bienes se compran y se consumen" (EKELUND y HÉBERT, 2013). Ejemplos que muestran el desdibujado de esta frontera son las "tecnologías del consumo", es decir, la aplicación de la teoría de la producción a las decisiones de consumo, como son el enfoque de características de KEVIN LANCASTER, la economía doméstica de GARY BECKER, la producción doméstica de REUBEN GRONAU o la economía de la información de GEORGE J. STIGLER (la información sobre los bienes de consumo, como bien económico o costoso, obliga a un proceso de búsqueda que debe combinarse con el bien de consumo físico).
Además, la microeconomía también estudia a otros agentes como las instituciones financieras o el Estado.

- La **economía del bienestar** está ligada a este enfoque normativo.
  - Concretamente, las **funciones** de la economía del bienestar son[3]:
    a. *Proporcionar instrumentos para valorar la deseabilidad social* de los estados alternativos, caracterizados por una asignación de recursos y una distribución de la renta. En principio, valorar esos estados alternativos requiere hacer juicios de valor explícitos sobre los mismos.
    b. *Proporcionar normas de política económica* que permitan maximizar el bienestar social (alcanzar el estado o estados realizables socialmente más preferidos).

- ***Relevancia***:
  - La comparación de diferentes alternativas es algo vital, ya que nos permite:
    - Por un lado, valorar si el resultado al que conduce un sistema económico es apropiado o mejorable.
    - Por otro lado, valorar si una política económica que cambie la situación inicial es deseable.
  - Como afirmó ATKINSON, el gran teórico de la desigualdad del siglo XX, la ciencia económica no existe sólo para describir el comportamiento humano y satisfacer la curiosidad y la vanidad de los economistas, sino para emitir recomendaciones y diseñar y valorar políticas que contribuyan a mejorar la vida de los ciudadanos[4].
    - Por ello, la economía del bienestar forma parte del "corazón" de la ciencia económica y como tal, debe ocupar un lugar preeminente en la formación de un economista y más aún de un *policy-maker*.

- ***Contextualización***:
  - *Desde un punto de vista histórico*,
    - En un sentido amplio, podría considerarse que todos los economistas, desde el momento en que han formulado valoraciones sobre los determinantes del bienestar social, son "economistas del bienestar":
      - Los mercantilistas, por ejemplo, relacionaban el bienestar social con la acumulación de metales preciosos en el país.
      - ADAM SMITH (1776), relacionaba implícitamente el bienestar social con el nivel total de producción (para así alejarnos del estado estacionario).
      - Ya en el siglo XIX, JEREMY BENTHAM consideraba que el concepto de bienestar está asociado al de utilidad (*utilitarismo*), por lo que consideraba el bienestar como la suma de las utilidades individuales.
      - Los marginalistas (JEVONS, MENGER y WALRAS), en la década de 1870, también consideraban que los 'bienestares individuales' se podían medir agregando las funciones de utilidad, pues concebían una *función de utilidad cardinal*. Así, proponen utilizar el concepto de *utilidad marginal decreciente* para sugerir que el bienestar social alcanzaría un máximo si el ingreso fuese distribuido equitativamente entre todos los miembros de la sociedad.
    - No obstante, el primer economista del bienestar propiamente dicho, se considera que ha sido ALFRED MARSHALL (1890) quien formula una concepción de bienestar en el que este es susceptible de medición en términos monetarios (*excedente del consumidor*[5]), permitiendo las comparaciones interpersonales de bienestar.

---

[3] Siguiendo a ANNA KOUTSOYIANNIS (*"Modern Microeconomics"*), las tareas de la teoría económica del bienestar son:
a) Demostrar que en el estado actual $W < W^*$, y
b) Sugerir las formas de elevar W hacia W^*.

[4] «*El principal motivo del análisis económico es contribuir a la mejora social*»
ARTHUR CECIL PIGOU (*Economics of Welfare*, 1920)

[5] Este concepto fue formalizado por MARSHALL (1890). Sin embargo, el concepto ya había sido propuesto anteriormente por el autor proto-marginalista JULES DUPUIT en su obra *De la Mesure de l´Utilité des Travaux Publics* (1884) [ver tema 3.A.9].

- Algunos años después VILFREDO PARETO, sucesor de WALRAS en la Escuela de Lausanne, rechaza esta concepción cardinal de la utilidad[6] y propone en su obra *Manual de Política Económica* (1906) un nuevo criterio para valorar una asignación de los recursos, el **criterio de Pareto** o criterio de optimalidad paretiana, que se convertirá en el criterio de referencia en la literatura al basarse en pocos juicios de valor que serán generalmente aceptados[7]. Según este criterio, un estado es eficiente si ningún agente puede mejorar sin que empeore otro.
  - La **importancia** del criterio de Pareto radica en que:
    a. Es un criterio *ordinal* que permite evitar comparaciones interpersonales de utilidad;
    b. Está basado en *pocos juicios de valor que son generalmente aceptados*; y
    c. *Permite valorar el sistema competitivo como mecanismo de asignación de los recursos de la economía*.
  - Sin embargo, a pesar de su popularidad, el criterio de Pareto *no está exento de limitaciones* (el principal inconveniente es que el criterio de Pareto resulta incompleto, en el sentido de que no permite comparar entre óptimos). Como consecuencia de ello, surgen una serie de *criterios alternativos*[8]:
    - Para resolver su problema de incompletitud, en las décadas de 1930 y 1940, surgen 2 soluciones:
      → Los **criterios de compensación** de la mano de autores como KALDOR, HICKS, SCITOVSKY (*línea neoparetiana* de la London School of Economics) y
      → La caracterización de una **Función de Bienestar Social** de BERGSON y SAMUELSON (*línea neoutilitarista* de Harvard).
    - Por otro lado, en la década de 1950, LIPSEY y LANCASTER formulan la **teoría del second best**, relacionada con el estudio de qué ocurre si se incumple alguna de las condiciones que define un óptimo de Pareto.
    - Además, también en los años 50, surge una revolución en la economía del bienestar de la mano de KENNETH ARROW[9] y su **Teorema de Imposibilidad**.

---

[6] Antes de la aportación de PARETO se produce una fusión entre *utilitarismo* y *marginalismo*. Sin embargo, son 2 cosas totalmente diferentes:
- *Marginalistas* (JEVONS, WALRAS): Consideraban que los 'bienestares individuales' se podían medir, pues concebían una función de utilidad cardinal.
- *Utilitaristas* (BENTHAM): Consideraban que el bienestar social era igual a la suma de los 'bienestares individuales'.

Partiendo de esas ideas, estos autores hallaban el bienestar social **sumando** las utilidades **cardinales**, valorando los efectos sobre el bienestar de la política económica en términos de si la suma total de utilidades **aumentaba o disminuía** (se asume que las comparaciones interpersonales de utilidad son posibles).

El problema es que estos enfoques se basaban en unos **supuestos muy fuertes y restrictivos**.

Quien rompe con esta alianza desgraciada es PARETO. Esto se ve reflejado en la introducción del concepto *Ofelimidad* (https://es.wikipedia.org/wiki/Ofelimidad):
- PARETO prefiere el término *ofelimidad* al término más común de utilidad para enfatizar que no siempre lo que el individuo desea (es decir, lo que es 'ventajoso') también es útil, en el sentido de favorable. La *ofelimidad* se diferencia de la utilidad por su carácter de subjetividad. En otras palabras, la *ofelimidad* representa la utilidad desde el punto de vista de la intensidad de la preferencia de un individuo, no de la comunidad.
- IRVING FISHER propuso reemplazar la *ofelimidad* (y, por lo tanto, la utilidad como se la interpreta comúnmente) con el término *wantability*.

[7] El criterio fundamental que ha surgido en el campo de la economía del bienestar para valorar el bienestar social es el de eficiencia económica, entendida como mejor uso posible de los recursos limitados. En la literatura económica han surgido varios enfoques para valorar la eficiencia económica, pero entre ellos destaca el criterio de Pareto.

[8] PARETO y ARROW son 2 puntos de inflexión en la historia de la economía del bienestar:
- Inicialmente los máximos exponentes son MARSHALL y PIGOU.
- VILFREDO PARETO supone un punto de inflexión que da lugar a la *New Welfare Economics*.
- KENNETH ARROW supone un punto de inflexión que da lugar a la *New New Welfare Economics*.

[9] KENNETH ARROW fue galardonado con el Premio Nobel de Economía en 1972 «Por sus contribuciones a la teoría del equilibrio económico y del bienestar».

- <u>*En esta exposición*</u>, nos vamos a centrar en una rama muy concreta de la economía del bienestar: la *teoría de la elección social*.
  - En concreto, nos centraremos en el criterio de la Función de Bienestar Social, la teoría de la elección social de ARROW y los desarrollos posteriores que surgen como respuesta a esta.

- ***Problemática (Preguntas clave)***:
  - ¿En qué medida ARROW cambia el programa de investigación y qué impacto tiene su Teorema de la Imposibilidad en la literatura?
  - ¿Cuáles son los desarrollos posteriores/vías de escape al Teorema de la Imposibilidad?

- ***Estructura***:

El orden cronológico sería que tras el criterio de Pareto, para resolver el problema de incompletitud surgen los criterios de compensación y la FBS de BERGSON y SAMUELSON. Posteriormente, ARROW cambia el enfoque y enuncia el TIA. Frente al TIA hay distintas reacciones: 1) Consideraciones éticas y juicios de valor (HARSANYI y RAWLS) 2) Levantamiento de los axiomas.

Hacer disclaimer de que pese a que BERGSON y SAMUELSON son anteriores a ARROW, *la FBS va a ser abordada en el segundo apartado*

1. TEORÍA DE LA ELECCIÓN SOCIAL Y TEOREMA DE LA IMPOSIBILIDAD DE ARROW. KENNETH ARROW Y SU OBRA *"SOCIAL CHOICE AND INDIVIDUAL VALUES"* (1951).

1.1. Contexto en el que surge el Teorema de Imposibilidad de Arrow (1951) – Incompletitud del criterio de PARETO

1.1.1. Criterio de PARETO

Definición

Requisitos de óptimo de Pareto y obtención de la Gran Frontera de Posibilidades de Utilidad

Teoremas Fundamentales de la Economía del Bienestar

Valoración

1.1.2. Criterios adicionales para resolver el problema de incompletitud

1.2. Teorema de la Imposibilidad de Arrow (1951)

Idea

Función de Elección Social, preferencias individuales y su agregación a una ordenación social

Definición de Función de Elección Social

Preferencias individuales (inputs) y ordenación social (output) con sus respectivas axiomáticas

Teorema de Imposibilidad de ARROW

1.3. Valoración y escapatorias al Teorema de la Imposibilidad de Arrow

2. ESCAPATORIAS AL TEOREMA DE LA IMPOSIBILIDAD DE ARROW (I). RECHAZO DE LA METODOLOGÍA DE ARROW, CONSIDERACIONES ÉTICAS Y JUICIOS DE VALOR. FUNCIONES DE BIENESTAR SOCIAL.

2.1. Idea principal: Rechazo de la metodología de ARROW y búsqueda de otros supuestos o conjuntos de axiomas alternativos

Críticas de LITTLE y BERGSON

Similitudes y diferencias entre el enfoque de ARROW y el enfoque de BERGSON y SAMUELSON

2.2. Línea neoutilitarista: Función de Bienestar Social

2.2.1. Idea

2.2.2. Función de Bienestar Social – BERGSON (1938) y SAMUELSON (1947)

Idea (definición de Función de Bienestar Social)

Índices de utilidad individual (inputs) y nivel de bienestar social (output)

2.2.3. Desarrollo

Propiedades deseables de la Función de Bienestar Social según SAMUELSON

Problema del planificador

Implicaciones

2.2.4. Valoración

2.3. Valores éticos subyacentes (elección de una forma funcional)

2.3.1. Idea

2.3.2. Consideraciones éticas: la filosofía utilitarista y el neocontractualismo

2.4. Justificación de criterios equitativos y de justicia

2.4.1. ANTHONY ATKINSON

2.4.2. Enfoque empírico de RICHARD THALER

Juego del ultimátum

Juego del dictador

2.4.3. Crítica a las Funciones de Bienestar Social de AMARTYA SEN

2.5. Valoración de las Funciones de Bienestar Social

3. ESCAPATORIAS AL TEOREMA DE LA IMPOSIBILIDAD DE ARROW (II): LEVANTAMIENTO DE LOS AXIOMAS, REGLAS DE VOTACIÓN Y PROBLEMAS DE INFORMACIÓN

3.1. Idea principal: Aceptación de la metodología de ARROW

3.2. Levantamiento de los axiomas

3.2.1. Renuncia al axioma de universalidad

Idea

Paradoja de Condorcet (intransitividad de las preferencias sociales cuando las preferencias individuales son multimodales)

Paradoja de Condorcet (intransitividad de las preferencias sociales)

Posibles soluciones al problema de intransitividad de las preferencias sociales

Preferencias unimodales

3.2.2. Renuncia al axioma de independencia de las alternativas irrelevantes

Utilidades cardinales

Reglas de votación

Regla de Borda y el problema de manipulación estratégica

Otros sistemas de votación con manipulación estratégica (Regla de segunda vuelta, Ley D'Hont)

Teorema de GIBBARD-SATTERTHWAITE

Evitar comportamientos estratégicos mediante el diseño de mecanismos de incentivos compatibles

Ejemplo: Mecanismo de Clarke y Groves para la provisión de un bien público

Otras reglas de elección diferentes a la mayoría simple

1. TEORÍA DE LA ELECCIÓN SOCIAL Y TEOREMA DE LA IMPOSIBILIDAD DE ARROW. KENNETH ARROW Y SU OBRA "*SOCIAL CHOICE AND INDIVIDUAL VALUES*" (1951).

1.1. Contexto en el que surge el Teorema de Imposibilidad de Arrow (1951) – Incompletitud del criterio de PARETO

1.1.1. Criterio de PARETO

Muy brevemente en esta exposición (de palabra) ya que este apartado se desarrolla en el tema 3.A.22.

Definición

- Como decíamos, un estado será óptimo de Pareto si no es posible que un agente mejore sin que empeore otro[10,11].

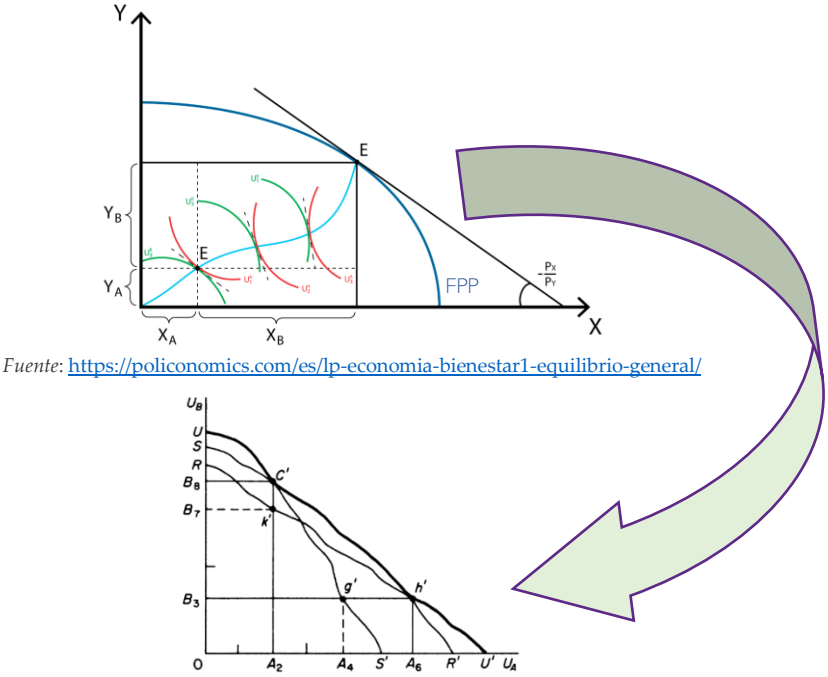
Requisitos de óptimo de Pareto y obtención de la Gran Frontera de Posibilidades de Utilidad

- Para que una asignación de recursos sea un **óptimo de Pareto**, es necesario que la asignación sea:
  - *i)* *Eficiente en la producción*: La eficiencia en la producción ocurrirá allí donde las *RMST* de producir ambos bienes se igualen (es decir, allí donde las isocuantas sean tangentes). Si no se cumple esta igualdad, podría aumentarse la producción de un bien sin disminuir la del otro, o aumentar la de ambos a la vez. Se puede representar a partir de una caja de Edgeworth que nos dará como resultado una curva de contrato en la que la producción es óptima y se podrá representar en una *Frontera de Posibilidades de Producción* como la de la Imagen 1.
  - *ii)* *Eficiente en el intercambio*: La eficiencia en el intercambio ocurrirá allí donde las *RMS* de consumir ambos bienes se igualen (es decir, allí donde las curvas de indiferencia sean tangentes). Si no se cumple esta igualdad, podría aumentarse la utilidad de un individuo sin disminuir la del otro, o aumentar la de ambos a la vez. En el gráfico de la Imagen 1 representado dentro de la caja de Edgeworth.
  - *iii)* *Eficiente globalmente*: En el óptimo global se igualan todas las valoraciones. Puesto que la *RMT* muestra la tasa a la cual un bien puede transformarse en otro, desde el punto de vista de la producción, y la *RMS* muestra la tasa a la cual los consumidores están dispuestos a cambiar un bien por el otro, el sistema no está en equilibrio a menos que todas las razones sean iguales. En el gráfico, en el punto *E* en la caja de Edgeworth las curvas de indiferencia tienen la misma pendiente que la *FPP* en el punto en el que se forma esa caja de Edgeworth. Los puntos que cumplan ese requisito (y no todos los que forman la línea de contrato) serán óptimos de Pareto. Podremos representar todos los puntos que cumplan estos criterios de optimalidad en el plano de las utilidades de los individuos mediante la *Gran Frontera de Posibilidades de Utilidad (GFPU)*.

[10] Un concepto ligado estrechamente al de óptimo de Pareto es el de *Pareto superioridad*: un estado *A* será Pareto superior a *B* si todos los individuos mejoran con el paso de *B* a *A*, o al menos uno mejora y el resto permanece igual. Es decir, es aquella nueva asignación *A* a la que nadie se opondría a cambiar desde la asignación *B*.
- Que una asignación sea Pareto superior a otra no significa que la primera sea un óptimo de Pareto.
- Al mismo tiempo, una asignación será un óptimo de Pareto si no existe ninguna otra asignación Pareto superior a ella.

[11] El principio de Pareto se basa en **4 juicios de valor** generalmente aceptados (según GRAVELLE):
1. *Independencia del proceso*: No importan los mecanismos utilizados para asignar los recursos y distribuir la renta. Es decir, no importa si una situación es fruto de la acción libre de los individuos o de decisiones forzadas.
2. *Individualismo*: El único aspecto relevante de una asignación es su efecto sobre las utilidades de los individuos individuales[11].
3. *No paternalismo*: El hecho de que las asignaciones se evalúen tomando como referencia las funciones de utilidad de los individuos implica que se considera que *el individuo es el mejor juez de su propio bienestar*[11].
4. *Benevolencia*: Un aumento en la utilidad de un individuo se considera, *ceteris paribus*, beneficioso desde el punto de vista social. Esto implica que el criterio de Pareto no emite juicios de valor sobre las desigualdades.

IMAGEN 1.– *Obtención de la Gran Frontera de Posibilidades de Producción*



Fuente: https://policonomics.com/es/lp-economia-bienestar1-equilibrio-general/

Fuente: A.Koutsoyiannis *"Modern Microeconomics". Ver pp. 544 y ss. para explicación obtención de la curva*

- En resumen, cada punto de la FPP da lugar a una Frontera de Posibilidades de Utilidad. La *envolvente* de estas Fronteras de Posibilidades de Utilidad da lugar a la *Gran Frontera de Posibilidades de Utilidad (GFPU)*.
- Al ser capaz de arrojar infinitos óptimos no comparables, el criterio de Pareto **puede no ser suficiente** para determinar qué asignación maximiza el bienestar social, que es el objetivo último de la economía del bienestar.
  - Así, la eficiencia u optimalidad paretiana global es una condición necesaria pero no suficiente para garantizar la maximización del bienestar social[12]. Es decir, una situación puede ser óptima en el sentido de Pareto pero no maximizar el bienestar social. Sin embargo, la maximización del bienestar sólo se alcanza en una situación óptima de Pareto.

Teoremas Fundamentales de la Economía del Bienestar

De este análisis es posible derivar los **2 teoremas fundamentales de la economía del bienestar:**

1. *Primer Teorema del Fundamental de la Economía del Bienestar (1TFEB)*: si un mercado funciona de forma competitiva y no presenta ningún tipo de imperfección (fallo de mercado), la asignación de mercado de dicho equilibrio será Pareto eficiente.
2. *Segundo Teorema del Fundamental de la Economía del Bienestar (2TFEB)*: cualquier asignación eficiente en sentido de Pareto puede alcanzarse como un equilibrio competitivo a través de una redistribución de las dotaciones iniciales.
   Una consecuencia directa del 2TFEB es que un *planificador benevolente* podría llevar a cabo *transferencias de suma fija* (no distorsionantes) y dejar posteriormente trabajar a los *mecanismos de mercado* para alcanzar el óptimo de Pareto global que *maximiza el bienestar social* ("punto de felicidad" o *"bliss point"*), esto es, el punto en el que la curva de indiferencia derivada de la función de bienestar social es tangente a la Gran Frontera de Posibilidades de Utilidad.
   Por lo tanto, el 2TFEB implica que es posible *separar las dimensiones eficiencia y equidad*. Así, los defensores de la equidad utilizarán el 2TFEB para defender que, aún en el caso de que no haya

[12] Suponiendo monotonía de la Función de Bienestar Social, es decir, que al aumentar la utilidad de algún individuo necesariamente aumente (o al menos no disminuya) el bienestar social.

imperfecciones, el Estado debe intervenir para mejorar la equidad, sin que por ello se tenga que resentir la eficiencia.

- El contraargumento que se ha esgrimido es que es prácticamente imposible introducir un sistema de impuestos y transferencias óptimo (i.e. que no distorsione las decisiones y el comportamiento de los agentes económicos), por lo que pueden surgir pérdidas de eficiencia y aparecer con fuerza el trade-off equidad vs. eficiencia.

- Por lo tanto, vemos cómo el **1TFEB** sirve de *defensa del libre mercado* (limitando el papel del Estado a la corrección de imperfecciones), mientras que el **2TFEB** abre la puerta a la *intervención pública ex-ante* para que, mediante la redistribución (y, posteriormente, la acción libre de los agentes en el mercado), se alcance un determinado equilibrio deseado.

Valoración

- Valorando el criterio de PARETO nos damos cuenta de que es un criterio que tiene 3 grandes **ventajas**:
  1. Es un criterio *ordinal* que permite evitar comparaciones interpersonales de utilidad y
  2. Está basado en *pocos juicios de valor que son generalmente aceptados*.
  3. *Permite valorar el sistema competitivo como mecanismo de asignación de los recursos de la economía*.

- Sin embargo, ha sido **criticado** por distintos motivos:
  1. La *escuela austriaca critica el criterio de Pareto por ser un concepto estático*. Lo verdaderamente relevante en el análisis económico no es hallar la asignación óptima de la FPP sino aplicar sistemáticamente el criterio de eficiencia dinámica que es aquel en el que se fija la capacidad del sistema para mover continuamente hacia la derecha la FPP.
  2. Debido al derecho a veto, el criterio de Pareto *favorece el statu quo*, por lo que la política económica puede resultar inoperante. Lo que implica la necesidad de unanimidad en las decisiones (esto es lo que BAUMOL denominaba *"tiranía de las minorías"*).
  3. El criterio de Pareto también ha sido criticado por introducir como juicio de valor *independencia del proceso*. Para el criterio de Pareto es igualmente válida, *ceteris paribus*, una decisión que ha sido adoptada democráticamente que otra impuesta de manera dictatorial.
  4. *Al evitar las comparaciones interpersonales de utilidad no tiene en cuenta aspectos redistributivos*. De hecho, la solución de esquina (un individuo consume todo y el otro nada) siempre es un óptimo en el consumo (pues para que el que no tiene nada mejore, el que tiene todo tiene que empeorar)[13]. El criterio de Pareto solo se fija en la eficiencia, dejando de lado la equidad. Como señaló AMARTYA SEN[14] es posible que lleguemos a una solución eficiente en la que una pequeña minoría del país viviera en la abundancia y la gran mayoría viviera en la más absoluta miseria.
  5. Pero posiblemente, el principal inconveniente del criterio de Pareto es que constituye un criterio *incompleto*, en el sentido de que no ofrece una ordenación completa de los estados económicos alternativos y existen infinitos óptimos de Pareto que no serían comparables entre sí.

[13] En respuesta a esto, se ha propuesto el concepto de "asignaciones justas en el sentido de FOLEY", que serían aquellas que, además de Pareto eficientes, son *equitativas* (entendiendo "equitativas" como aquellas asignaciones *libres de envidia*, es decir, en las que ningún agente prefiere la asignación que tiene el otro).

[14] AMARTYA SEN fue galardonado con el Premio Nobel de Economía en 1998 «Por sus contribuciones al análisis del bienestar económico». En 2021 también fue galardonado con el Premio Princesa de Asturias de Ciencias Sociales.

1.1.2. Criterios adicionales para resolver el problema de incompletitud

- Para determinar qué óptimo de Pareto maximiza el bienestar social, es necesario algún **criterio adicional**. En este sentido, surgió durante los años 30 y 40 una nueva línea de pensamiento, la **Nueva Economía del Bienestar**, que trató de superar esta indefinición a través de 2 vías:
  a. *Línea neoparetiana*, que introduce los criterios de compensación con el objetivo de hacer comparables los distintos óptimos paretianos y que la sociedad pudiera ordenarlos en el espacio de utilidades mediante un sistema de compensaciones potenciales entre los grupos ganadores y perdedores involucrados (KALDOR, HICKS, SCITOVSKY, SAMUELSON y LITTLE).
  b. *Línea neoutilitarista*, que introduce el criterio de la función de bienestar social (SAMUELSON y BERGSON).

Aquí mencionar en 3 palabras lo que hacen BERGSON y SAMUELSON para superar el problema de incompletitud y dibujar una curva de indiferencia social en la GFPU que hayas dibujado en la pizarra. Pero la FBS se va a analizar luego en el segundo apartado. Es importante entender que las FBS nacen antes del TIA, pero que luego fueron retomadas posteriormente por RAWLS y HARSANYI, por lo que es necesario hacer un *disclaimer* de que no estamos siguiendo un orden cronológico, ya que vamos a hablar antes de ARROW que de BERGSON y SAMUELSON.

Habiendo mencionado la FBS, decir que ARROW cambia radicalmente el enfoque y desarrollar Arrow con más detalle.

1.2. Teorema de la Imposibilidad de Arrow (1951)

Idea

- KENNETH ARROW, en su obra *"Social Choice and Individual Values"* (1951), cambia radicalmente el enfoque.
- ARROW abre una línea de análisis: la **teoría de la elección social**. ARROW cambia el enfoque al plantearse la pregunta: *¿cuáles son las condiciones o principios que debería cumplir una agregación racional de las preferencias de una sociedad de forma consistente y coherente?*
  - Es decir, en lugar de proponer una Función de Bienestar Social basada en juicios de valor como los autores anteriores, ARROW se pregunta si **partiendo de las preferencias individuales ordinales** (asumidas consistentes) se puede definir **una** preferencia social que representara las preferencias sociales, garantizando la **eficiencia** del resultado.
  - Por lo tanto, se persigue que la Función de Bienestar Social sea una regla de elección social no controvertida[15]. En otras palabras, ARROW quiere agregar las preferencias individuales en una Función de Elección Social de forma consistente y coherente con las preferencias individuales.
    - Este es el primer punto clave de la teoría de la elección social de ARROW.

Función de Elección Social, preferencias individuales y su agregación a una ordenación social

Definición de Función de Elección Social

- De esta manera plantea una *Función de Elección Social* (FES)[16]. Es una regla[17] (R_G) que asigna una **preferencia social** a cada posible perfil de órdenes de preferencias débiles ($\succcurlyeq$) individuales (R_i):

$$R_G = f(R_1, R_2, ..., R_I)$$

  - ARROW, al trabajar con un funcional, busca una mayor generalidad y preservar el supuesto de ordinalidad.

[15] Es importante aquí tener claro lo que queremos decir con *no controvertido*. Es controvertido en el sentido de que cambia el enfoque ya existente de cómo buscar una agregación de las preferencias individuales en una Función de Bienestar Social. Sin embargo, decimos que es *no controvertida* en el sentido de que la agregación se realiza de forma *consistente* y *coherente* con las preferencias individuales.

[16] ARROW habla de Función de Bienestar Social, pero la redenominaremos como Función de Elección Social para marcar una distinción. Esta terminología ha sido adoptada de forma habitual en la literatura económica.

[17] A veces se la denomina *funcional* (función de funciones) o *agregador del bienestar social*.

Preferencias individuales (inputs) y ordenación social (output) con sus respectivas axiomáticas

- Vemos como este funcional parte de unos *inputs* que son las preferencias individuales para obtener como *output* una ordenación social. Procedamos a analizar cada uno de estos componentes:
  - Por un lado, tenemos como *inputs* las **preferencias individuales**.
    - Se supone que cada individuo, tiene una serie de preferencias sobre un perfil de alternativas finitas, N (p.ej. bienes de consumo)[18]:
    $$\left(\alpha_i^1 \succ_i \alpha_i^2 \succ_i \ldots \succ_i \alpha_i^N\right) \in R_i \; \forall i$$
      - Por tanto, los inputs de la Función de Elección Social arrowiana son funciones (R_i) que representan una ordenación individual de preferencias.
    - ARROW supone que el perfil de ordenación de las **preferencias individuales**, R_i, cumplirán con los axiomas de pre-orden completo débil (o de racionalidad) de DEBREU (1959) de la teoría del consumidor [ver tema 3.A.8], es decir, cumplirá los axiomas de *completitud*, *reflexividad* y *transitividad* además del axioma de *consistencia interperfiles*[19], consistente en que las elecciones de todos los individuos deben estar basadas en la misma escala, es decir, que todos los individuos puedan ordenar las mismas N opciones y tengan la misma forma de comparar las alternativas.
    - A partir de estos inputs, obtenemos como output una ordenación social de todas las alternativas, es decir, una regla que va a permitir un orden de preferencia para cada una de las posibles alternativas (R_G).
  - Por otro lado, tenemos como *output* la **ordenación social**.
    - ARROW considera que dicha ordenación social, R_G, debe cumplir una serie de axiomas deseables. Por lo tanto, dota a su teoría de un elevado grado de formalización al desarrollar una teoría de la elección social **axiomática**. Es decir, ARROW realizó en 1951 dentro de la rama de la elección colectiva de la economía del bienestar lo que DEBREU llevaría a cabo en 1959 para la teoría de la elección individual.
      - Este es el segundo punto clave de la teoría de la elección social de ARROW.
    - En concreto, según ARROW cualquier regla de elección colectiva (R_G) que se quiera considerar *apropiada* deberá cumplir **5 propiedades** básicas (*axiomas*):
      1. *Universalidad o dominio no restringido*: La regla de elección colectiva ha de ser definida para cualquier conjunto de preferencias que puedan tener los individuos. Este axioma da generalidad, ya que el mecanismo debe ser capaz de agregar cualesquiera preferencias individuales.
      2. *Transitividad*: Este axioma asegura que si la regla de elección social prefiere la opción A sobre la opción B y la opción B sobre la opción C, preferirá la opción A sobre la opción C:
      $$A \succcurlyeq_G B \text{ y } B \succcurlyeq_G C \Longrightarrow A \succcurlyeq_G C$$
      3. *Independencia de las alternativas irrelevantes*: Las preferencias individuales sobre 2 alternativas dependen únicamente de las preferencias sobre esas alternativas, y no de su posición relativa frente a otras[20]. Este axioma excluye la cardinalidad de las preferencias.

[18] Es decir, α_i^1 sería la primera elección del individuo i, α_i^2 sería su segunda elección, así hasta ordenar las N alternativas disponibles.

[19] Un elemento importante es que los perfiles de preferencias individuales tienen que ser *consistentes*, es decir, los perfiles de preferencias individuales deben definirse sobre la misma escala (p.ej. mismas alternativas y misma forma de comparar las alternativas).

ARROW trabaja con *invarianza ordinal*. Esto está relacionado con su voluntad de trabajar con preferencias y no con funciones de utilidad. Requerimos invarianza ordinal para que una función de utilidad venga de unas preferencias ordinales. Esto es criticable, pues existen muchas otras formas de invarianza (D'ASPREMONT y GEVERS, 2002 – https://philarchive.org/archive/NEBAWI).

[20] Así, por ejemplo, si la elección es entre gasto en sanidad y en educación, el resultado debe ser independiente de cómo valoren los individuos la sanidad y la educación en relación a un tercer bien (como, por ejemplo, la justicia).

4. *Pareto optimalidad débil*: Si *todos* los individuos prefieren la alternativa A a la B, la regla de elección colectiva debe preferir también la A a la B.

$$A \succ_i B \; \forall i \Longrightarrow A \succ_G B$$

Este axioma asegura la eficiencia del resultado[21].

5. *Ausencia de dictadores*: Este axioma asegura que ningún individuo particular controla el mecanismo de elección. En concreto, este axioma postula que no hay ningún individuo i tal que para todos los perfiles del dominio de f y para todos los pares de alternativas, si $A \succcurlyeq_i B$ entonces $A \succcurlyeq_G B$.

Teorema de Imposibilidad de ARROW

- ARROW demuestra que, si existen al menos 3 alternativas y el número de individuos es finito, no existe ninguna regla de elección colectiva que cumpla simultáneamente los 5 axiomas.
  - Lo que el Teorema de la Imposibilidad de ARROW nos dice es que las *reglas de votación no son un buen mecanismo de decisión*, ya que todas ellas deben renunciar a alguna de las 5 propiedades mencionadas, y todas ellas se consideran propiedades básicas y deseables [Para una demostración, ver el anexo A.4].

- El **Teorema de Imposibilidad de Arrow** dejaba claro que en la medida que los anteriores autores intentaban hallar:
  - una única Función de Bienestar Social (*dominio universal*),
  - que suponga una ordenación coherente de preferencias (*transitividad*),
  - que representara las preferencias sociales a partir de preferencias individuales ordinales (*irrelevancia de las alternativas relevantes*),
  - garantizando la eficiencia del resultado (*Pareto-optimalidad débil*),
  - sin que ningún individuo particular controle el mecanismo de elección (*ausencia de dictadores*)

  estaban abocados al fracaso pues se estaban proponiendo un imposible.

- La razón última del Teorema de Imposibilidad hay que buscarla en la *pérdida de información* que se produce al pasar de tantos órdenes como existen entre los individuos de la población a un solo orden social. Estamos reduciendo una realidad de muchas dimensiones a otra de solamente una.

1.3. Valoración y escapatorias al Teorema de la Imposibilidad de Arrow

- El Teorema de Imposibilidad de Arrow, es una **aportación constructiva** (que no busca zanjar el debate existente en la literatura de la economía del bienestar en materia de elección social).
  - En este sentido, podemos decir que ARROW abre una nueva línea de análisis sobre la que van a trabajar un gran número de autores desde varias ópticas. Es decir, el trabajo de ARROW no es destructivo sino **constructivo**, pues **abre nuevos enfoques** y aproximaciones.

Tiene que quedar claro, que ARROW no busca zanjar el debate, sino ser constructivo, de ahí que surjan otros debates posteriores en la literatura.

- Ante el pesimismo que supone este teorema, los **desarrollos posteriores** han ido en 2 direcciones:
  a. Desarrollos que *rechazan* la metodología empleada por ARROW.
    - Volver a la vía ya propuesta de la Función de Bienestar Social *à la* BERGSON-SAMUELSON.
    -
  b. Desarrollos que *aceptan* la metodología de ARROW y estudian cuál es el axioma del que se puede *prescindir*.
    -

[21] Si se exigiera Pareto optimalidad fuerte, entonces para que la alternativa A fuese preferida a la B, no haría falta que todos los individuos la prefiriesen, sino que bastaría con que un individuo la prefiriese y el resto estuviesen indiferentes.

Morreau, M. (2019). Arrow's Theorem. En E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2019). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/entries/arrows-theorem/

2. ESCAPATORIAS AL TEOREMA DE LA IMPOSIBILIDAD DE ARROW (I). RECHAZO DE LA METODOLOGÍA DE ARROW, CONSIDERACIONES ÉTICAS Y JUICIOS DE VALOR. FUNCIONES DE BIENESTAR SOCIAL.

2.1. Idea principal: Rechazo de la metodología de ARROW y búsqueda de otros supuestos o conjuntos de axiomas alternativos

Críticas de LITTLE y BERGSON

- Para autores como LITTLE y BERGSON, la economía del bienestar se ocupa de juicios de bienestar social, más que de una regla mecánica de elección colectiva.
  - **LITTLE** criticó que la regla de elección colectiva de ARROW simplemente actúe como una máquina que agrega valores y permite alcanzar una solución, pero que esta solución no representaría necesariamente los valores de la sociedad.
  - **BERGSON**, por su parte, introdujo una distinción entre aconsejar a un individuo y aconsejar a un funcionario. El objetivo real de la economía del bienestar según BERGSON, es el primero, mientras que la regla de elección colectiva de ARROW indicaría al funcionario lo que debe hacer.
    - Esta crítica va a ser muy útil para trazar una distinción entre la *evaluación del bienestar* de las distintas alternativas por parte de un individuo y las *reglas de elección colectiva* que agregan esas evaluaciones individuales. Esto es útil dentro de la democracia deliberativa, en la que cabe distinguir 2 fases:
      - *Fase I*: los individuos expresan sus juicios de valor.
      - *Fase II*: agregación de las preferencias para la toma de decisiones gubernamentales.
    - Así, la *función de bienestar social* tendría más que ver con la fase I, *deliberativa*, mientras que la *regla colectiva de ARROW* tendría más que ver con la fase II, *decisiva*.
  - ARROW, de hecho, está de acuerdo en que su teoría de la elección social es un mecanismo para la toma de decisiones de la sociedad, y que constituye, pues, más una herramienta para el funcionario, representante de la colectividad, que para el individuo. Pero reclama que la preocupación central de la economía del bienestar debe ser precisamente esa: qué acciones o políticas debe adoptar la sociedad en su conjunto.

Similitudes y diferencias entre el enfoque de ARROW y el enfoque de BERGSON y SAMUELSON

Similitudes y diferencias entre el enfoque de ARROW y el enfoque de BERGSON y SAMUELSON	
Similitudes	
Ambos buscan hallar una preferencia social y resolver el problema de incompletitud del criterio de Pareto.	
Diferencias	
ARROW	BERGSON y SAMUELSON
Busca respetar la consistencia y coherencia de las preferencias individuales	Aplican juicios de valor para hallar esa preferencia social

2.2. Línea neoutilitarista: Función de Bienestar Social

2.2.1. Idea

- El problema fundamental del criterio paretiano es que puede **arrojar infinitos óptimos no comparables entre sí** desde el punto de vista de la eficiencia.
  - Los distintos óptimos globales se recogen en la **Gran Frontera de Posibilidades de Utilidad**, que no es más que la envolvente (de puntos o de tramos) de las Fronteras de Posibilidades de

Utilidad que se corresponden con los puntos de la FPP para los que existe un reparto (o varios) de los bienes producidos que sea globalmente óptimo.

– Al ser capaz de arrojar infinitos óptimos no comparables, el criterio de Pareto puede no ser suficiente para determinar qué asignación maximiza el bienestar social, que es el objetivo último de la economía del bienestar.

  o Además, el *Segundo Teorema Fundamental de la Economía del Bienestar* nos dice que, mediante una redistribución apropiada, cualquier óptimo de Pareto será alcanzable, por lo que resulta especialmente importante determinar cuál es preferible.

- Así, la **eficiencia u optimalidad paretiana global** es una condición necesaria pero no suficiente para garantizar la maximización del bienestar social. Es decir, una situación puede ser óptima en el sentido de Pareto pero no maximizar el bienestar social. Sin embargo, la maximización del bienestar sólo se alcanza en una situación óptima de Pareto.
  – Para determinar qué óptimo de Pareto maximiza el bienestar social, es necesario introducir algún criterio adicional. Como decíamos, aquí nos vamos a enfocar en el criterio de la Función de Bienestar Social.

2.2.2. *Función de Bienestar Social – BERGSON (1938) y SAMUELSON (1947)*

Idea (definición de Función de Bienestar Social)

- El primero en estudiar la Función de Bienestar Social (FBS) fue BERGSON en 1938, siendo su trabajo completado por SAMUELSON en 1947. Siguiendo el enfoque BERGSON-SAMUELSON plantearemos una función de bienestar social en el espacio de utilidades de los individuos[22].
  – **Analíticamente**, podríamos definir la Función de Bienestar Social como:

$$W = W\left(U_1\left(x_1^1, x_2^1, \ldots, x_N^1\right), U_2\left(x_1^2, x_2^2, \ldots, x_N^2\right), \ldots, U_I\left(x_1^I, x_2^I, \ldots, x_N^I\right)\right)$$

  de forma que el "bienestar de una sociedad", W, es simplemente una función de las utilidades de los miembros que la constituyen.

Índices de utilidad individual (inputs) y nivel de bienestar social (output)

- Al igual que la FES de ARROW partía de las preferencias sociales como *inputs* para obtener como *output* una ordenación social, vemos como la FBS tiene como *inputs* las funciones de utilidad individuales definidas sobre bienes de consumo, y obtiene como *output* una medida numérica para cada nivel de bien del consumo de los agentes.

2.2.3. *Desarrollo*

Propiedades deseables de la Función de Bienestar Social según SAMUELSON

- A pesar de tener un objetivo diferente al de ARROW, SAMUELSON considera que la FBS debe satisfacer 3 propiedades deseables:
  1. Se trata de una *FBS individualista*, porque los individuos determinan la utilidad que reciben para cada cesta consumida. Esto está asociado con el no-paternalismo, que implica que cada individuo es el mejor juez de su propio bienestar.
  2. Los *índices de utilidad deben ser ordinales*, sin admitir comparaciones interpersonales de utilidad (CIPU).
     - Una cuestión crucial era determinar si las utilidades que formaban el espacio en el que se definía la Función de Bienestar Social eran ordinales o cardinales.
       - Dado que la teoría económica había rechazado de forma prácticamente unánime las funciones de utilidad cardinales (pues permitían comparaciones

[22] Existe otro tipo de Función de Bienestar Social: la de BERGSON-SCITOVSKY, definida para el espacio de bienes (de manera que las curvas de indiferencia social se combinarían con la FPP, y no con la GFPU como en el caso de la Función de Bienestar Social de BERGSON-SAMUELSON).

interpersonales de la utilidad, algo que exigiría un exceso de racionalidad difícil de sostener[23]), lo ideal sería prescindir de ellas.

¡Cuidado! Tal y como se indica en la nota al pie de página 43, las funciones de utilidad cardinales no necesariamente permiten comparaciones interpersonales de la utilidad ni viceversa, puede haber ordinales con CIPU (p.ej. Rawls)

- Sin embargo, esto no va a ser posible, ya que la ordinalidad no es compatible con las decisiones sociales, que deben valorar, por definición, *trade-offs* entre individuos, por lo que es imposible no hacer comparaciones cardinales interpersonales[24].

3. Debe ser consistente con el *criterio de Pareto*.
   - Matemáticamente, se puede trabajar con la primera derivada parcial de $W(\cdot)$, $\partial W(\cdot)/\partial U_i$ como la utilidad marginal social de la persona i, es decir, el incremento en el bienestar social asociado a un incremento marginal de la utilidad de la persona i, manteniendo constante la utilidad del resto de los agentes.
   - Para que la FBS sea consistente con el criterio de Pareto se tiene que cumplir que $\partial W(\cdot)/\partial U_i > 0 \; \forall i$, es decir, si la utilidad de una persona aumenta y la del resto de personas se mantiene constante, el bienestar debe aumentar (monotonía de la FBS).
   - Sin embargo, la FBS va más allá del criterio de Pareto. Como desarrollaremos más adelante, es necesario definir una forma funcional de la función $W(\cdot)$.
     - Nótese que la FBS, $W(\cdot)$, permite comparar asignaciones que no eran comparables por el criterio de Pareto.
     - Permite comparar 2 asignaciones cualquiera, incluso cuando el cambio de una asignación a otra aumente la utilidad de algunos individuos y disminuya la de otros.
   - **Gráficamente**, una vez se define una forma funcional de la FBS, si suponemos que existen 2 individuos, la FBS se puede representar a través de las *curvas de indiferencia social*, que son el lugar geométrico de las diferentes combinaciones de utilidad de los individuos A y B que suponen el mismo bienestar social.
     - La curvas de indiferencia social son ubicuas, no se cortan, de mayor utilidad cuanto más alejadas del origen, con pendiente negativa y convexas.

Problema del planificador

- Una vez se tiene definida la forma funcional, la elección social representaría la solución del problema del planificador de maximizar la FBS sujeto a las restricciones de recursos que dan lugar a la Gran Frontera de Posibilidades de Utilidad[25]:

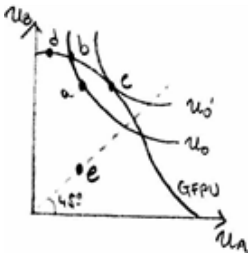
$$\max_{\{U_1,U_2\}} W\big(U_1(x_1,x_2,\dots,x_N),\, U_2(x_1,x_2,\dots,x_N)\big)$$
$$s.a. \quad GFPU$$

[23] Como señaló LIONEL ROBBINS (1938), *"las comparaciones interpersonales de utilidad carecen de fundamento científico porque cada mente es inescrutable para otra mente"*.

[24] Nótese que las comparaciones interpersonales de utilidad no tienen por qué exigir el "exceso de racionalidad" del que hablan sus detractores. Así, por ejemplo, podríamos decir que el incendio de Roma produjo más desutilidad a sus ciudadanos que la ganancia de utilidad de NERÓN, aunque no pudiéramos dar valores concretos. Sobre cuestiones como éstas se fundamenta el argumento de que sí se pueden realizar comparaciones interpersonales de utilidad.

[25] El concepto de la Gran Frontera de Posibilidades de Utilidad es un concepto que crea SAMUELSON para resolver su problema matemático.

- **Gráficamente**,
  - La solución se encontraría en el punto de tangencia entre la GFPU y la curva de indiferencia social, es decir, donde la pendiente de la GFPU se igualara a la Relación Marginal de Sustitución Social[26].
  - En el ejemplo gráfico, el *optimum optimorum* (el mejor de todos los óptimos) se daría en el punto *c*.
  - Esta representación también nos permite analizar el efecto de cambios en la situación social.
    - Por ejemplo, un cambio que desplazara a la sociedad desde el punto *a* hacia el *c*, aumentaría el bienestar de la sociedad ya que se encontraría sobre una curva de indiferencia social más alejada del origen. De la misma manera, un cambio que desplazara a la sociedad de *a* a *b* no lo alteraría.
    - El punto de máximo bienestar social se alcanzará en el punto de tangencia entre la Gran Frontera de Posibilidades de Utilidad y la curva de indiferencia social más alejada del origen. Este punto se denomina *"punto de felicidad" o "bliss point"* (punto *c*)[27].
      - Ahora podemos ver por qué la optimalidad de Pareto es una *condición necesaria pero no suficiente* para la maximización del bienestar. La maximización del bienestar se producirá en un punto sobre la GFPU, cuyos puntos son todos óptimos paretianos. Sin embargo, un gran número de puntos por debajo de la GFPU, si bien no constituyen óptimos en el sentido de Pareto, implican un nivel de bienestar social más alto que algunos puntos sobre la GFPU (p.ej. el punto *a*, ineficiente, genera mayor bienestar social que el punto *d*, eficiente). No obstante, dado cualquier punto ineficiente (i.e. por debajo de la GFPU), existirá algún punto sobre la GFPU que represente una mejora en el bienestar social.



Implicaciones

- Hemos demostrado, pues, que bajo ciertos supuestos una economía puede alcanzar el **punto de máximo bienestar social**.
  - Debería destacarse que el punto de felicidad (y la solución del sistema para los valores de las 10 variables que constituyen las incógnitas en el problema de maximización del bienestar del modelo 2×2×2×2) es puramente **tecnocrático**.
    - Es decir, se alcanza por las formas funcionales supuestas, que permiten la igualación de valoraciones de los distintos agentes ($RMS = RMT$) con la igualación social (pendiente de la curva de indiferencia social), **con independencia de los precios**.
  - Si introducimos los precios en un contexto de competencia perfecta, se puede demostrar que el equilibrio general competitivo será óptimo de Pareto global, pero no necesariamente maximizará el bienestar social.
    - Que el *punto de equilibrio general competitivo* (que es un óptimo global) maximice, a su vez, el bienestar social, dependerá de si dicho equilibrio coincide con el punto de máxima felicidad, esto es, con el punto donde la Gran Frontera de Posibilidades de Utilidad es tangente a la curva de indiferencia derivada de la función de bienestar social.

[26] Podríamos definir la Relación Marginal de Sustitución Social como la pendiente de la curva de indiferencia social en un punto, es decir, la cantidad de utilidad del individuo 2 a la que la sociedad está dispuesta a renunciar a cambio de un aumento unitario de la utilidad del individuo 1. En valor absoluto:

$$\left|RMS_{Social}{}^{U_2}_{U_1}\right| = -\frac{dU_2}{dU_1} = \frac{\partial W/\partial U_1}{\partial W/\partial U_2}$$

La convexidad implicará que la RMS aumenta (se reduce en valor absoluto) a medida que descendemos a lo largo de la curva de indiferencia.

[27] Una vez determinado el punto que maximiza el bienestar social, se puede desandar el proceso seguido para determinar los óptimos paretianos y obtener los valores únicos optimizados de las 10 incógnitas del modelo 2×2×2×2: L_X^*, L_Y^*, K_X^*, K_Y^*, X^*, Y^*, X_A^*, X_B^*, Y_A^*, Y_B^*.

- Precisamente, esta posible divergencia entre el punto eficiente de equilibrio y el punto eficiente de maximización del bienestar social es lo que da una mayor relevancia al *Segundo Teorema Fundamental de la Economía del Bienestar (2TFEB)*, que defiende que, mediante una redistribución de los recursos vía transferencias de suma fija y dejando después actuar al mercado, se puede alcanzar cualquier solución eficiente en sentido de Pareto.
  - De este modo, a través de una adecuada redistribución, dejando actuar al mercado se podría alcanzar cualquier punto de la GFPU y maximizar el bienestar social.
- No obstante, en la práctica es prácticamente imposible introducir un sistema de impuestos y transferencias que no distorsione las decisiones y el comportamiento de los agentes económicos, por lo que pueden surgir pérdidas de eficiencia y aparecer con fuerza el *trade-off equidad vs. eficiencia*.
  - Imaginemos que la situación de partida es *d*, y el Gobierno quiere alcanzar la asignación *c*, que es más igualitaria (i.e. más cerca de la bisectriz) y tiene asociado un mayor nivel de bienestar social. Para ello, aplica una serie de actuaciones que, en la medida en que son distorsionantes, hacen que no se llegue a *c*, sino a *a*, ineficiente pero socialmente preferida a la inicial.
  - Pero imaginemos que el Gobierno quiere continuar con la redistribución. Si continúa aplicando medidas, puede que los beneficios sociales en equidad no superen a los costes en términos de eficiencia económica, y la economía se acabe situando en *e*, más igualitario pero con un bienestar social menor que *a* y, posiblemente, peor a la situación inicial *d*.

2.2.4. Valoración

- La Función de Bienestar Social presenta una serie de **aspectos positivos**:
  - *Resuelve el problema de la incompletitud* habiendo seleccionado una forma funcional.
  - La definición de una Función de Bienestar Social *es útil para ramas como la Economía Pública donde existen trade-offs entre eficiencia y equidad* (teoría de la imposición óptima).
- Sin embargo, también presenta **inconvenientes**:
  1) SAMUELSON reconoce limitaciones de su trabajo. Se puede partir de índices de utilidades individuales, pero *la FBS debe tener algún elemento de cardinalidad que permita medir el bienestar social*.
     - En la teoría económica habían primado las funciones de utilidad ordinales por su mayor generalidad y por evitar comparaciones individuales de utilidad.
     - Sin embargo, SAMUELSON encuentra que precisamente dicha generalidad la hace no apta para el análisis de la decisión social, donde los *trade-offs* entre los individuos deben ser considerados.
       - Por tanto, SAMUELSON concluye que la única manera de comparar los puntos eficientes de la GFPU es con una *FBS cardinal* que permita *comparaciones interpersonales de utilidad*.
  2) Otro aspecto negativo es la *dificultad a la hora de definir una forma funcional específica* de la FBS. BERGSON se remite a caracterizar la FBS sin proponer una forma funcional concreta.
     - Por tanto, BERGSON admite que cada forma funcional incorporaría propiedades adicionales a la eficiencia. En concreto, propiedades sobre justicia y equidad.
     - En este sentido, sería *necesario incorporar juicios de valor*.

2.3. Valores éticos subyacentes (elección de una forma funcional)

2.3.1. Idea

- Para definir una forma funcional de la FBS, estableciendo juicios de valor, se puede teorizar sobre qué sería lo justo en una sociedad.
  - El objetivo de este apartado será justificar la elección de una forma funcional sobre otra.

2.3.2. Consideraciones éticas: la filosofía utilitarista y el neocontractualismo

- Hasta ahora hemos asumido una Función de Bienestar Social con curvas de indiferencia convexas. Esta es una situación intermedia entre las 2 posturas más extremas, propuestas por autores que introducen en la teoría económica visiones de filosofía política:
  - *Filosofía utilitarista* (HARSANYI[28], quien hereda los postulados de BENTHAM).
    - Supone que el bienestar social es la *suma* de los bienestares individuales ($W = \sum U_i$), por lo que será socialmente deseable toda asignación que incremente el bienestar de los beneficiados más de lo que disminuye el de los perjudicados, aunque esto implique una ausencia absoluta de equidad.
    - Según HARSANYI, cuando los agentes actúan bajo el "velo de la ignorancia"[29] (i.e. desconocen *ex-ante* su lugar en la sociedad), serán *neutrales* al riesgo.
    - Las curvas de indiferencia social son *líneas rectas con pendiente unitaria*, ya que los niveles de utilidad de uno y otro individuo son sustitutivos perfectos y ponderan igual.
    - Para establecer este resultado, debemos distinguir entre 2 clases de preferencias expresadas por cualquier individuo dado:
      - Una clase incluye sus *preferencias personales* basadas en sus gustos y sus intereses personales y sus preferencias morales o sociales (i.e. sus puntos de vista sobre los intereses generales de la sociedad). Sus preferencias personales pueden ser representadas por una función de utilidad von Neumann-Morgenstern (i.e. por una función de utilidad cardinal que mide la utilidad de cualquier lotería por su utilidad esperada [ver tema 3.A.10]).
      - Por otro lado, las *preferencias morales* se representan por su particular propuesta de Funciones de Bienestar Social. Por lo tanto, cada individuo tendrá una FBS propia, posiblemente diferente de la FBS del resto de la sociedad.
    - Ahora supongamos que un individuo dado eligiese entre 2 situaciones sociales en completa incertidumbre: lo único que conoce es que tendría exactamente la misma probabilidad de ocupar cualquiera de las posiciones sociales disponibles, es decir, adoptar la identidad personal de cualquier individuo de la sociedad[30].
    - Según HARSANYI, lo que este individuo desearía sería elegir la situación social que le proporcionase una mayor utilidad esperada, que en este caso elegir la situación que concediese un nivel medio de utilidad más alto a los individuos de la sociedad (la mayor utilidad esperada posible). Es decir, HARSANYI considera que la FBS viene dada por una función de utilidad esperada social de von Neumann-Morgenstern, que por definición es lineal en probabilidades. Por ello, HARSANYI apuesta porque la FBS sea una función lineal de las utilidades individuales:

$$W = \sum U_i$$

    - Esta función aditiva da lugar a unas Curvas de Indiferencia Social (CIS) rectas con pendiente unitaria, ya que los niveles de utilidad de uno y otro individuo son sustitutivos perfectos y se ponderan igual[31].

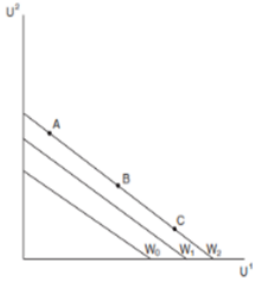
[28] Harsanyi, J. C. (1955). Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility. *Journal of Political Economy*. https://doi.org/10.1086/257678.

[29] HARSANYI no usa el término "velo de la ignorancia", pero se desprende de su teoría. El término es de RAWLS.

[30] El teorema de Harsanyi mostró que dados algunos supuestos relativamente comunes sobre las preferencias de la sociedad, la FBS es necesariamente utilitarista. Así, si la sociedad maximiza el bienestar social esperado, los agentes maximizan la utilidad esperada y la sociedad es indiferente entre 2 distribuciones de probabilidad sobre estados sociales cuando todos los agentes lo son, la FBS será necesariamente utilitarista.

[31] Aunque pueda parecer razonable a priori por su sencillez y compatibilidad con el criterio de Pareto, varias paradojas han sido descritas. La más importante, denominada *"resultado repugnante"* hace referencia al hecho de que el simple aumento de los individuos de una sociedad puede hacerla preferible a otra con menos agentes aunque en la primera se encuentran todos en la miseria. Una solución a esta paradoja implica la transformación de la FBS utilitarista simple en una media de las utilidades en vez de una simple suma.

IMAGEN 2.– *Curvas de Indiferencia Social de la FBS utilitarista (HARSANYI)*



Fuente: …

– <u>*Filosofía rawlsiana o neocontractualismo*</u> (RAWLS[32], quien hereda las ideas de ROUSSEAU, HOBBES y LOCKE sobre el contrato social[33]).
  - Supone que el bienestar social viene determinado por la *utilidad del individuo más desfavorecido* ($W = min\{U_1, U_2, \ldots, U_I\}$), por lo que el bienestar social mejora sólo si mejora él, de forma que el óptimo social es uno de igualdad absoluta.
  - Según RAWLS, cuando los agentes actúan bajo el "velo de la ignorancia"[34] (i.e. desconocen ex ante su lugar en la sociedad), tendrán *aversión infinita a la incertidumbre*, dando lugar a un criterio social *maximin* [ver tema 3.A.10], es decir, el bienestar social sólo mejora si aumenta la utilidad del individuo más desaparecido.
  - Las curvas de indiferencia social tendrán *forma de L* y sus vértices se contienen en la bisectriz, ya que sólo aumenta el bienestar social si las utilidades individuales aumentan en la proporción 1 a 1 (es decir, si, tras el aumento, cada individuo sigue teniendo la misma utilidad que el otro)[35].

[32] RAWLS (*"Una Teoría de la Justicia"*, 1971) basa esta aportación en sus trabajos sobre filosofía política.

[33] Podríamos resumir las ideas de estos autores de la siguiente manera:
- HOBBES: Habla del Leviatán y afirma que en la naturaleza humana sobresale el enfrentamiento entre personas. En el orden social los individuos renuncian a ser libres y el poder debe ser absoluto para evitar los enfrentamientos.
- ROUSSEAU: Para vivir en sociedad, los individuos firman por voluntad propia un contrato social en virtud del cual admiten la existencia de una autoridad.
- LOCKE: La vida del hombre pertenece a Dios. El hombre no es súbdito de otro hombre sino que es libre. Se necesita un orden social que permita dirimir qué hacer en caso de conflicto entre 2 individuos. La autoridad ha de ser un parlamento que represente a todos.

[34] RAWLS hace uso del concepto <u>*velo de ignorancia*</u> como la situación de los participantes en la posición original. Plantea que si no tuviéramos información sobre nuestras características sociales (i.e. desconocemos nuestra raza, religión, etc.), buscaríamos intereses universalizables para establecer los principios básicos de justicia de modo equitativo.

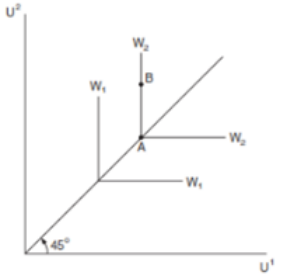
Los principios de justicia son resultado de un acuerdo original hipotético al que llegarían personas libres y racionales interesadas en promover sus propios fines, en una situación inicial de igualdad. Serían los cimientos de una sociedad justa y servirían de criterio para establecer si las leyes o decisiones políticas son justas o no. Introduce 2 *principios de justicia*:

1. *Principio de libertad*: Cada persona tiene el mismo derecho a un esquema de libertades básicos iguales que sea compatible con un esquema similar de libertades para todos. Así, todos deben tener igual derecho a estas libertades, pues son requisito para lograr la realización o modificación de cualquier plan de vida. RAWLS considera que las libertades básicas son un bien de suprema importancia.
2. *Principio de diferencia*: Establece las 2 condiciones que han de darse para que las desigualdades sociales y económicas sean admisibles. En primer lugar, el Estado deberá garantizar que una persona con una posición social y económica inferior a otra pueda hacer efectivos sus derechos y libertades, así como realizar su propio plan de vida si su mérito y esfuerzo lo permiten. Así, el Estado deberá conceder ventajas a estas personas (becas, exenciones fiscales…). En segundo lugar, solo admitiremos el grado de desigualdad si esta supone un beneficio para los más desfavorecidos, atenuando así las desigualdades reales.

[35] ROBERT NOZICK, en su obra *Anarquía, Estado y utopía* (1974) responde a la obra *Teoría de la justicia* (1971) de JOHN RAWLS, afirmando que si existiese igualdad de oportunidades las desigualdades no serían indeseables. Por el principio de diferencia, uno de los criterios por los cuales la desigualdad es legítima es si beneficia la posición de los más desfavorecidos.

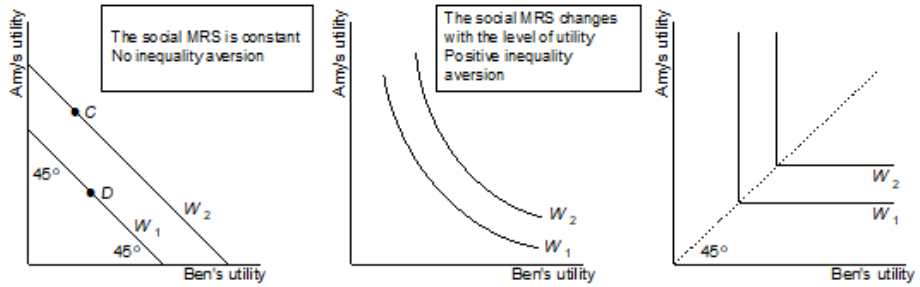
NOZICK considera que una distribución es justa si, inicialmente, los bienes de cualquier individuo o grupo han sido adquiridos legalmente. RAWLS, en cambio, considera que en esa distribución inicial influyen extraordinariamente las diferencias en las capacidades naturales y en las condiciones sociales y que, para contrarrestar estos factores, no se contempla ningún mecanismo que dé lugar a una mayor igualdad.

IMAGEN 3.– *Curvas de Indiferencia Social de la FBS contractualista (RAWLS)*



Fuente: …

IMAGEN 4.– *Curvas de indiferencia social para distintas Funciones de Bienestar Social*



Fuente: …

2.4. Justificación de criterios equitativos y de justicia

- Hemos visto que RAWLS propone una asignación más equitativa haciendo uso del concepto del velo de la ignorancia, es decir, utilizando una lógica de justicia bajo incertidumbre.
  - Esta metáfora es útil a nivel teórico, pero ¿cómo justificamos la inclusión de criterios equitativos y de justicia en la teoría de la elección social?
  - En la teoría económica hay distintos autores que han buscado dar una respuesta a esta pregunta. Aquí revisaremos las siguientes contribuciones:
    - ANTHONY ATKINSON
    - Enfoque empírico de RICHARD THALER
    - Crítica a las Funciones de Bienestar Social de AMARTYA SEN

2.4.1. ANTHONY ATKINSON

- http://blognewdeal.com/andrea-lucai/tony-atkinson-nos-deja-una-frontera-mas/
- https://nadaesgratis.es/juan-francisco-jimeno/professor-sir-tony-atkinson-la-tenaz-lucha-contra-la-desigualdad-i
- https://nadaesgratis.es/juan-francisco-jimeno/professor-sir-tony-atkinson-la-tenaz-lucha-contra-la-desigualdad-ii
- https://nadaesgratis.es/admin/desigualdad-movilidad-e-innovacion
- https://nadaesgratis.es/bentolila/la-desigualdad-en-perspectiva-historica-comparada

- ANTHONY ATKINSON considera que el criterio de eficiencia paretiana presenta un problema, y es que incorpora un juicio de valor implícito de aceptar como dada la distribución actual de recursos, y que no permite valorar la distancia en la distribución.
- En consecuencia, propone que una ordenación social debe tener en cuenta índices internacionalmente aceptados como los índices de desigualdad de renta.
  - En este sentido, propone el índice de ATKINSON en el que, a diferencia del más citado índice de Gini, se incorpora in parámetro para valorar el grado de aversión de la sociedad (o del investigador) a la desigualdad.
  - Cuanto mayor es el parámetro, mayor es la preocupación por las desigualdades en la parte baja de la distribución de la renta, de forma que el país con mayor desigualdad en este tramo, sale reflejado con un índice peor.

2.4.2. Enfoque empírico de RICHARD THALER

- Por su parte, THALER[36], KAHNEMAN[37] y KNETSCH en su obra *"Fairness and the Assumptions of Economics"* (1986) afirman que "la ausencia de consideraciones de justicia en las teorías estándar de la economía es uno de los grandes contrastes entre la economía y el resto de las ciencias sociales".
  - Se puede argumentar que el propio comportamiento humano en algunas situaciones le llevan a interesarse por la suerte de los demás, tal y como subrayaba ADAM SMITH en su *Teoría de los Sentimientos Morales* (1759).
- THALER, desde un enfoque más empírico y microeconómico, pretende mostrar la existencia de consideraciones éticas y de justicia en nuestras preferencias mediante el juego del ultimátum y el juego del dictador.

Juego del ultimátum

- Es un juego en el que participan 2 individuos que se deben repartir una renta, digamos de 10 €. El primer individuo decide el reparto y el segundo decide aceptarlo (en cuyo caso cada uno se lleva su parte) o rechazarlo (en cuyo caso ningún jugador se lleva nada).
  - Siguiendo una lógica racional, de teoría de juegos [ver tema 3.A.14] podríamos considerar este juego como un juego secuencial de 2 etapas, y podríamos resolverlo por *inducción hacia atrás*.
    - Así, vemos que el Equilibrio de Nash Perfecto en Subjuegos sería que el jugador 1 proponga una distribución en el que se queda con los 10 € menos una parte infinitesimal que daría al jugador 2 para que este esté mejor aceptando el reparto.
  - Sin embargo, THALER observa mediante experimentos que cuando este juego se pone en práctica, el jugador 1 ofrece de media cerca de mitad de la renta a repartir y cuando se realizan ofertas por debajo del 30 % de la renta a repartir el jugador 2 suele rechazar la oferta y preferir irse sin nada.
- Podemos realizar 2 lecturas de este resultado:
  - Los individuos son egoístas, en el sentido de que el jugador 1 ofrece una cantidad positiva y cercana al 50 % por miedo al rechazo de la oferta por parte del jugador 2; o
  - Los individuos tienen cierto sentido de justicia, que justificaría introducir criterios equitativos en las preferencias de los individuos.

Juego del dictador

- Para saber porque se da el resultado observado en el juego del ultimátum, THALER propone un segundo juego: el **juego del dictador**.
  - Este juego es idéntico al anterior, pero sin dar al jugador 2 la opción de rechazar, es decir, se llevará a cabo el reparto propuesto por el jugador 1.
  - Este juego está creado para discernir si el individuo 1 en el juego del ultimátum ofrece cantidades positivas y cercanas al 50 % por miedo al rechazo de la oferta por parte del jugador 2 o si lo hace así porque existen consideraciones de justicia en la toma de decisiones individuales.
- Pues bien, en esta situación, las decisiones de reparto del jugador 1 apenas se ven modificadas y de media se sigue ofreciendo al jugador 2 cerca del 50 %, por lo que este estudio parece concluir que existen consideraciones de justicia en la toma de decisiones individuales.

[36] RICHARD THALER fue galardonado con el Premio Nobel de Economía en 2017 «Por sus contribuciones a la economía conductual.»

[37] DANIEL KAHNEMAN fue galardonado con el Premio Nobel de Economía en 2002 junto con VERNON SMITH «Por integrar aspectos de la teoría psicológica sobre el comportamiento económico del ser humano en momentos de incertidumbre y realizar análisis empíricos de laboratorio, especialmente sobre mecanismos alternativos de mercado».

- Pero ¿cómo podríamos introducir en la función de utilidad el altruismo, la generosidad, la envidia y la aversión a la desigualdad? Esto abriría un nuevo campo en la literatura que fue abordado por AMARTYA SEN.

2.4.3. Crítica a las Funciones de Bienestar Social de AMARTYA SEN

La obra de SEN con axiomática y más cosas es complicada. Pero aquí daremos 2 ideas básicas (la crítica a la modelización por el abuso del concepto de utilidad y la crítica al concepto de Welfare que propone cambiarlo por Wellbeing)

- AMARTYA SEN criticaría posteriormente a la FBS, afirmando que era una Función de Bienestar Social "Bienestarista" (*"welfarist" social wefare function*).
  - La crítica de AMARTYA SEN parte de la idea de que trabajar con maximización de la utilidad ha llegado demasiado lejos en la literatura[38]. SEN critica este análisis argumentando que desde ADAM SMITH se reconoce que existen una pluralidad de motivaciones para explicar la conducta humana y afirma que la utilidad no es el único concepto para medir el bienestar propio.
  - De este modo, SEN critica el concepto de Bienestar (*"Welfare"*) y propone sustituirlo por el concepto de Bienestar (*"Wellbeing"*)[39], un concepto más amplio que engloba primordialmente la introducción de capacidades o tener acceso a bienes de primera necesidad.
- En segundo lugar, SEN señala que desprenderse de los supuestos de ordinalidad de las preferencias y permitir comparaciones interpersonales de utilidad permite ampliar el panorama de reglas de agregación de manera interesante.
  - Cuando las utilidades individuales pueden ser medidas cardinalmente pero sin permitir comparaciones interpersonales de utilidad seguimos teniendo un resultado de imposibilidad.
  - Por lo tanto, es necesario introducir comparaciones interpersonales de utilidad, es decir, se necesitan juicios de valor más allá del criterio de Pareto. En la práctica muchos de estos juicios de valor han primado la preocupación por una FBS que prime en cierta medida una mayor equidad social. ¿Se puede justificar el empleo de criterios equitativos?

2.5. Valoración de las Funciones de Bienestar Social

- En este apartado hemos estudiado una primera escapatoria al Teorema de Imposibilidad de Arrow.
  - Esta rama de la literatura surge del rechazo de la metodología empleada por KENNETH ARROW, y sugiere que el enfoque de la Función de Bienestar Social no busca ser una máquina que agregue de forma perfectamente consistente y coherente las preferencias individuales, sino que busca medir el bienestar social realizando determinados juicios de valor.
  - En la mayoría de las aportaciones resalta una preocupación por equidad, lo que se puede representar con curvas de indiferencia social convexas.
- Por lo tanto, estos desarrollos no son más que **disquisiciones teóricas** acerca de la forma funcional que podría adoptar una Función de Bienestar Social que pudiera agregar las preferencias individuales.
- En el siguiente apartado, pasamos a estudiar la segunda línea de desarrollo (levantamiento de axiomas), que ha sido más fructífera al tener un **mayor sentido práctico**.

[38] Un ejemplo claro ha sido la extensión del análisis utilitarista desde la teoría microeconómica hasta otros ámbitos como el matrimonio las drogas o el crimen en la teoría de GARY BECKER.

[39] El concepto de *Wellbeing* ha servido en las propuestas de medición que recomienda la ONU.

3. ESCAPATORIAS AL TEOREMA DE LA IMPOSIBILIDAD DE ARROW (II): LEVANTAMIENTO DE LOS AXIOMAS, REGLAS DE VOTACIÓN Y PROBLEMAS DE INFORMACIÓN

3.1. Idea principal: Aceptación de la metodología de ARROW

- Otra escapatoria adoptada por la literatura económica al Teorema de Imposibilidad de Arrow ha sido aceptar el resultado, es decir, aceptamos que un criterio de elección social no puede cumplir los 5 axiomas propuestos por KENNETH ARROW.
  - Sin embargo, **en la práctica se necesitan tomar decisiones sociales** (no todas las elecciones a las que nos enfrentamos son individuales) y, por tanto, **se hace necesario aplicar alguna regla de elección colectiva, aunque sea imperfecta** (i.e. incumplimientos de los axiomas).
- En este sentido, podemos definir la **votación**, como una *regla de elección social que permite agregar preferencias individuales en una colectiva*.
  - Por tanto, una regla de votación no deja de ser una relación entre el perfil de preferencias individuales y el conjunto de preferencias sociales. De ahí que el trabajo de ARROW tuviera como consecuencia el auge de la literatura del análisis matemático de las reglas de votación.
  - Como veremos, las distintas reglas propuestas se van a topar con problemas de información.

3.2. Levantamiento de los axiomas

- De este modo, en este apartado vamos a levantar algunos de los axiomas planteados por ARROW de modo que podamos encontrar alguna regla de elección colectiva aunque sea imperfecta:
  1) En primer lugar, levantaremos el axioma de universalidad.
  2) En segundo lugar, levantaremos el axioma de independencia de las alternativas irrelevantes.
- Vamos a estudiar la regla de elección colectiva de la **votación por mayoría simple**, que consiste en la toma de decisiones cuando existen más votos a favor que en contra, y es la regla predominante en las democracias.
  - El **problema** de la mayoría simple como regla de elección colectiva es que para que **exista una solución única y eficiente** es necesario levantar los supuestos de **universalidad** y de **independencia de las alternativas irrelevantes**, respectivamente[40]. Lo vemos a continuación.

3.2.1. Renuncia al axioma de universalidad

Idea

- El axioma de universalidad dotaba de generalidad a la regla de elección social, al permitir que los individuos pudieran llevar a cabo cualquier ordenación de sus preferencias.

Paradoja de Condorcet (intransitividad de las preferencias sociales cuando las preferencias individuales son multimodales)

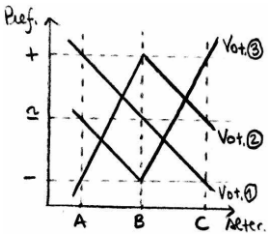
Paradoja de Condorcet (intransitividad de las preferencias sociales)

- Sin embargo, algunos perfiles de preferencias individuales podrían llegar a dar lugar a intransitividad de las preferencias sociales.
  - CONDORCET (1785) demostró que las votaciones por mayoría simple podían dar lugar a situaciones de **circularidad o de ciclicidad**, **incumpliéndose**, por tanto, el **axioma de transitividad**, lo que impedía la **existencia de una solución colectiva** (impidiendo la ruptura de la indeterminación paretiana).

[40] Por suerte, parece lo más razonable que, si es inevitable tener que levantar o relajar algún axioma, se opte por el de universalidad o el de la independencia de las alternativas irrelevantes, ya que son los más débiles, y se mantengan los otros axiomas: Pareto-optimalidad débil, ausencia de dictadores, y transitividad.

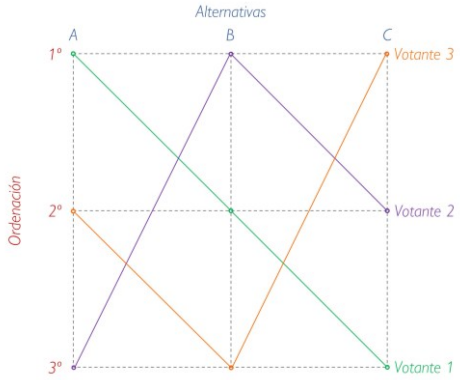
– Esto dio lugar a lo que se conoce como **paradoja de CONDORCET**: las preferencias *individuales* pueden ser transitivas, pero las *colectivas* (resultantes de la regla de elección) pueden no serlo, de manera que la regla de elección colectiva no será capaz de adoptar una solución.

- Veamos un ejemplo:
  – Supuestos:
    - 3 votantes (individuos *1*, *2*, y *3*).
    - 3 alternativas o estados sociales (*A*, *B*, y *C*).
    - Preferencias:
      - Votante 1: $A \succ B \succ C$
      - Votante 2: $B \succ C \succ A$
      - Votante 3: $C \succ A \succ B$
    - Se vota por pares de alternativas.



  – De la confrontación entre *A* y *B* saldría ganadora *A*[41], y de la comparación entre *B* y *C* saldría ganadora la *B*. Si existiera transitividad, al ser preferida la *A* a la *B* y la *B* a la *C*, la *A* debería ser preferida a la *C*. Sin embargo, si enfrentamos la *A* a la *C* resultaría ganadora la *C*. Por lo tanto, no existe transitividad en las preferencias sociales.
  – La intransitividad la originan las **preferencias multimodales** del individuo *3* (i.e. tiene 2 máximos relativos).
    - Pero como la **transitividad** es una propiedad **muy deseable** en la regla de la elección colectiva (pues asegura la racionalidad), tenemos que **asegurarla**.

IMAGEN 5.– *Paradoja de Condorcet (preferencias multimodales)*



Fuente: Elaboración propia

Posibles soluciones al problema de intransitividad de las preferencias sociales

- La transitividad en las preferencias sociales (y, por lo tanto, la existencia de una solución única) se puede conseguir si:
  1. Todas las preferencias son idénticas. *Problema*: es poco realista y en ese caso no sería necesaria una regla de elección colectiva.
  2. Se eliminan las alternativas que hayan sido rechazadas alguna vez. *Problema*: es susceptible de manipulación (la conocida como "manipulación de agenda"), pues la alternativa elegida será siempre la que no es comparada inicialmente, por lo que el resultado dependerá del orden de discusión de las alternativas.
     a. Si en la primera ronda enfrentamos a la alternativa *A* y a la alternativa *B*, ganará la *A*, pero en la segunda ronda, al enfrentarse a la alternativa *C*, ésta última saldrá elegida.
     b. Si en la primera ronda enfrentamos a la alternativa *B* y a la alternativa *C*, ganará la *B*, pero en la segunda ronda, al enfrentarse a la alternativa *A*, ésta última saldrá elegida.
     c. Si en la primera ronda enfrentamos a la alternativa *A* y a la alternativa *C*, ganará la *C*, pero en la segunda ronda, al enfrentarse a la alternativa *B*, ésta última saldrá elegida.

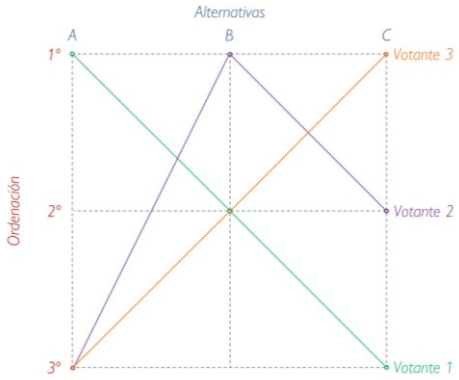
[41] La alternativa *A* obtendría los votos de los votantes 1 y 3, mientras que la alternativa *B* sólo obtendría el voto del individuo 2.

3. Las preferencias son unimodales (DUNCAN BLACK), lo que implica renunciar al axioma de universalidad[42], en el sentido de que ya no vamos a permitir que los individuos lleven a cabo cualquier ordenación de preferencias, sino que vamos a imponer que para todos los individuos las preferencias sean unimodales. Nos vamos a centrar en esta opción.

Preferencias unimodales

- En el ejemplo anterior, para **hacer las preferencias unimodales**, deberíamos modificar las preferencias del individuo 3 (que en un principio eran $C \succ A \succ B$) por $C \succ B \succ A$, de manera que la alternativa ganadora sería la B, y el gráfico quedaría de la siguiente forma (i.e. existe un solo pico –máximo– para cada individuo).

IMAGEN 6.– *Solución al problema de intransitividad (renuncia al axioma de universalidad y preferencias unimodales)*



Fuente: Elaboración propia

¿Es restrictivo?

- Pero, ¿hasta qué punto es **restrictivo** renunciar a la universalidad exigiendo preferencias unimodales? Depende de a qué tipo de elecciones nos enfrentemos.
  - Si la elección se realiza sobre variables cuantitativas esta renuncia podría ser poco restrictiva. Por ejemplo, si describimos las preferencias de un individuo sobre sanidad pública, parece lógico que, si su opción preferida es *mucho* gasto público, su siguiente opción preferida sea gasto público *medio*, y su última opción sea gasto público *nulo*.
    - En este caso, no es demasiado restrictivo exigir la unimodalidad, pues ésta implica que, a medida que un individuo se aleja de su opción preferida, debería estar siempre peor, de manera que las preferencias se modulen y no salten de extremo a extremo.
  - Pero, si por el contrario, la elección se realiza sobre variables cualitativas (p. ej. el nombre que se le pondrá al estadio de una ciudad o el color de los bancos) podría tener menos sentido imponer preferencias unimodales.

- En definitiva, abandonar la condición de universalidad implica un **cierto sacrificio** (pues no se podrán comparar todas las alternativas), pero este sacrificio se juzga menos importante que el que resulta del incumplimiento de la condición de transitividad (i.e. la no existencia de una solución única).

Teorema del votante mediano

- Una implicación importante de imponer unimodalidad y unidimensionalidad en las preferencias de los individuos es que la elección vendrá determinada por las **preferencias del votante mediano** (en el ejemplo anterior, la elección vendrá determinada por quien prefiera B, esto es, por el individuo *2*).
  - Esto otorga a la teoría de la votación una capacidad predictiva, de modo que si conocemos las preferencias del votante mediano podremos predecir el resultado de la votación.

[42] Esta manera de evitar la intransitividad de las preferencias sociales sólo es posible cuando se trata de alternativas unidimensionales, es decir, en las que sólo se decide sobre una variable. Si fueran multidimensionales (p.ej. nivel de gasto público y su distribución), KRAMER demostró en 1973 que la ciclicidad sólo era evitable cuando las preferencias de los individuos eran idénticas, es decir, cuando existía unanimidad sobre las alternativas.

- Y el problema es que las preferencias del votante mediano no tienen por qué coincidir con la *condición de SAMUELSON de provisión óptima de los bienes públicos* ($|RMT_X^G| = \sum|RMS_X^G|$).
- La razón por la cual la regla mayoritaria no conduce necesariamente a provisiones eficientes se debe a que no tiene en cuenta la intensidad de las preferencias, es decir, se debe a que se mantiene el axioma de independencia de las alternativas irrelevantes. Pasemos, por tanto, a levantar este axioma.

3.2.2. *Renuncia al axioma de independencia de las alternativas irrelevantes*

Utilidades cardinales

- Si trabajamos con utilidades cardinales (lo que nos permitirá tener en cuenta la intensidad de las preferencias) podemos levantar el axioma de independencia de alternativas irrelevantes y trabajar con otras reglas de votación.

Reglas de votación

Regla de Borda y el problema de manipulación estratégica

- De este modo, se podrá votar haciendo uso de la **regla de Borda**, consistente en puntuar cada opción reflejando la intensidad de tus preferencias. Al final se suman los puntos que obtiene cada alternativa y gana la que más puntos tiene.
  - La principal ventaja de este método es que refleja la intensidad de las preferencias.
  - Sin embargo, es susceptible de *manipulación estratégica*. Esto es un problema ya que un agente que prevea que sus alternativas preferidas no serán escogidas puede dar una mayor puntuación a las segundas más preferidas en lugar de a las primeras, optando por el "voto útil". Por ejemplo, si el resultado final es la elección de un partido político, un votante que prevea que su partido favorito no va a ser elegido podría otorgar la máxima puntuación a su segundo partido favorito, y además, puede otorgar la mínima puntuación a otros partidos que no sean los últimos en su orden de preferencias pero que considere que pueden ganar por encima de sus partidos favoritos.

Otros sistemas de votación con manipulación estratégica (Regla de segunda vuelta, Ley D'Hont)

- Este problema de manipulación estratégica se da en muchos otros sistemas de votación. Sin ánimo de ser exhaustivo:
  - *Regla de segunda vuelta*: En el primer recuento de votos (en la primera vuelta) eliminamos todos los candidatos excepto los dos que más votos tienen. Luego en la segunda vuelta se enfrentan estos dos y gana el que más votos tiene. La principal ventaja de este método es una mayor gobernabilidad, pero también es susceptible de votaciones estratégicas ya que concentra el voto en partidos más susceptibles de ser elegidos.
  - *Ley D'Hont*: Esta regla electoral tendría el beneficio de una mayor gobernabilidad por introducir un ligero sesgo al alza a favor de las mayorías, pero también existen problemas de manipulación estratégica. Un elector en una circunscripción pequeña que anticipe que un partido minoritario no va a conseguir escaño puede no votar a ese partido a pesar de que sea el partido más cercano a sus preferencias y opte por otro partido en el espacio electoral cercano ideológicamente.

Teorema de GIBBARD-SATTERTHWAITE

- Según el **Teorema de GIBBARD-SATTERTHWAITE**, en elecciones en que concurran 3 o más candidatos, se va a producir uno de los siguientes 3 efectos para cualquier regla de elección:
  1. La regla es *dictatorial* (i.e. un solo individuo puede elegir al ganador), o
  2. Hay candidatos que *nunca podrán ganar*, o
  3. La regla puede ser objeto de *manipulación estratégica* (en el sentido de que el votante no vota respetando sus propias preferencias).

- El teorema de GIBBARD-SATTERTHWAITE es un resultado de gran importancia en el ámbito del *diseño de mecanismos* que fue derivado como resultado de la aparición del teorema de Arrow y su formulación axiomática de la elección social.
  - Uno de los supuestos implícitos al teorema de Arrow es que se conocen perfectamente las preferencias de los agentes, o que pueden conocerse con certeza a la hora de introducir el perfil de preferencias en el funcional. Sin embargo, en la práctica difícilmente esto sucede y los agentes no tienen por qué revelar sus verdaderas preferencias.
  - GIBBARD (1973) y SATTERTHWAITE (1975) demuestran que *dados los supuestos del Teorema de Imposibilidad de Arrow, no existe ninguna agregación de las preferencias individuales tal que los agentes tengan incentivos a revelar sus verdaderas preferencias al tiempo que no existe un dictador.*
    - Esto es, salvo que la relación de preferencia social coincida perfectamente con las preferencias de un agente (i.e. un dictador), los agentes tienen incentivos a revelar preferencias de forma estratégica, que no tiene por qué corresponderse con la realidad.
  - En el ámbito del *diseño de mecanismos* este es un resultado esencial porque pone de manifiesto una serie de condiciones a priori deseables bajo las cuales los agentes no revelan sus verdaderas preferencias y la necesidad de abordar el problema de la compatibilidad de los incentivos.

Evitar comportamientos estratégicos mediante el diseño de mecanismos de incentivos compatibles

- Por tanto, el diseño de mecanismos es la vertiente de la teoría económica que pretende crear un sistema (mecanismo) para evitar estos comportamientos estratégicos, permitiendo la revelación correcta de preferencias y que en definitiva los agentes actúen conforme a sus propias preferencias.
  - Este es un campo de gran relevancia y por el cual HURWICZ, MASKIN y MYERSON reciben el Premio Nobel de Economía en 2007, concretamente, lo reciben «*Por establecer las bases de la teoría del diseño de los mecanismos, que determina cuándo los mercados están trabajando de manera efectiva*».

Ejemplo: Mecanismo de Clarke y Groves para la provisión de un bien público

- Acabamos de ver cómo, para evitar la ciclicidad y garantizar la **existencia** de una solución, hemos tenido que renunciar al axioma de universalidad. En este subapartado vamos a ver que, para garantizar que la solución existente sea, además, **eficiente**, tendremos que renunciar también al **axioma de la independencia de las alternativas irrelevantes**.
- El debate sobre la necesidad de incluir el axioma de la independencia de las alternativas irrelevantes está directamente relacionado con el **debate entre ordinalidad y cardinalidad**[43].
  - Precisamente, ARROW incluyó este axioma para blindar la teoría de la elección social de la cardinalidad, es decir, para garantizar que la solución social fuese independiente de la intensidad de las preferencias.
- Al contrario que en la teoría de la demanda del consumidor, donde existe consenso en relación a la superioridad de la ordinalidad respecto a la cardinalidad (debido a que esta última exige un supuesto adicional –medir el grado de utilidad– sin añadir nada a los resultados), en el caso de la

43

Respecto a la cardinalidad/ordinalidad y la posibilidad de comparaciones interpersonales, el cuadro con las diferentes opciones está en el GAERTNER, capítulo 7, página 112.

Let us recapitulate the different forms of informational set-up in the following scheme:

OMN	no interpersonal comparisons possible, neither of utility levels nor of gains and losses	ordinal Arrovian approach
OMCL	interpersonal comparisons of utility levels possible, but not of gains and losses	ordinal Rawlsian approach
CMN	no interpersonal comparisons possible, neither of utility levels nor of gains and losses	cardinal bargaining approach
CMCU	an interpersonal comparison of gains and losses possible, not of utility levels	cardinal utilitarian approach

utilidad social las conclusiones no son tan claras, ya que la introducción de la **intensidad de las preferencias** es de gran valor.

- En efecto, sin la intensidad podría darse el caso de que una medida no alcance los votos necesarios a pesar de que, sin embargo, el aumento de utilidad que experimentaría la minoría sería mayor que la pérdida de utilidad de la mayoría.
  - En ese caso, el rechazo de una propuesta de este tipo por mayoría sería ineficiente, ya que el bien público no se proveerá hasta que la RMT iguale a la suma de las valoraciones de todos los individuos ($\sum RMS$), sino hasta que iguale a la suma de las valoraciones de los individuos que ostentan la mayoría (así, si la valoración de la medida de gasto público para la mayoría es negativa, votarán en contra y no se llevará a cabo, sin tener en cuenta las valoraciones de la minoría).

- La solución pasa, pues, por **levantar el axioma de la independencia de las alternativas irrelevantes**, introduciendo, pues, **cardinalidad**, lo cual se puede hacer a través de **2 maneras**:
  1. *Intercambio de votos o log-rolling*. Las partes negocian antes de la votación, en forma de intercambio de votos: los votantes apoyan mutuamente sus propuestas, de manera que se tienen en cuenta las intensidades de las preferencias[44]. No obstante, esta solución presenta 3 problemas:
     - Sólo es factible si el número de votantes es reducido, pues genera altos costes de negociación.
     - Existe un riesgo de explotación de las minorías.
     - Podría generar soluciones ineficientes por exceso de provisión: si la regla de la mayoría podía generar soluciones ineficientes por infraprovisión del bien público, el intercambio de votos puede generar ineficiencias en el sentido contrario: podrían aprobarse medidas netamente beneficiosas para determinados votantes, pero que, en conjunto, generasen pérdidas (supraprovisión del bien público).
  2. *Voto ponderado* (o regla de BORDA). Suponiendo que existan 4 alternativas, el individuo asignará 4 puntos a la más preferida, 3 puntos a la siguiente más preferida, y así sucesivamente. Se suman los puntos que consiga cada alternativa y se escoge la que tenga más puntos. Sin embargo, esta manera de introducir la cardinalidad también presenta inconvenientes:
     - Aumenta las posibilidades de empate y, por lo tanto, el coste de adopción de decisiones.
     - Es manipulable, ya que el individuo, sabiendo que su alternativa preferida no va a resultar elegida, puede cambiar el orden y dar más puntos a su segunda preferida (i.e. *voto útil*).

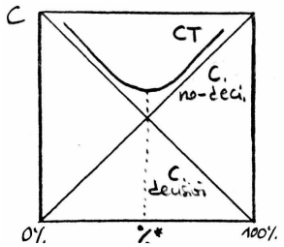
[44] Supongamos que existen tres votantes (*1*, *2* y *3*), y dos proyectos de gasto presentados por *1* y por *2*: el proyecto del votante *1* sólo le beneficia a él, mientras que el proyecto del votante *2* sólo le beneficia a él. Los proyectos, no obstante, deben ser financiados por todos a partes iguales, y a nivel global los beneficios de ambos proyectos son superiores a los costes.

Lógicamente, por la regla de la mayoría no se aceptarían los proyectos, ya que el votante *3* siempre votaría, en contra junto con el votante *2*, del proyecto del votante *1*, y viceversa. Sin embargo, esto no sería eficiente para la sociedad, ya que se estarían rechazando proyectos beneficiosos. Así, lo eficiente sería que los votantes *1* y *2* llegaran a un acuerdo para apoyarse mutuamente sus proyectos.

Nótese que para que no haya incentivos a romper el acuerdo, es necesario que se cree una nueva propuesta de gasto que englobe a las dos anteriores, para que así los votantes *1* y *2* voten simultáneamente –*logrolling*– (si fuesen consecutivos, cuando ya ha sido aprobado, digamos, el proyecto del votante *1*, éste no tendría incentivos para apoyar después el proyecto *2*).

Otras reglas de elección diferentes a la mayoría simple

- En toda la sección de levantamiento de supuestos nos hemos basado en la regla de elección colectiva de la mayoría simple en democracia directa. Vamos a ver ahora otras reglas de elección colectiva:
  - *Regla de la unanimidad* (WICKSELL):
    - Esta regla de votación fue introducida por WICKSELL para resolver el problema de la provisión de los bienes públicos de forma coherente con el criterio de optimalidad de Pareto, de forma que se realizará una provisión determinada sólo si es aceptada por todos los individuos (o, al menos, no rechazada por ninguno), lo cual ocurrirá si ningún individuo empeora y al menos uno mejora.
    - Problemas de esta regla de votación:
      i) elevados costes de decisión, sobre todo cuando los grupos son amplios; y
      ii) mantenimiento del *statu quo*, pues las minorías siempre podrán bloquear una decisión (lo que BAUMOL llamó "tiranía de la minoría").

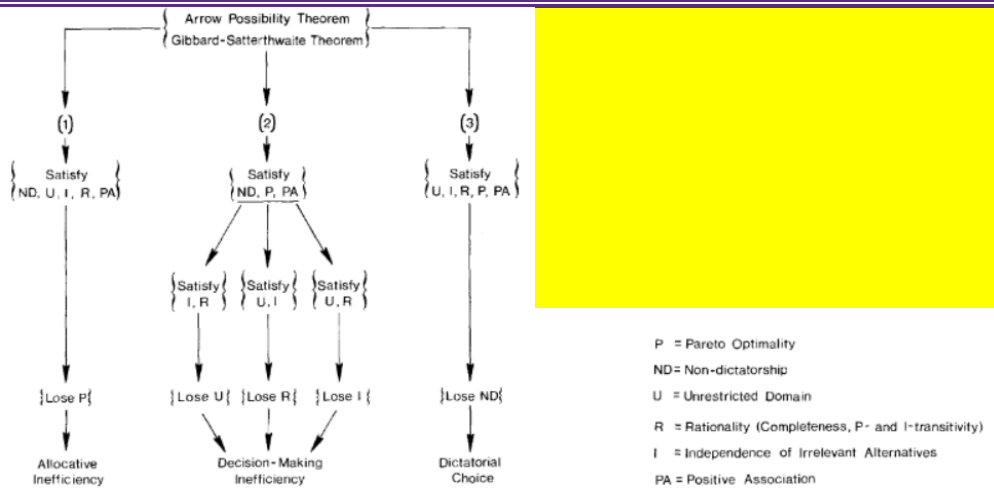


  - *Regla de la mayoría óptima en cada elección* (BUCHANAN y TULLOCK):
    - BUCHANAN y TULLOCK defendieron que, para cada votación, hay una *mayoría óptima*, que surge de la *ponderación* de los 2 tipos de costes que conlleva cualquier regla de votación:
      1. *Costes de la decisión*, entendidos como los costes que tendría la decisión sobre los individuos perjudicados. Cuanto mayor sea la mayoría requerida, menores serán estos costes (pues menores serán los potenciales perjudicados –de hecho, estos costes son nulos en el caso de elección por unanimidad–).
      2. *Costes de la no-decisión (según José Luis, costes externos)*, entendidos como los costes que tendría la no-decisión sobre los beneficiarios potenciales. Cuanto mayor sea la mayoría requerida, mayores serán estos costes (siendo máximos en el caso de elección por unanimidad).
    - *Gráficamente*, la mayoría óptima de la elección vendrá determinada por el mínimo de la curva de costes totales, y variará en cada elección al ser los costes diferentes.

- **Democracia representativa**. Hasta ahora hemos visto cómo se toman las decisiones en un sistema de democracia directa (con mayoría simple y con otros niveles de mayoría). Sin embargo, lo más frecuente en las sociedades democráticas es la democracia representativa, esto es, que en lugar de que los individuos elijan los proyectos, **elijan a los políticos que decidirán sobre los proyectos**. Así, el mecanismo que se genera es el siguiente:

Ciudadanos	→	*Políticos*	→	*Burócratas o funcionarios*
Eligen a los políticos		Eligen los programas		Aplican los programas
		Democracia representativa		Producción pública y burocracia

  - No obstante, la intervención del sector público que exige la democracia representativa (y la producción pública y burocracia que le siguen), suele generar una serie de *problemas* [ver someramente tema 3.A.23].



CONCLUSIÓN

- ***Recapitulación (Ideas clave)***:
  - El **criterio de eficiencia de Pareto** puede arrojar *infinitos óptimos no comparables entre sí*. Determinar el *mejor* de entre todos ellos es particularmente importante, ya que, según el 2TFEB, mediante una redistribución apropiada, cualquier óptimo de Pareto será alcanzable.
    - Para resolver este problema de incompletitud surgieron 2 líneas de investigación: los **criterios de compensación** y la **función de bienestar social**. En esta exposición nos hemos enfocado en esta última.
  - La función de bienestar social nos permite determinar el óptimo de Pareto que **maximiza el bienestar social**, es decir, aquel óptimo en el que la curva de indiferencia social es tangente a la Gran Frontera de Posibilidades de Utilidad.
  - KENNETH ARROW cambia el enfoque y busca unas condiciones deseables (axiomas) que debe cumplir la agregación de las preferencias individuales a través de **reglas de elección colectiva**.
    - No obstante, el propio ARROW concluye que ninguna regla de elección colectiva podrá reunir todos esos axiomas. Es lo que se conoce como **Teorema de la imposibilidad de ARROW**.
  - La exposición ha pivotado entorno a este resultado y a las reacciones que suscitó en la literatura económica:
    - Por un lado, una rama de la literatura rechazó la metodología de ARROW al considerar que la regla de elección colectiva de ARROW simplemente actúe como una máquina que agrega valores y permite alcanzar una solución, pero que esta solución no representaría necesariamente los valores de la sociedad. Estos autores consideraban que la economía del bienestar se ocupa de juicios de bienestar social, más que de una regla mecánica de elección colectiva.
    - Por otro lado, otra rama de la literatura aceptó los resultados de ARROW, y ante la necesidad de tomar decisiones colectivas estudió cuál de los axiomas se podría levantar alcanzando una elección social que fuera lo más justa posible. Concretamente, se centraron en **determinar cuál de los axiomas era más prescindible**, y lo hicieron con respecto a la regla de elección colectiva más común: la **regla de la mayoría simple**. La conclusión es que para que la regla de elección colectiva de la mayoría simple tenga una solución y dicha solución sea eficiente, es necesario renunciar al axioma de la **universalidad** y al de la **independencia de las alternativas irrelevantes**.

    - Por último, otra rama de la literatura se centró en los problemas de la elección social en un contexto de democracia representativa: *Public Choice*.
  - Por último, el **análisis coste-beneficio** [ver anexo A.2] es un mecanismo de asignación de recursos particularmente útil cuando los costes y beneficios privados no coinciden con los sociales.

- ***Relevancia***:
  -

- ***Extensiones y relación con otras partes del temario***:
  -

- ***Opinión***:
  -

- ***Idea final (Salida o cierre)***:
  - En definitiva, la toma de decisiones colectivas se topa con numerosos obstáculos y se puede concluir que es imperfecta. Dada la dificultad de tomar decisiones colectivas una propuesta sería minimizarlas a lo estrictamente necesario, es decir, limitar al máximo la discrecionalidad del sector público y dejar que el mercado descentralizadamente determine una asignación de recursos.
  - Sin embargo, existen decisiones eminentemente colectivas que resulta problemático dejar al arbitrio del mercado. De hecho, la economía pública abre la puerta a la intervención del sector público en la economía por motivos de eficiencia, equidad y estabilización.
  - Por tanto, dado que el sector público tiene un rol importante que jugar en nuestras economías, debemos aspirar a conseguir un sector público funcional.

Bibliografía

Tema ICEX-CECO

Tema Juan Luis Cordero Tarifa

Tema A. Sahuquillo

A. Koutsoyiannis

Tresch

Inman

Gaertner, W. (2006). *A primer in social choice theory*. Oxford University Press.

Jehle, G. A. & Reny, P. J. (2011). *Advanced microeconomic theory* (3. ed., 1. publ). Financial Times, Prentice Hall.

Lucai, A. (s. f.). *Tony Atkinson nos deja una frontera más: Recuperar para la economía su carácter de ciencia moral – Blog NewDeal*. https://blognewdeal.com/andrea-lucai/tony-atkinson-nos-deja-una-frontera-mas/

Morreau, M. (2019). Arrow's Theorem. En E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2019). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2019/entries/arrows-theorem/

Preguntas de otros exámenes

- ¿Cómo casan el Teorema de la Imposibilidad de Arrow con la existencia de óptimos únicos en otras ramas de la Economía?
- ¿Qué aplicación práctica tiene el tema? En concreto, en la Constitución Española.

Enlace a preguntas tipo test

https://www.quia.com/quiz/6550375.html

Anexos

A.1. Anexo 1: Modelo de DOWNS-HOTELLING

- El modelo de DOWNS-HOTELLING muestra cómo, en un sistema bipartidista con votación unidimensional en el espectro ideológico, los 2 partidos tenderán a adoptar una ideología de centro, proponiendo un programa lo más similar posible a la ideología del votante mediano (*principio de diferenciación mínima*). En línea con el teorema de BLACK del votante mediano.

A.2. Anexo 2: Análisis coste-beneficio

- Durante la década de los 70, cuando la Economía del bienestar estaba en uno de sus puntos más bajos debido al éxito del teorema de la imposibilidad de ARROW, empezó a desarrollarse una nueva técnica de asignación de recursos: el **análisis coste-beneficio** (ACB).
- El ACB parte de la existencia de **diferencias** entre los costes/beneficios privados y los sociales. Esta discrepancia lleva a que la **asignación privada** de los recursos, basada en la maximización de beneficios particulares, **no sea óptima**.
- Así, la función del ACB es la de considerar **todos** los costes y beneficios, privados y públicos, derivados de la provisión de un bien o servicio.
- El ACB, tal y como está planteado, tiene su base en los **criterios de compensación**: al decidir sobre la conveniencia de un proyecto, se analiza el valor presente de beneficios y costes *sociales*, de manera que se aceptará el proyecto si los beneficios son mayores que los costes, aunque no recaigan sobre las mismas personas.

- El ACB es especialmente útil para la **asignación de recursos públicos**, dado que a menudo los bienes y servicios ofrecidos por este sector tienen **externalidades importantes**.
- El **problema** más grave del ACB proviene de la ***Teoría del second best***. Supongamos una economía de mercado, con unos precios que no reflejan todos los costes y beneficios, pero que son los que emplean los agentes. Si el sector público utiliza precios sombra representativos de dichos costes y beneficios, nos habremos acercado en un sector (el público) al óptimo paretiano, pero, según LIPSEY y LANCASTER, un movimiento en dirección al óptimo dejando aún imperfecciones en otros sectores, no implica necesariamente una mejor situación que la de partida, e incluso puede empeorarla.

A.3. Anexo 3: Relación de los temas 3.A.16, 3.A.21, 3.A.22, 3.A.23 y 3.A.24

- **Tema 3.A.16**: Competencia perfecta con equilibrio *parcial*.
- **Tema 3.A.21**: Competencia perfecta con equilibrio *general* (haciendo hincapié en la economía como un sistema de ecuaciones simultáneas, nº de ecuaciones vs. nº de incógnitas, ecuaciones de exceso de demanda, simultaneidad, precios de vaciado endógenos, etc.).
- **Tema 3.A.22**: Optimalidad global de Pareto (no hablar de precios, pues la economía del bienestar presenta un sistema tecnocrático –i.e. que se resuelve por las formas funcionales supuestas, con independencia de los precios–) y equilibrio general competitivo como óptimo global de Pareto en ausencia de imperfecciones.
- **Tema 3.A.23**: Si se dan imperfecciones, el equilibrio general competitivo no será óptimo.
- **Tema 3.A.24**: Punto eficiente de equilibrio general competitivo vs. punto eficiente de maximización del bienestar social: ambos no tienen por qué coincidir. Esta divergencia se puede resolver con una redistribución previa de los recursos por parte del sector público, con la intención de que la economía alcance un equilibrio en el punto eficiente donde se maximiza el bienestar social (2TFEB).

A.4. *Anexo 4*: Demostración del Teorema de Imposibilidad de Arrow (JEHLE y RENY, 2011)

Teorema de Imposibilidad de Arrow: Si hay al menos tres estados sociales en X, entonces no existe una función de elección social f que satisfaga simultáneamente las condiciones de universalidad (U), Pareto-optimalidad débil (POD), independencia de las alternativas irrelevantes (IAI) y ausencia de dictador (AD).

THEOREM 6.1 ***Arrow's Impossibility Theorem***

If there are at least three social states in X, then there is no social welfare function f that simultaneously satisfies conditions U, WP, IIA, and D.

Proof: The strategy of the proof is to show that conditions *U, WP*, and *IIA* imply the existence of a dictator. Consequently, if *U*, *WP*, and *IIA* hold, then *D* must fail to hold, and so no social welfare function can satisfy all four conditions.

The proof, following Geanakoplos (1996), proceeds in four steps. Note that axiom *U*, unrestricted domain, is used in each step whenever we choose or alter the preference profile under consideration. Unrestricted domain ensures that every such profile of preferences is admissible.

Step 1: Consider any social state, c. Suppose each individual places state c at the bottom of his ranking. By WP, the social ranking must place c at the bottom as well. See Fig. 6.2.

Step 2: Imagine now moving c to the top of individual 1's ranking, leaving the ranking of all other states unchanged. Next, do the same with individual 2: move c to the top of 2's ranking. Continue doing this one individual at a time, keeping in mind that each of these changes in individual preferences might have an effect on the social ranking. Eventually, c will be at the top of every individual's ranking, and so it must then also be at the top of the social ranking by WP. Consequently, there must be a *first* time during this process that the social ranking of c increases. Let individual n be the first such that raising c to the top of his ranking causes the social ranking of c to increase.

R^1	R^2	$\cdots$	R^n	$\cdots$	R^N	R
c	c	$\cdots$	c	$\cdots$	x''	c
x	x'	$\cdots$		$\cdots$	y''	
y	y'				$\cdot$	$\cdot$
$\cdot$	$\cdot$				$\cdot$	$\cdot$
$\cdot$	$\cdot$				$\cdot$	$\cdot$
$\cdot$	$\cdot$				$\cdot$	$\cdot$
w	w'	$\cdots$		$\cdots$	c	w'''

Figure 6.3. Axioms *WP*, *U*, and *IIA* yield a pivotal individual.

We claim that, as shown in Fig. 6.3, when c moves to the top of individual n's ranking, the social ranking of c not only increases but c also moves to the *top* of the social ranking.

To see this, assume by way of contradiction that the social ranking of c increases, but not to the top; i.e., $\alpha R c$ and $c R \beta$ for some states $\alpha, \beta \neq c$.

Now, because c is either at the bottom or at the top of every individual's ranking, we can change each individual i's preferences so that $\beta P^i \alpha$, while leaving the position of c unchanged for that individual. But this produces our desired contradiction because, on the one hand, $\beta P^i \alpha$ for every individual implies by WP that β must be strictly preferred to α according to the social ranking; i.e., $\beta P \alpha$. But, on the other hand, because the rankings of c relative to α and of c relative to β have not changed in any individual's ranking, *IIA* implies that the social rankings of c relative to α and of c relative to β must be unchanged; i.e., as initially assumed, we must have $\alpha R c$ and $c R \beta$. But transitivity then implies $\alpha R \beta$, contradicting $\beta P \alpha$. This establishes our claim that c must have moved to the top of the social ranking as in Fig. 6.3.

Step 3: Consider now any two distinct social states a and b, each distinct from c. In Fig. 6.3, change the profile of preferences as follows: change individual n's ranking so that $a P^n c P^n b$, and for every other individual rank a and b in any way so long as the position of c is unchanged for that individual. Note that in the new profile of preferences the ranking of a to c is the same for every individual as it was just *before* raising c to the top of individual n's ranking in Step 2. Therefore, by *IIA*, the social ranking of a and c must be the same as it was at that moment. But this means that aPc because at that moment c was still at the bottom of the social ranking.

Similarly, in the new profile of preferences, the ranking of c to b is the same for every individual as it was just *after* raising c to the top of individual n's ranking in Step 2. Therefore by *IIA*, the social ranking of c and b must be the same as it was at that moment. But this means that cPb because at that moment c had just risen to the top of the social ranking.

So, because aPc and cPb, we may conclude by transitivity that aPb. Note then that no matter how the others rank a and b, the social ranking agrees with individual n's ranking. By *IIA*, and because a and b were arbitrary, we may therefore conclude that for all social states a and b distinct from c

$$a P^n b \text{ implies } aPb.$$

That is, individual n is a dictator on all pairs of social states not involving c. The final step shows that individual n is in fact a dictator.

Step 4: Let a be distinct from c. We may repeat the above steps with a playing the role of c to conclude that some individual is a dictator on all pairs not involving a. However, recall that individual n's ranking of c (bottom or top) in Fig. 6.3 affects the social ranking of c (bottom or top). Hence, it must be individual n who is the dictator on all pairs not involving a. Because a was an arbitrary state distinct from c, and together with our previous conclusion about individual n, this implies that n is a dictator. ∎

Although here we have cast Arrow's theorem as an 'impossibility' result, the proof just sketched suggests it can also be stated as a 'possibility' result. That is, we have shown that any social welfare function satisfying the three conditions U, WP, and IIA must yield a social preference relation that exactly coincides with one person's preferences whenever that person's preferences are strict. As you are asked to explore in Exercise 6.3 this leaves several 'possibilities' for the social welfare function, although all of them are dictatorial according to condition D.

37.9 El mecanismo de Vickrey-Clarke-Groves

Analicemos el problema de los bienes públicos en un modelo muy general. El objetivo es elegir un resultado (por ejemplo, instalar o no una farola) que maximice la suma de las utilidades de los agentes interesados. El reto es averiguar cuáles son exactamente esas funciones de utilidad individuales, ya que los consumidores pueden tener incentivos para no declarar el verdadero valor que el bien tiene para ellos.

En el caso más sencillo, el problema podría consistir en elegir entre cero y uno: si $x = 1$, se instala la farola y si $x = 0$, no se instala. En un caso más general, podría consistir en decidir la cantidad que debe suministrarse de un bien: cuántas farolas deben instalarse o cuánta luz deben dar o dónde deben colocarse. Utilizaremos la letra x para representar las opciones posibles, cualesquiera que sean éstas. Suponemos que hay n agentes y que $u_i(x)$ es la utilidad del agente i. El objetivo es elegir el valor de x que maximiza la suma de la utilidad de los agentes, $\sum_i u_i(x)$.

Eso sería fácil si el responsable de tomar la decisión supiera cuáles son las funciones de utilidad. Desgraciadamente, en ninguna situación realista lo sabe. Y como hemos visto, los agentes pueden muy bien tener razones para ocultar sus verdaderas funciones de utilidad.

Aunque parezca mentira, existe una ingeniosa manera de conseguir que los agentes digan la verdad y lograr un resultado eficiente. Este **mecanismo económico** se conoce con el nombre de **mecanismo de Vickrey-Clarke-Groves** o **mecanismo VCG**.

El mecanismo de Groves

Dividiremos la descripción del mecanismo VCG en dos partes. Primero hablaremos de lo que se conoce con el nombre de **mecanismo de Groves**:

1. El centro pide a cada agente i que declare cuánto estaría dispuesto a pagar para tener x unidades del bien público suministrado. Representamos esta utilidad declarada de x unidades del bien público por medio de $r_i(x)$.
2. El centro elige el nivel del bien público x^* que maximiza la suma de las utilidades declaradas, $R = \sum_{i=1}^{n} r_i(x)$.
3. Cada agente i recibe un pago que es la suma de las utilidades declaradas de todos los demás, evaluado al nivel de x determinado en el paso 2. Representamos este pago por medio de $R_i = \sum_{j\neq 1} r_j(x^*)$.

Resulta que, en este mecanismo, cada agente tiene una **estrategia dominante** que consiste en declarar su verdadera función de utilidad. Para ver por qué, consideremos la ganancia total del agente i, que es su utilidad más el pago

$$u_i(x) + \sum_{j\neq i} r_j(x).$$

Obsérvese que al agente i le interesa su verdadera función de utilidad, pero su pago depende de la suma de las funciones de utilidad declaradas de los demás.

El agente i se da cuenta de que el responsable de tomar la decisión maximizará la suma de las utilidades utilizando su utilidad declarada,

$$r_i(x) + \sum_{j\neq i} r_j(x).$$

Sin embargo, el agente i quiere que el responsable de tomar la decisión maximice su propia utilidad (verdadera) más el pago,

$$u_i(x) + \sum_{j\neq i} r_j(x).$$

El agente i puede conseguir que el responsable de tomar la decisión maximice esta expresión declarando su verdadera utilidad; es decir, fijando $r_i(x) = u_i(x)$.

El mecanismo de Groves lo que hace esencialmente es "internalizar la externalidad" entre los agentes. Hace que cada agente se enfrente a los costes y los beneficios que su declaración impone a los demás agentes. Cada agen-

te quiere declarar su verdadera utilidad, ya que eso es lo que quiere que se maximice.

El mecanismo VCG

El problema de la utilización del mecanismo de Groves solamente es que puede resultar caro: el centro tiene que pagar a cada agente una cantidad igual a la suma de las utilidades declaradas de los demás. ¿Cómo puede reducirse la magnitud de estos pagos?

Obsérvese que podemos establecer un "impuesto" sobre cada agente mientras este impuesto sea independiente de la decisión del agente. Si el impuesto es independiente de la decisión de i, entonces no puede afectar a dicha decisión.[3] Elegiremos, pues, el impuesto de modo que garantice que los pagos netos que recibirá el centro sean positivos o cero. Por tanto, el centro siempre tendrá al menos la cantidad necesaria de dinero para financiar el bien público.

Un impuesto especialmente indicado es gravar al agente i con una cantidad igual a la suma máxima de las utilidades declaradas, excluido el agente i. Es decir, cobramos a cada agente la suma de las utilidades declaradas que resultaría si no estuviera presente. El impuesto neto establecido sobre el agente i es, pues,

$$W_i - R_i = \sum_{j \neq i} r_j(x) - \max_z \sum_{j \neq i} r_j(z).$$

Obsérvese que este número es positivo o cero. ¿Por qué? Porque la suma *máxima* de las $n - 1$ utilidades declaradas tiene que ser mayor que cualquier otro valor de esa suma.

Lo que estamos calculando aquí es la diferencia entre lo que ocurriría si el agente i estuviera presente y lo que ocurriría si estuviera ausente. Por tanto, mide el coste neto que impone el agente i a los demás agentes. Mientras i tenga que hacer frente al coste que impone a los demás agentes, tendrá los incentivos necesarios para declarar su verdadera utilidad.

Ahora podemos terminar la descripción del mecanismo VCG. Seguimos los pasos 1 y 2 anteriores, pero sustituimos el paso 3 por los pasos siguientes.

3. El centro también calcula el resultado que maximiza la suma de las $n - 1$ utilidades declaradas si el agente 1, 2, …, n no estuvieran presentes. Sea W_i la suma máxima de las utilidades declaradas que resulta sin el agente i.
4. Cada agente i paga un impuesto igual a $W_i - R_i$.

[3] Es aquí donde es importante el supuesto de utilidad cuasilineal.

37.10 Ejemplos de VCG

Hay que reconocer que el análisis del apartado anterior era abstracto, por lo que resulta conveniente examinar algunos casos concretos.

Subasta de Vickrey

El primer caso que analizamos es la **subasta de Vickrey**, que describimos en el capítulo 17. En este caso, el resultado es sencillo: qué persona debe recibir el artículo que se subasta. Sean $v_1 > v_2$ los verdaderos valores de dos postores y $r_1 > r_2$ los valores declarados.

Si el agente 1 está presente, obtiene una utilidad de v_1. Si está ausente, el artículo se adjudica al otro agente, por lo que la ganancia total del agente 1 es $v_1 - r_2$. El agente 2 obtiene una ganancia de cero en todo caso. Cada agente tiene un incentivo para declarar su verdadero valor, por lo que acabamos obteniendo el resultado óptimo.

El mecanismo de Clark-Groves

El siguiente ejemplo es un problema de bienes públicos como el del juego de la compra de un televisor que describimos en la tabla 37.1. Al igual que en ese ejemplo, supongamos que hay dos compañeros de habitación que están tratando de decidir si compran o no un televisor. Sea c_i la cantidad que pagará el agente i si compran el televisor. Dado que el coste total del televisor es de 150 euros, debemos tener que $c_1 + c_2 = 150$.

Según el mecanismo VCG, cada agente declara el valor que tiene el televisor para él, representado por r_i. Si $r_1 + r_2 > 150$, los agentes comprarán el televisor y lo pagarán de acuerdo con el mecanismo. Sea $x = 1$ si compran el televisor y $x = 0$ si no lo compran.

Antes de analizar el mecanismo VCG, pensemos qué ocurriría si utilizáramos el siguiente mecanismo: pedir a cada agente que declarase su valor y adquirir el televisor si la suma de los valores declarados fuese superior al coste del televisor.

Supongamos que el valor de la persona 1 es mayor que la parte que le corresponde pagar del coste, por lo que $v_1 - c_1 > 0$. En ese caso, la persona 1 también puede declarar, por ejemplo, un millón de euros; eso garantizará la compra del televisor, que es lo que quiere conseguir. En cambio, si $v_1 < c_1$, la persona 1 también puede declarar, por ejemplo, un millón de euros negativo.

El problema se halla en que cada agente, actuando independientemente, no tiene ninguna razón para tener en cuenta los valores del otro agente. Los agentes tienen muchos incentivos para exagerar de una u otra forma los valores que declaran.

Veamos cómo resuelve el mecanismo VCG este problema. La ganancia del agente 1 es

$$(v_1 - c_1)x + (r_2 - c_2)x - max_y(r_2 - c_2)_y.$$

El primer término es la utilidad neta que le reporta el televisor: el valor que tiene para él menos el coste que tiene que pagar. El segundo término es la utilidad neta declarada de su compañero de habitación. El último término es la utilidad máxima que obtendría su compañero de habitación si el agente 1 no estuviera presente. Dado que el agente 1 no puede influir en él, podemos dejarlo de lado por el momento.

Reordenando los 2 primeros términos, tenemos la ganancia del agente:

$$[(v_1 + r_2) - (c_1 + c_2)]x.$$

Si es positiva, puede conseguir que se compre el televisor si declara $r_1 = v_1$, ya que en ese caso la suma de los valores *declarados* será mayor que el coste total. Si es negativa, puede conseguir que no se compre el televisor declarando $r_1 = v_1$. De cualquiera de las dos formas, resulta óptimo declarar el verdadero valor. Lo mismo ocurre con el agente 2. Si ambos dicen la verdad, sólo se comprará el televisor cuando $v_1 + v_2 > 150$, que es lo óptimo.

Obsérvese que el agente *i* sólo tendrá que pagar si cambia la decisión colectiva. En este caso, decimos que el agente *i* es el agente bisagra. La cantidad que paga un agente bisagra es simplemente el coste que impone a los demás agentes.

37.11 Problemas del VCG

El mecanismo VCG lleva a decir la verdad y al nivel óptimo del bien público. Sin embargo, no está exento de problemas.

El primero es que sólo funciona cuando las preferencias son cuasilineales, ya que la cantidad que debe pagarse no tiene que influir en la demanda del bien público. Es importante que sólo haya un nivel óptimo único del bien público.

El segundo problema es que el mecanismo VCG no genera, en realidad, un resultado eficiente en el sentido de Pareto. El nivel del bien público es óptimo, pero el consumo derivado podría ser mayor debido a la recaudación del impuesto. Recuérdese que para que los incentivos sean los adecuados, las personas bisagra deben pagar, de hecho, algunos impuestos que reflejen el perjuicio que causan a las demás. Y estos impuestos no pueden ir a parar a ninguna de las personas implicadas en el proceso de decisión, ya que eso podría afectar sus decisiones. Los impuestos tienen que desaparecer del sistema. Y ése es el problema: si hay que pagar

los impuestos, el consumo privado termina siendo menor y, por lo tanto, es ineficiente en el sentido de Pareto.

Sin embargo, los impuestos sólo tienen que pagarse si una persona actúa de bisagra. Si la decisión afecta a muchas, la probabilidad de que una de ellas haga de bisagra puede o no ser muy grande, por lo que normalmente cabe esperar que la recaudación sea bastante pequeña.

El tercer problema del VCG es que es vulnerable a la colusión. Consideremos, por ejemplo, el problema de los bienes públicos que hemos descrito antes. Supongamos que en la subasta del televisor participan 3 compañeros de habitación, pero dos coluden. Los que coluden acuerdan en cada estado que el beneficio neto que obtienen con el televisor es de 1 millón de dólares. Eso garantiza que se comprará el televisor, pero como ninguno de los dos es el agente bisagra (es decir, ninguno de los dos agentes que coluden cambia la decisión), ninguno de los dos tiene que pagar el impuesto.

El último problema se refiere a la disyuntiva entre la equidad y la eficiencia inherente al mecanismo VCG. Dado que el sistema de pago debe fijarse de antemano, generalmente hay situaciones en las que el suministro del bien público empeora el bienestar de algunas personas, incluso aunque se suministre la *cantidad* eficiente en el sentido de Pareto. Decir que es preferible en el sentido de Pareto suministrar el bien es decir que *existe* un sistema de pago con el que todo el mundo disfruta de un mayor bienestar si se suministra el bien que si no se suministra. Pero eso no significa que con un sistema de pago *arbitrario* todo el mundo disfrute de un mayor bienestar. El impuesto de Clarke garantiza que si todo el mundo *puede* disfrutar de un mayor bienestar si se suministra el bien, se suministrará. Pero eso no significa que todo el mundo disfrute, de hecho, de un mayor bienestar.

Sería bueno que existiera un sistema que no sólo determinara si debe suministrarse o no el bien público, sino que también proporcionara un método eficiente en el sentido de Pareto para pagarlo, es decir, un plan de pago que mejorara el bienestar de todo el mundo. Sin embargo, no parece que exista ningún plan general de esas características.