

RACIONALIDAD INDIVIDUAL- IRRACIONALIDAD SOCIAL: EL CONFLICTO JUSTICIA-EFICIENCIA¹

1. Introducción

Esta charla se enmarca en una reunión de profesionales de Filosofía del Derecho donde se trata de abordar diversas cuestiones relativas al liberalismo. La idea de hablar de algún aspecto de esta cuestión desde el «punto de vista de los economistas» resulta a la vez atrayente y difícil. Confío en encontrar un lenguaje adecuado para transmitir algunas de las cuestiones que a nosotros nos preocupan en el análisis del liberalismo económico.

Los economistas tenemos la «costumbre» de valorar las decisiones de política económica tomando en consideración algunos elementos que tienen que ver con la eficiencia, o funcionamiento óptimo del sistema. Y entendemos tal eficiencia en dos aspectos distintos: a) *eficiencia económica* (nos preocupa que el sistema no desaproveche oportunidades), y b) *eficiencia informacional* (nos preocupa que las decisiones políticas puedan ser puestas en práctica a bajo coste).

Aparte de los dos elementos de eficiencia mencionados, es claro que hay otros dos elementos importantes que considerar a la hora de valorar las decisiones gubernamentales: c) *Respeto a las libertades individuales*, y d) *Justicia (o equidad)*.

La defensa a ultranza de la libertad individual subyace a las bases filosóficas del liberalismo económico. En este sentido, el liberalismo considera que el respeto a las libertades individuales es tal vez el más importante entre los valores mencionados, y el primero que hay que garantizar. La idea fundamental que defienden los liberalistas descansa en suponer *que nadie conoce mejor que el propio individuo qué es lo bueno para sí mismo*, y por tanto, la única misión

¹ Agradezco a Antonio Villar sus comentarios y sugerencias.

del Gobierno consiste en la defensa de los derechos individuales, sin realizar ningún tipo de intervención en la actividad económica.

Llevado este principio al funcionamiento de las economías, puede encontrarse que, *en determinadas circunstancias*, al preservar la libertad individual se obtienen (gratis) algunas otras buenas propiedades. En efecto, si las economías reales se ajustan bien al modelo competitivo, opera el llamado *Teorema de la Mano Invisible*, de Adam Smith, por el cual, las acciones individualistas de los diferentes agentes se coordinan de tal forma que el resultado final es *económicamente eficiente*. Por otro lado, este tipo de Institución (a la que denominaremos *Gobierno Mínimo*) realiza una política que resulta ser *informativamente eficiente*, ya que el Estado no necesita ninguna información para que el mecanismo de mercado funcione correctamente, pues cada agente, actuando de manera individual, va a conseguir un «buen» resultado colectivo. En estas condiciones, pues, con la sola defensa de la libertad individual, se obtiene eficiencia económica y eficiencia informacional, y la política liberalista pasaría correctamente tres de los cuatro puntos que considerábamos relevantes.

No obstante, hay dos elementos que oscurecen la aparente nitidez de la conclusión anterior. En primer lugar, el observar que el apelativo «óptimo social» (o socialmente eficiente), puede resultar engañoso. El concepto de eficiencia social se corresponde con la idea de que algo (accesible) es eficiente si no existe otra cosa (accesible) que todos los agentes prefieren. En este sentido, la eficiencia y la equidad (o algún sentido de justicia) pueden resultar contradictorias. Basta notar, por ejemplo, que una asignación en la que un solo individuo disfruta de todos los bienes, mientras que los demás no poseen nada es *socialmente eficiente*. Así, la política liberalista podría llevar a la sociedad a situaciones que resultasen muy malas desde el punto de vista de la justicia distributiva.

En segundo lugar, es evidente que la incuestionabilidad del razonamiento previo se rompe si, por alguna razón, el funcionamiento de mercado por sí solo no conduce a la economía a un resultado socialmente eficiente. Cuando esto sucede, los economistas hablamos de fallos de mercado, y entonces el Gobierno mínimo (liberalista) puede incumplir, además del criterio de justicia, el de eficiencia económica.

Las reflexiones anteriores conducen a considerar las posibles soluciones a los problemas mencionados, como una justificación para la creación de Gobiernos con mayor capacidad de intervención.

Consideremos en primer lugar el problema de la falta de equidad. En el contexto ideal en el que opera el Teorema de la Mano

Invisible (una economía perfectamente competitiva, sin fallos de mercado), se verifica el llamado *Segundo Teorema del Bienestar*, que asegura que cualquier asignación eficiente puede ser el resultado final del funcionamiento del mercado competitivo, para *una distribución específica de la riqueza*. En estas condiciones, si somos capaces de elegir una situación final socialmente deseable (eficiente y justa), ésta puede ser el resultado de la acción combinada de una redistribución de los recursos, seguida del libre funcionamiento del mercado. Supongamos entonces que el Gobierno posee la facultad de alterar la distribución de los recursos iniciales. Entonces, éste podría conseguir asignaciones finales *justas y económicamente eficientes*. Llamaremos *Gobierno Redistributivo* a este tipo de institución.

Por otra parte, en el caso de producirse fallos de mercado, se abre la posibilidad de otro tipo de actuación del Gobierno, encaminada a corregir dichos fallos, forzando la cooperación entre los agentes. Llamaremos a esta institución *Gobierno Reformista*, que podría garantizar asignaciones finales económicamente eficientes.

En esta charla intentaremos explicitar los límites de la conclusión liberalista, deteniéndonos fundamentalmente en el primero de los problemas apuntados. Es preciso aclarar bien qué es lo que se entiende por lo *bueno colectivamente*, y observaremos que este intento presenta grandes dificultades. Comenzaremos analizando las bases del comportamiento individual, en términos de *racionalidad*, para pasar de ahí a la definición de racionalidad colectiva. Nos limitaremos a los problemas derivados de las dificultades de agregación de las preferencias individuales, y a los problemas que se plantean cuando se pretenden conjuntar criterios de eficiencia y equidad. Una vez aclarado lo difícil que es llegar a definir un resultado final satisfactorio, comentaremos brevemente las posibilidades de buen comportamiento del Gobierno Redistributivo como una alternativa al Gobierno Mínimo.

A lo largo de la charla aparecerá de manera recurrente el problema de la puesta *en práctica* de las diferentes soluciones. No entraremos en ningún momento a analizarlo, puesto que ello será el objeto de otra de las exposiciones.

2. Racionalidad individual: aspectos descriptivo y normativo

En la Teoría Económica tradicional, los agentes económicos son considerados como *racionales*. El significado del apelativo «racional» se concreta en suponer que el agente, a) tiene objetivos perfectamente definidos, b) actúa con la única motivación de acercarse tanto

como pueda a dichos objetivos, y c) posee la habilidad para distinguir perfectamente cuál entre todas sus posibles líneas de acción es la que mejor se adapta a la consecución de sus objetivos.

Es bastante obvio que las consecuencias que se derivan de una aplicación estricta del concepto previo de racionalidad no puede esperarse que concuerden fielmente con el comportamiento observado de los agentes. Ello se debe a que, en general, los humanos están bastante lejos del modelo del *homo rationalis*: incluso si se da el caso de que tengan objetivos bien definidos, su motivación para alcanzarlos no suele ser completa; lejos de poseer infinita habilidad para calcular las mejores opciones, con frecuencia son punto menos que estúpidos, e incluso cuando son inteligentes, pueden estar cansados, hambrientos o distraídos, ser incapaces de razonar bajo presión, o dejarse guiar más por sus emociones que por su cerebro.

En consecuencia, no podemos esperar que la teoría económica sea *totalmente descriptiva* en el sentido en que lo son la física o la astronomía. Es muy frecuente encontrarse con patrones de comportamiento que no se pueden explicar mediante el paradigma del *hombre racional*. Un gran número de teorías alternativas a la racionalidad sustantiva, desde el trabajo de Simon en 1955 sobre «Racionalidad Acotada», han intentado compaginar la teoría y el comportamiento observado. Ésta es una de las direcciones en que la teoría económica se está desarrollando actualmente².

La observación anterior respecto de la inconsistencia aparente entre el modelo de hombre racional y el real pueden llevar a la conclusión errónea de que la teoría económica no va a tener ninguna relevancia a la hora de explicar el comportamiento de las economías reales. No obstante, esto no es así: procesos de selección y de aprendizaje llevan a los humanos a comportarse en multitud de ocasiones «como si» fueran racionales. Así como, por ejemplo, la teoría *biológica* de los juegos está resultando tremendamente acertada desde el punto de vista empírico, a pesar de que nadie supone que los animales se comportan racionalmente, sino que son las *fuerzas de la selección natural* las que señalan pautas de comportamiento de tipo optimizador, también aparecen procesos de selección en la vida económica³. Por otro lado, los *procesos de aprendizaje*, condicionan el comportamiento de los agentes, de tal forma que, cuanto más y más familiar resulta una determinada situación, los agentes van ajustando su comportamiento más y más al modelo racional *en dicha situación o en situaciones similares*.

² Para un ejemplo de los desarrollos recientes en esta dirección, puede verse Rubinstein (1990).

³ Véase Corchón (1990).

Las observaciones anteriores indican, pues, que dentro de determinados límites aún queda un espacio importante para encontrar aplicaciones descriptivas de la teoría económica⁴.

No obstante lo anterior, podemos decir que la visión ofrecida por la teoría económica resulta particularmente adecuada en los aspectos *normativos*, en el sentido de que la teoría prescribe que cada agente debería elegir aquella acción que resulte óptima para él, dada la información de que dispone.

Es importante señalar que las prescripciones anteriormente mencionadas se refieren al agente individual, con el objeto de preservar su comportamiento racional. No obstante, las diversas situaciones en que un individuo debe tomar decisiones son de naturaleza diversa, y puede muy bien ocurrir que, al seguir las indicaciones de la racionalidad, se obtengan resultados extraños o socialmente indeseables. Quizá sea interesante señalar algunas propiedades patológicas que aparecen cuando el proceso de decisión es *interactivo*, en el sentido de que los resultados obtenidos para cada agente no dependen sólo de su decisión, sino también de las decisiones de otros agentes. Un ejemplo de tales propiedades sería el que, cuando estamos ante un proceso de decisión interactivo, *aumentar las opciones no supone que los resultados finales vayan a ser más favorables*. Ello sucede, por ejemplo, en la situación siguiente, modelizada en forma de «juego» que, dadas las fechas en que nos movemos, sería oportuno identificar con la decisión de Hernán Cortés de quemar sus naves en el enfrentamiento con Moctezuma, lo cual le llevó a un resultado mejor que el que habría obtenido si hubiera mantenido abierta la opción de regresar⁵:

	β_1	β_2
α_1	5,5	10,4
α_2	4,-5	6,0

En el ejemplo anterior, la opción α_1 se puede identificar con «luchar sin entusiasmo», o si se prefiere con las espaldas cubiertas, por parte de Hernán Cortés, mientras que α_2 sería «luchar con entusiasmo». Por otra parte, las opciones β_1 y β_2 serían, para Moctezuma, «resistir hasta el límite», o no hacerlo. Dada la superioridad tecnológica de Hernán Cortés, los resultados de las combinaciones

⁴ Véase Selten (1985).

⁵ Véase Corchón (1991).

de las diferentes opciones pueden interpretarse fácilmente. Nótese que, desde el punto de vista de la racionalidad individual, la opción α_1 domina a la opción α_2 . Si Hernán Cortés elige α_1 , la mejor opción de Moctezuma es elegir β_1 , con lo que ambos obtienen el resultado de 5. No obstante, al quemar las naves, Hernán Cortés deshace la estrategia α_1 , y sólo deja abierta la opción α_2 , con lo que la mejor opción para Moctezuma pasa a ser β_2 y Hernán Cortés se garantiza un resultado mejor que antes, cuando tenía dos opciones a su disposición.

Otro resultado atípico lo constituye el hecho de que, con frecuencia, la aplicación del criterio de racionalidad individual conlleva ineficiencia social. Para ilustrar esto utilizaremos un ejemplo muy querido para los economistas, el llamado *dilema de los prisioneros*.

Dos hermanos, que tras las vicisitudes de una herencia, no se hablan, comparten una linde de sus terrenos, en la cual deberían construir una valla de separación. El coste de la valla es de 400.000 pesetas, mientras que cada uno de ellos valora el beneficio derivado de la existencia de la valla en 300.000 pesetas. Cada uno de los hermanos puede tomar una de las dos decisiones siguientes: 1) Contribuir al coste de construcción de la valla. 2) No contribuir. Los beneficios derivados de cada una de las cuatro opciones se pueden reflejar en la siguiente tabla:

	C	NC
C	1,1	-1,3
NC	3,-1	0,0

Si ninguno de los dos contribuye, la valla no se realiza. Si ambos contribuyen a partes iguales, la valla se hace y ambos obtienen un beneficio neto de 100.000 pesetas. Si uno de los dos construye la valla en solitario, éste experimenta una pérdida neta de 100.000 pesetas, mientras que el otro obtiene un beneficio neto de 300.000 pesetas.

Obsérvese que, haga lo que haga su hermano, la mejor opción de cada uno de los dos agentes es NC. Ello supone que la racionalidad individual llevaría a recomendar a cada uno de los agentes la no contribución. No obstante, este es un ejemplo en el que sólo hay una opción socialmente eficiente, que es aquella en la que los dos hermanos colaboran a la construcción de la valla. El problema que aquí aparece es el de la contradicción entre la eficiencia social y la racionalidad individual. Sin más, no hay forma de conseguir la coordinación necesaria para que la solución adoptada sea la que corresponde al óptimo social.

3. Racionalidad Colectiva. Preferencias Sociales

Hemos comentado previamente las implicaciones de la racionalidad individual en los procesos de decisión interactiva. La vertiente normativa de comportamiento individual se limita a prescribir a cada individuo su *mejor opción*, posiblemente a expensas de otro individuo, o a expensas de la eficiencia social. Pretendemos ahora analizar lo que debería ser la *racionalidad colectiva*, esto es, si miramos al grupo como un agente único, cuáles deberían ser desde el punto de vista social las prescripciones para obtener las *mejores opciones sociales*.

Es fácil ver que, en este caso, resulta inevitable un *proceso en dos etapas*: en primer lugar, se deben recomendar las acciones que resulten óptimas para el grupo, pero luego queda el problema de conseguir que cada uno de los individuos que componen el grupo encuentre individualmente razonable seguir el comportamiento prescrito para él, como miembro del colectivo.

Es importante señalar desde el principio que las dos etapas del proceso de decisión racional *para grupos* presentan muchas y muy graves dificultades. En primer lugar, aparece el problema de *determinar cuáles son las preferencias del grupo*, es decir, cuáles son las formas razonables de entender *lo que es bueno para el grupo*. En este punto aparecen algunas preguntas inmediatas cuya respuesta no es ni mucho menos obvia: (a) ¿tiene que existir relación entre lo bueno para el grupo y lo bueno para los individuos que lo componen?, (b) si los intereses de los individuos entran en conflicto, ¿cómo se debe decidir qué individuos van a marcar la pauta en las decisiones del grupo y cuáles deben ser relegados?, (c) si la *acción coordinada* del grupo permite obtener mejores resultados que los alcanzables individualmente, ¿cómo se deben repartir entre los componentes los beneficios de la cooperación?, (d) alternativamente, ¿debe definirse lo *bueno para el grupo* en términos de determinadas propiedades éticas?

Por su parte, en la segunda etapa nos encontramos con dificultades no menores. Supongamos que ya hemos definido exactamente qué es lo mejor para el grupo. Si la única forma de conseguir los resultados sociales es a través de las decisiones individuales, aparece el problema que los economistas denominamos *compatibilidad de incentivos*, es decir, cómo conseguir que los agentes, individualmente, decidan tomar aquellas acciones que conduzcan al óptimo social. Volvamos al ejemplo relativo al vallado de las parcelas contiguas mencionado en la sección anterior, y supongamos ahora que el Ayuntamiento decide que las parcelas deben estar debidamente separadas y fuerza a los dos hermanos a contribuir, mediante un impuesto para cada uno de 200.000 pesetas, a la construcción de la valla, que corre a cargo del propio

Ayuntamiento. En este ejemplo, esta Institución (el sector público local), mediante un decreto, constituye un mecanismo que permite alcanzar un resultado cooperativo, imposible de obtener mediante la pura decisión individualista de los agentes.

A lo largo de esta sección nos ocuparemos únicamente de la primera etapa del proceso mencionado. El concepto de eficiencia económica, en contra de lo que a primera vista puede parecer, en general no reduce drásticamente las opciones que pueden ser consideradas socialmente como buenas. Recordemos que una asignación era económicamente eficiente cuando no era posible encontrar otra asignación en la que todos los individuos estuvieran mejor. Así pues, si cualquier cambio en una asignación dada implica que alguien empeora, entonces resulta que tal asignación es socialmente eficiente desde el punto de vista económico. En consecuencia, el problema con que nos enfrentamos consiste en *aislar* una asignación específica, como aquella *socialmente más deseable* para cada sociedad particular. *Deseamos, en definitiva, poder recomendar a los grupos su mejor opción, de forma análoga a como la teoría de la racionalidad individual recomienda a los individuos su mejor opción.*

¿Qué significa seleccionar un óptimo social? Básicamente, dada una determinada sociedad (individuos con sus recursos, preferencias, etc., empresas con sus posibilidades de producción...), aislar un determinado resultado en base a un proceso específico de decisión colectiva. La vida real proporciona muchos ejemplos de procesos de decisión colectiva: toma de decisiones en base a la opinión de una mayoría especificada (como en los sistemas de asambleas o referéndum); elección por la colectividad de una determinada persona, a la que se otorga un voto de confianza de tal forma que las decisiones sociales son tomadas por esta persona concreta (como pueden ser un alcalde, o un director de Departamento); imposición por la fuerza de las opiniones de un individuo o grupo social, que son los que deciden (como en el caso de las dictaduras políticas), etc.

Nos enfrentamos, pues, al problema de diseñar un proceso de elección colectiva que, en definitiva, seleccionará en cada caso el «óptimo social» adecuado. Por consiguiente, el óptimo social será aceptado como tal si el proceso de decisión colectiva se diseña cumpliendo propiedades que sean admisibles para los individuos que componen la sociedad⁶. La aceptación de las propiedades de un proceso de decisión supone asumir a priori un determinado tipo de compromisos éticos. Nos centraremos aquí en aquellos procesos de decisión

⁶ El asumir a priori los requisitos mínimos que debe cumplir el proceso de elección colectiva es el único camino de romper la recursión al infinito de los acuerdos necesarios sobre la forma de agregar las preferencias, para obtener el óptimo social.

colectiva que seleccionan los resultados *tomando en cuenta las preferencias de los individuos que componen la sociedad de referencia*. Para este tipo de procesos vamos a considerar una serie de requisitos mínimos, que sería razonable exigir para que el proceso pudiera calificarse de «satisfactorio». Teniendo en cuenta que lo que se puede considerar razonable es un tema discutible, es útil pedir sólo condiciones muy débiles, universalmente aceptables, para no restringir demasiado las opciones plausibles. Por ello, vamos a buscar procesos de decisión que *ordenen socialmente todas las alternativas*⁷, y que, además sean:

1. *Generales*, esto es, aplicables a cualquier tipo de sociedad⁸.
2. *Respetuosos con la Unanimidad*, o dicho de otro modo, que verifiquen el Principio de Pareto: si todos los individuos prefieren la opción x a la y, la regla social debe preferir también x a y.
3. *Informacionalmente eficientes*, esto es, tener un bajo coste de puesta en práctica. Ello supone que deben funcionar con poca cantidad de información⁹.
4. *No dictatoriales*, es decir, no pueden limitarse a tomar en consideración las preferencias de un solo individuo¹⁰.

Arrow (1951) probó que las propiedades anteriores son conjuntamente inconsistentes. Esto es, *no existen procesos de elección colectiva que ordenen todas las alternativas y sean simultáneamente generales, respetuosos con la unanimidad, eficientes y no dictatoriales*. Este es, en resumen, el contenido del llamado *Teorema de Imposibilidad*¹¹. Ello supone que, a la hora de construir las preferencias sociales (o alternativamente, los índices de *bienestar social* de una determinada sociedad, en base a los cuales se tomarán las decisiones sociales), no podemos pretender preservar todas las

⁷ La racionalidad arrowiana se identifica con el hecho de que el resultado de agregar las preferencias individuales sea un orden social.

⁸ Esta propiedad es llamada por Arrow «dominio Universal».

⁹ La eficiencia arrowiana está vinculada a la propiedad conocida como «independencia de alternativas irrelevantes», cuyo espíritu consiste en asumir la no necesidad de información acerca de aquellas opciones que no se comparan directamente, a la hora de decidir socialmente entre varias alternativas.

¹⁰ La democracia arrowiana se identifica con la no existencia de ningún individuo cuyas preferencias sean las sociales (dictador).

¹¹ El enunciado del teorema especifica que no existe ningún agregador de órdenes individuales (o función del bienestar social) capaz de satisfacer simultáneamente las condiciones de dominio universal, principio de Pareto, no dictatorial e independencia de alternativas irrelevantes. Diversas pruebas del teorema, además de la original de Arrow, pueden verse en Kelly (1978), Suzumura (1983) y Sen (1986).

propiedades anteriores: hay que renunciar a alguna de ellas.

Ya era sabido desde muy antiguo que algunos métodos de combinar preferencias individuales para construir preferencias sociales conducían a inconsistencias. Condorcet (1785) ya observó problemas con la regla de decisión mayoritaria. Las inconsistencias de esta regla atrajeron la atención de pensadores tan pintorescos como C. L. Dogson, más conocido como Lewis Carroll. El caso más famoso es la llamada «paradoja del voto», presentada por Nanson (1882), que proporciona una buena introducción a la naturaleza del problema: Consideremos un ejecutivo que planea una determinada campaña publicitaria. En principio, hay tres opciones a considerar: A) un sistema de pequeños regalos; B) un sistema de descuentos, y C) un sorteo para un gran regalo. El ejecutivo encarga a tres técnicos en publicidad que le aconsejen sobre las diferentes alternativas. En el informe del primer técnico, se considera A) la mejor opción, seguida de B) y ésta de C); para el segundo técnico, la mejor opción es B), seguida de C) y ésta de A); finalmente, el tercer técnico considera preferida la opción C) a continuación A) y finalmente B). A la vista de estos informes, ¿qué alternativa debería elegir el ejecutivo? Nótese que en este ejemplo, si se pretende utilizar la regla de decisión mayoritaria, *no hay ninguna decisión posible. En efecto, la alternativa A es socialmente preferida a B, puesto que hay dos individuos (el primero y el tercero) que la colocan por delante; por otro lado, B es socialmente mejor que C, puesto que una mayoría de individuos (primero y segundo), la prefieren. Si la preferencia social fuera consistente, A tendría que ser socialmente preferida a C, pero lo que ocurre es que C es socialmente preferida a A, puesto que una mayoría de agentes (segundo y tercero), la prefieren.* En este ejemplo, el proceso de decisión mayoritaria no permite tomar ninguna decisión.

El interés del resultado anterior radica en el hecho de que la regla de decisión mayoritaria es muy atractiva como procedimiento de decisión social. En particular, verifica las condiciones de respeto a la unanimidad, eficiencia y no es dictatorial. Sin embargo, *no es general*. Es importante notar que el Teorema de Imposibilidad nos advierte de que *cualquier* procedimiento de decisión tiene que fallar en alguna de las propiedades especificadas.

Puede resultar ilustrativo el observar los fallos de alguna otra regla de decisión. Un método de decisión muy antiguo es la llamada *Cuenta de Borda*, por la cual se otorga un cierto número de puntos a la alternativa más preferida, una cierta cantidad (menor) a la segunda, y así sucesivamente. Luego resulta «socialmente» elegida la alternativa con más puntos. Este es un procedimiento de decisión muy usado en concursos literarios o festivos. Consideremos, de nuevo, el ejemplo de la campaña publicitaria, y supongamos que la

ordenación del primer técnico es: 1, 2, 3; la ordenación del segundo técnico es 2, 3, 1, y, finalmente, la ordenación del tercer técnico es 3, 1, 2. Si, al disponer de tres alternativas, otorgamos tres puntos a la más preferida, dos a la segunda, y uno a la tercera, en este ejemplo cada una de las alternativas recibe seis puntos, y el resultado es un empate, no una inconsistencia. Sin embargo, la Cuenta de Borda no es informacionalmente eficiente, pues es necesario conocer toda la ordenación de cada individuo para poderla aplicar.

Consideremos, finalmente, una regla de decisión social quizá absurda, pero no por ello ilusoria: supongamos que la preferencia social está totalmente determinada por un cierto código tradicional que ordena totalmente los estados sociales. Este tipo de regla satisface las condiciones de generalidad, eficiencia informacional, no dictatorial, pero no respeta la unanimidad, pues la ordenación social ha de seguir la prescrita por el código, *incluso si todos los individuos prefirieran otra diferente*. Un ejemplo reciente de este tipo de procedimiento podría aplicarse a la discusión en Italia respecto a la despenalización de la blasfemia, donde las voces en contra, curiosamente, surgen del Vaticano, un Estado diferente.

Los ejemplos anteriores sirven como ilustración del Teorema de Imposibilidad. La fuerza de este resultado radica en el hecho de que *siempre* podemos predecir el fallo de alguna de las propiedades, ante cualquier tipo de proceso de decisión social con que nos encontremos.

4. Dificultades con el concepto de Justicia o Equidad

El Teorema de Imposibilidad de Arrow señala las limitaciones que debemos imponer a los posibles procesos de decisión colectiva. No podemos imponer todas las condiciones anteriores, pues resultan inconsistentes. Emprendamos, entonces, un camino distinto. De entre las propiedades anteriores, vamos a conservar la más indiscutible: *el respeto a la unanimidad*, con lo que excluimos las reglas de «inspiración divina», y veamos si es posible completar esta propiedad con algún requisito ético del tipo justicia o equidad, de forma que obtengamos criterios de decisión colectiva operativos. La propiedad de respeto a la unanimidad implica que la asignación social elegida debe ser *eficiente económicamente, o en otras palabras, debe ser un Óptimo de Pareto*. En efecto, tomemos una asignación que no es un Óptimo de Pareto. Ello significa que existe otra posibilidad tal que algunos individuos mejoran, sin que nadie empeore. Para los primeros, la segunda opción es *igual de buena* que la primera. Por tanto, toda la sociedad declararía como preferida la segunda opción, y en consecuencia, la primera nunca puede ser declarada

como *la opción preferida socialmente*.

No obstante, el criterio de respeto a la unanimidad sólo permite comparar asignaciones en las que la preferencia de los individuos es unánime. El problema que aparece con este tipo de criterio es que *no compara* aquellas asignaciones en las que unos individuos mejoran y otros empeoran. Tomemos un caso en el que todos los individuos menos uno mejoran, mientras uno empeora: el criterio de respeto a la unanimidad tomado en solitario, no compara estas dos asignaciones, y por tanto, restringe muy poco las opciones que sí pueden ser consideradas como socialmente preferidas.

Queremos avanzar más, y ser capaces de discriminar entre los diferentes óptimos, quedándonos únicamente con aquellos que «pasen» algún test de justicia o equidad.

Una forma de construir test de equidad consiste en intercambiar los roles o las dotaciones de los diferentes individuos. La idea de colocarse en la posición de otro individuo aparece de diferentes formas en muchas culturas, aunque su empleo ha conducido a muy variadas consecuencias en las diversas sociedades. En la ley judaica se dice «haz con los demás lo que quisieras que los demás hiciesen contigo». La «ley moral» de Kant está muy relacionada también con esta idea de ponerse en la posición de los demás. En este sentido, merece la pena citar la interpretación de Sidgwick al enunciar su principio de «equidad»: *cualquier acción que cualquiera de nosotros juzgue buena para sí mismo, es implícitamente asumida como buena para cualquier otra persona similar en similares circunstancias*¹².

Un criterio de justicia relacionado en cierta forma con la idea anterior es el criterio de *ausencia de envidia*, de Foley (1967), que cualifica a una asignación como satisfactoria desde el punto de vista de la equidad cuando ningún individuo envidia la suerte de ningún otro individuo de la sociedad. De forma más precisa, una asignación posee la propiedad de ser libre de envidia si dada una pareja arbitraria de individuos de la sociedad, se cumple que ninguno de ellos desearía cambiar su asignación con la que tiene el otro.

Cuando se consideran economías en las que no hay producción, el concepto de ausencia de envidia es particularmente interesante, pues es fácil encontrar asignaciones que sean, simultáneamente, libres de envidia y eficientes¹³. La idea es muy sencilla: supongamos

¹² Véase Sidgwick (1907), Libro III, Cap. XIII. Sidgwick atribuye esta idea a Kant «lo que es bueno para mí debe ser bueno para cualquier otra persona en circunstancias similares que es el modo en que yo acepto la máxima kantiana- entiendo que es ciertamente fundamental, efectivamente cierto y no carece de importancia práctica».

¹³ Acerca de las asignaciones eficientes y libres de envidia en economías de intercambio, así como criterios relacionados, véase Thomson y Varian (1985).

inicialmente que en la economía sólo hay un bien, y dividamos este bien a partes iguales entre todos los individuos. Evidentemente, esta asignación es libre de envidia y eficiente (nadie envidia la suerte de nadie y no se desaprovecha ninguna oportunidad). Si ahora disponemos de diversos bienes, el proceso de asignar idénticas cantidades a todos los individuos, si bien continúa siendo libre de envidia ya no resulta eficiente, a no ser que todos los individuos tengan idénticos gustos. Si los individuos tienen gustos diversos, habrá alguna otra asignación que todos preferirán. Esta nueva asignación puede obtenerse dejando a los individuos intercambiar competitivamente sus bienes, obteniendo otra que ahora sí es eficiente y conserva el principio de ausencia de envidia, ya que los intercambios han sido voluntarios y por tanto nadie envidia la asignación de ningún otro agente.

El hecho de que, en economías de intercambio el criterio de ausencia de envidia resulte compatible con la eficiencia, no se extiende a las economías con producción. Consideremos el ejemplo siguiente¹⁴: dos agentes, por medio de trabajo, producen un bien. El salario conjunto de estos agentes se mide en términos de la cantidad de bien que han producido, esto es, si producen una unidad de bien, se repartirán un salario de una unidad monetaria. Por cada unidad de trabajo del agente 1 (digamos una jornada), se obtiene una unidad de bien. Sin embargo, el agente 2 es mucho más torpe, y necesita 10 jornadas de trabajo para producir una unidad de bien (o dicho de otro modo, con una jornada de trabajo produce 1/10 de bien). Así pues, si el agente 1 trabaja una proporción x de su jornada, y el agente 2 trabaja una proporción y de su jornada, generarán un «salario conjunto» de $x + y/10$ unidades monetarias. Cada agente dispone de una jornada de trabajo.

Por otra parte, la utilidad que el agente 1 deriva de cobrar un salario de a pesetas, junto con tener una proporción $(1 - x)$ de su jornada libre es de $u_1 = (1 - x)a + 2a$; la utilidad del agente 2 cuando cobra un salario de b pesetas, y tiene una proporción $(1 - y)$ de su jornada libre, es de $u_2 = (1 - y)b + 4b$.

Una asignación $\{(x, a); (y, b)\}$ será eficiente cuando no sea posible encontrar otra factible en la que alguien mejore sin perjudicar al otro. Nótese que, para que la asignación $\{(x, a); (y, b)\}$ sea eficiente, se tiene que cumplir que $x = 1$, puesto que si no, $\{(1 - \varepsilon, a); (y, b)\}$ se podría sustituir por $\{(1, a + \varepsilon); (y, b)\}$, que es mejor para el agente 1. Análogamente, $y = 0$, puesto que si no, $\{(1, a); (\delta, b)\}$ puede sustituirse por $\{(1, a); (0, b - \delta/10)\}$, que es mejor para el agente 2. Así pues, las posibles asignaciones eficientes son aquellas en que el

¹⁴ Este ejemplo está tomado de Pazner y Schmeidler (1974).

agente 1 trabaja toda la jornada, mientras que el agente 2 holgazanea toda la jornada. Dicho de otro modo, asignaciones del tipo $\{(1, a); (0, b)\}$, con $a + b = 1$.

Tomemos una asignación eficiente arbitraria. ¿En qué condiciones el agente 1 envidia al 2? Siempre que $2a < 1 + 2b$, esto es, siempre que $b > 1/4$. Por otro lado, el agente 2 envidiará al 1 siempre que $1 + 4b < 4a$, esto es, si $b < 3/8$. Por consiguiente, cuando $b < 1/4$, el agente 2 envidia al 1; si $1/4 \leq b \leq 3/8$, ambos agentes se envidian mutuamente, y si $b > 3/8$, el agente 1 envidia al 2. En este ejemplo, por tanto, no existen asignaciones que sean, simultáneamente, eficientes y libres de envidia. En términos más coloquiales, si el agente «poco productivo» cobra un salario bajo (aun sin trabajar), envidia al agente productivo. Si por el contrario, cobra un salario alto, entonces es el agente productivo el que le envidia.

Otra alternativa para interpretar la idea de *ponerse en la situación de otro*, consiste en suponer que el juicio acerca de la «bondad», desde el punto de vista ético de una determinada asignación se realiza suponiendo que a priori el individuo no sabe qué posición exacta es la que le tocará ocupar en la sociedad que está siendo analizada. Más precisamente, la idea sería la siguiente: supongamos que se trata de analizar si una determinada asignación es «justa» para una determinada sociedad. Tomemos entonces el caso de un individuo *que sabe va a ser miembro de dicha sociedad, pero no lo es todavía*. Este individuo se enfrenta a una lotería por la cual le tocará ocupar uno (indeterminado a priori) de los puestos de esta sociedad. No se trata ahora de intercambiar los roles de cada pareja de individuos, sino de analizar la sociedad desde el punto de vista de este *entrante hipotético*.

Uno de los criterios de equidad que contemplan las sociedades desde esta perspectiva es el criterio de justicia de Rawls (1958) (1963). La idea de Rawls consiste en defender como justos aquellos resultados que se derivan de una posición inicial de equidad. Dos son los tipos de principios que Rawls considera que se deben satisfacer en la «posición inicial»: a) *La máxima libertad para cada individuo, compatible con la máxima libertad para los demás, y b) las desigualdades sociales deben estar limitadas de forma que la sociedad consiga la mejor situación posible para el individuo con peor nivel de bienestar*.

Para que un criterio como el anterior sea operativo, tenemos que ser capaces de *comparar el bienestar de los individuos*. No podemos identificar «quién es el individuo que está peor» sin realizar este tipo de comparaciones. En este punto se plantea un problema de carácter algo espinoso: si el *bienestar individual* es algo subjetivo, ¿cómo se puede pretender establecer comparaciones de bienestar entre los

individuos? Si bien el supuesto de comparabilidad es un truco para poder construir una teoría operativa, es bien cierto que nada impide pensar que en un futuro los avances de la ciencia podrán crear un «felicidómetro»¹⁵, de forma que el nivel de bienestar individual se pueda correlacionar perfectamente con algún tipo de reacción cerebral, que pueda ser propiamente cuantificada.

El criterio de Rawls, pues, hace referencia a las comparaciones de bienestar entre los individuos, y opera de tal forma que ningún potencial entrante puede objetar la «buena fe» de esta sociedad, ya que ésta intenta conseguir que la peor posición sea lo mejor posible.

No obstante su indudable atractivo como criterio de justicia, el criterio de Rawls (*o criterio maximin*), plantea una serie de problemas que merece la pena comentar. En primer lugar, no está muy claro que se pueda considerar un criterio inapelable de equidad, puesto que, al mirar sólo el máximo bienestar posible del individuo que está peor, es ciego respecto del resto de la sociedad. Tomemos el siguiente ejemplo: consideremos una sociedad compuesta por tres agentes, de forma que el bienestar se mide en términos de renta. El agente A posee 100 millones de pesetas; el B posee 10 millones, y el C 500.000 pesetas. Tomemos ahora la situación siguiente, en la que A posee 109.300.000 pesetas, y tanto B como C poseen 600.000. Desde el punto de vista de Rawls, esta distribución es preferida a la primera, sin embargo, no se puede considerar muy satisfactoria desde el punto de vista de la equidad. En segundo lugar, el criterio de Rawls no contempla ningún tipo de trade-off, en el sentido de que prefiere una pequeña ganancia del individuo que está peor, antes que ganancias grandes de los demás individuos. Por ejemplo, consideremos el caso de una sociedad en la que el individuo que está peor vive en un lugar muy aislado, de forma que para incrementar su bienestar con algún tipo de servicio (digamos, ponerle teléfono), es necesario realizar una gran inversión. Alternativamente, con ese mismo dinero, podría construirse un parque público que incrementaría mucho el bienestar de todo el resto de la sociedad. El criterio de Rawls prima el pequeño incremento del único individuo aislado sobre el gran incremento del resto de la sociedad.

El criterio *utilitarista* puede ser defendido también sobre la base del *entrante hipotético*. Harsanyi (1955) defiende la idea de que las preferencias *éticas* de los individuos deben satisfacer un criterio de impersonalidad, en el sentido de que *el estado social elegido debe pasar el test de ser aceptable cuando el individuo no sabe cuál será su posición, ya que se enfrenta con igual probabilidad de ocupar cualquier de las posiciones de la sociedad*. La diferencia entre las interpretaciones

¹⁵ Véase, en este sentido, Ng (1983, Introducción).

de Harsanyi y Rawls radica en que Rawls no asigna probabilidades a las distintas posiciones de la sociedad, mientras que Harsanyi las considera equiprobables. Podríamos decir que la posición de Rawls es *la más pesimista posible*, en el sentido de que el entrante potencial se coloca como si le fuese a tocar ocupar la peor posición de la sociedad. Alternativamente, la posición de Harsanyi consiste en considerar todas las posiciones equiprobables, y suponer, además, que los individuos satisfacen los postulados de Von Neumann y Morgenstern de comportamiento ante el riesgo¹⁶. Así, las preferencias éticas se determinan mediante la maximización del bienestar esperado, y bajo el supuesto de equiprobabilidad, ello conduce a la maximización de la suma de utilidades individuales. Este criterio, de maximización de la suma utilidades individuales es normalmente conocido como *criterio utilitarista*, y de esta forma queda sustentado sobre un criterio de equidad, entendida como *impersonalidad*.

Analicemos un poco más las consecuencias del criterio utilitarista, para ser conscientes de si es o no satisfactorio como criterio de equidad. Los ejemplos siguientes nos pueden dar una pista de sus posibles dificultades. Consideremos una sociedad esclavista compuesta por 99 hombres libres y un esclavo. Este último sirve a todos los demás, y es terriblemente desgraciado. Dada idéntica probabilidad de estar en cualquiera de las posiciones de esta sociedad, es posible que algunos estuvieran dispuestos a tomar un riesgo de un 1 por 100 de ser el esclavo, dado que existe un 99 por 100 de opciones de no serlo. En este sentido, podría llegar a concluirse que una sociedad esclavista puede ser moralmente aceptable. Seguramente que mucha gente no estaría de acuerdo con este tipo de test. Otro ejemplo que sirve para observar las dificultades que presenta el criterio utilitarista como criterio de equidad sería el siguiente: supongamos que el bienestar de los individuos se identifica con su renta. Entonces, el criterio utilitarista asocia a cada sociedad como índice de bienestar, su renta per cápita. Ello significa que, dadas dos posibilidades, el criterio utilitarista prefiere aquella con renta per cápita más alta. Así, si consideramos las dos posibilidades siguientes: 1) la mitad de la población a renta 100, y la otra mitad a renta cero; 2) todos los individuos de la población a renta 49, el criterio utilitarista declara preferida 1), en contra de 2).

Los casos mencionados ponen de manifiesto que definir un criterio de justicia universalmente aceptable es un problema sin solución. Todos los criterios presentados poseen propiedades que los hacen atractivos desde el punto de vista de la equidad, pero ninguno de ellos está exento de problemas. La elección, por tanto, entre uno u otro es un tema abierto a la voluntad de cada grupo, sociedad o planificador.

¹⁶ Véase Von Neumann Morgenstern (1944).

No hay razones objetivas incuestionables que lleven a preferir un criterio sobre otro, o una función de bienestar social específica sobre otra diferente. La elección es, simplemente, un tema de preferencias de unas determinadas propiedades éticas sobre otras, y, en general, la posición más sensata sería aquella que consiste en elegir un determinado criterio de equidad para cada tipo de problema, sin pretender encontrar uno universalmente aceptable.

5. Conclusiones

Hemos repasado a lo largo de las secciones anteriores las dificultades que se plantean a la hora de seleccionar *la mejor opción desde el punto de vista social*. Supongamos, no obstante, que tenemos un criterio de justicia aceptado unánimemente por la sociedad, y que somos capaces de seleccionar la situación final *eficiente y justa* que deseamos para nuestra sociedad. Imaginemos, además, que nos encontramos en una economía *sin fallos de mercado*.

Sabemos entonces que, todo lo que hay que hacer para conducir a la sociedad a la situación deseada es *redistribuir los recursos entre los agentes, y luego dejar funcionar a cada uno según su propio interés*. Este era el cometido del gobierno que hemos llamado «redistributivo». Es importante señalar que la redistribución de los recursos sólo puede realizarse si el gobierno conoce cuáles son los recursos que cada individuo posee. Por tanto, para poder realizar su tarea redistributiva, el gobierno necesita información que en principio sólo poseen los agentes individuales. Esta necesidad de información rebaja considerablemente la eficiencia informacional del Gobierno Redistributivo respecto del Gobierno Mínimo.

En estas condiciones, si quisiéramos enfrentar el comportamiento de los gobiernos mínimo y redistributivo, en un mundo sin fallos de mercado, ante los cuatro tipos de valores que mencionábamos en la Introducción, encontraríamos el siguiente esquema:

	Libertad	Justicia	Ef. Económ.	Ef. Inform.
Gobierno mínimo	B	M	B	B
Gobierno Redistributivo	R	B	B	R

Para terminar, dos comentarios. En primer lugar, la clasificación anterior no resulta tan nítida para el Gobierno Redistributivo. Es bien sabido que en la vida real, los gobiernos no realizan redistribuciones de recursos, sino que lo que hacen es colocar un sistema

redistributivo posterior al funcionamiento del mercado, esto es, un sistema de impuestos y transferencias. Este tipo de redistribución, aunque con menores requerimientos informacionales, conduce a la economía a un subóptimo¹⁷, por lo que las ganancias en equidad se hacen, de alguna manera, a costa de pérdidas en eficiencia económica. En segundo lugar, la clasificación anterior está realizada en el supuesto *de que no haya fallos de mercado*. Si los hay, ambos tipos de Gobierno fallan en la obtención de la eficiencia económica¹⁸.

Finalmente, comentar que no hemos mencionado para nada los aspectos relativos a los problemas de la *puesta en práctica*, que complican aún más el realizar una valoración de los distintos sistemas gubernamentales en los cuatro aspectos mencionados.

BIBLIOGRAFÍA

- Arrow, K. J., *Social Choice and Individual Values*, Yale U. Press, New Haven, 1951.
- Arrow, K. J. y Intriligator, M. D. (eds.), *Handbook of Mathematical Economics*, Amsterdam, North Holland, 1986.
- Arrow, K. J. y Honkapohja (eds.), *Frontiers of Economics*, Oxford, Blackwell, 1985.
- Auerbach, A. J. y Feldstein, M. (eds.), *Handbook of Public Economics*, vol. II, Amsterdam, North Holland, 1987.
- Auman, R. J., «What is Game Theory trying to Accomplish?», en Arrow & Honkapohja (eds.), 1985.
- Condorcet, M., *Essai sur l'Application de l'Analyse a la Probabilité des Decisions Rendues a la Pluralité des Voix*, París, 1785.
- Corchón, L. C., «On Natural Selection in Oligopolistic Markets», *A Discusión N.º 32*, U. Alicante, 1990a.
- Corchón, L. C., «Una Introducción a los Juegos en dos etapas o por qué Hernán Cortés quemó sus naves y salió ganando con ello», *De Economía Pública*, N.º 9:3-26, 1990b.
- Foley, D., «Resource Allocations in The Public Sector», *Yale Economic Essays*, N.º 7:73-76, 1967.
- Hurwicz, L.; Schmeidler, D. y Sonnenschein, H. (eds.), *Social Goals and Social Organization*, Cambridge, Cambridge U. Press, 1985.
- Inman, R., «Markets, Government and the Political Economy», en Auerbach & Feldstein (eds.), 1987, pp. 674-777.
- Kelly, J. S., *Arrow Impossibility Theorems*, New York, Academic Press, 1978.
- Nanson, E. L., «Methods of Elections», *Transactions and Proceedings of the Royal Society of Victoria*, 18, 1882.
- NG, Y. K., *Welfare Economics*, Hong Kong, McMillan Press, 1983.

¹⁷ Véase Stiglitz (1982).

¹⁸ Para un análisis más detallado, véase Inman (1987).

- Pazner, E. y Schmeidler, D., «A Difficulty in the Concept of Fairness», *Review of Economic Studies*, 41:441-443, 1974.
- Rawls, J., «Justice as Fairness», *Philosophical Review*, 67, 1958.
- Rawls, J., «The sense of Justice», *Philosophical Review*, 72, 1963.
- Rubinstein, A., «New Directions in Economic Theory-Bounded Rationality», *Revista Española de Economía*, 7:3-15, 1990.
- Selten, R., «Comments to Auman's paper», en Arrow & Honkapolija (eds.), 1985.
- Sen, A., *Collective Choice and Social Welfare*, Amsterdam, North Holland, 1970.
- Sen, A., «Social Choice Theory», en Arrow Intriligator (eds.), vol. III: 1073-1181, 1986.
- Sidwick, H., *The Method of Ethics*, London, McMillan, 1907.
- Suzumura, K., *Rational Choice, Collective Decisions and Social Welfare*; Cambridge, Cambridge U. Press, 1983.
- Thomson, W. y Varian, A., «Theories of Justice based on Symmetry», en Hurwicz, Schmeidler y Sonnenschein (eds.), 107-130, 1985.
- Von Neumann, J. y Morgenstern, O.; *Theory of Games and Economic Behavior*, Princeton, Princeton U. Press, 1944.