

AMARTYA SEN

FUNCIONES DE BIENESTAR SOCIAL NO LINEALES: UNA RÉPLICA AL PROFESOR HARSANYI

En primer lugar, una cuestión preliminar. La respuesta de John Harsanyi (1977) se ocupa exclusivamente de una parte relativamente pequeña de mi conferencia (1977), a saber, aquélla que trata de la crítica de Harsanyi a las funciones de bienestar social no lineales. Todos los resultados formales —por ejemplo, del (T.1) al (T.4)— y gran parte de la discusión informal de mi conferencia tenían que ver con otros asuntos. Mientras que encuentro interesante e importante la crítica de Harsanyi a las funciones de bienestar social no lineales y agradezco esta oportunidad de discutir más a fondo por qué estoy en desacuerdo con él, me gustaría ciertamente evitar que se diera la impresión de que Harsanyi y yo discrepamos sobre lo esencial de mi conferencia.

Pasemos ahora a los argumentos de Harsanyi. Su primer punto es que, puesto que yo apoyo el rechazo de Peter Diamond (1967) a la invulnerabilidad ética del ‘axioma de independencia fuerte’ aplicado a la elección social, se deduce que debo aceptar ciertas ‘curiosas implicaciones de los argumentos de Diamond’ según Harsanyi. Por ejemplo, debo aceptar que “las desigualdades sociales y económicas hereditarias más extremas serían moralmente *menos censurables* cuando los recién nacidos fueran regularmente intercambiados entre diferentes familias por una gran lotería gubernamental” (Harsanyi, 1977, p. 294; p. 104 de esta ed.). Debo confesar que no me convence este ejemplo, al que llamaré —por mor de la brevedad— LH (‘leyenda de Harsanyi del

intercambio de niños por lotería'*). Para rechazar el axioma de independencia fuerte, es suficiente demostrar que *existen* circunstancias en las que las implicaciones del axioma deberían ser negadas; no hay ninguna obligación de negar las consecuencias de ese axioma en *toda* circunstancia concebible. Lo que se requiere es una negación de una ley universal y eso puede basarse en una afirmación existencial. Por el contrario, una indicación existencial en un ejemplo (LH, por ejemplo) que satisfaga el axioma para una circunstancia particular no establece la regla. Para el tipo de 'imparcialidad' que se requiere en la concepción de una 'posición original', el 'axioma de independencia fuerte' parece ser bastante objetable; el ejemplo de Diamond era apropiado para este caso. El hecho de que podamos decidir no violar el axioma de independencia fuerte en el caso LH, no restituye el axioma como una regla de dominio universal.

El resumen que hace Harsanyi de mi posición con respecto al axioma de independencia fuerte, a saber, que "no puede aplicarse legítimamente a las elecciones entre políticas sociales alternativas" (Harsanyi, 1977, p. 293; p. 104 de esta ed.), no es totalmente exacto. Si el axioma tuviera *algunas* implicaciones legítimas en circunstancias específicas pero tuviera que ser violado en otras, entonces la defensa axiomática de Harsanyi del utilitarismo también se vendría abajo, resultando ilegítima su conclusión de que el bienestar social "debe ser definido como la media aritmética de los niveles de utilidad de todos los individuos de la sociedad" (Harsanyi, 1975, p. 312; p. 57 de esta ed.).

A propósito, debería observarse que el ejemplo LH está muy inteligentemente escogido, ya que pone en juego la repulsión innata que mucha gente sentiría ante la idea de redistribuir los recién nacidos mediante una lotería, aun cuando ello no implica más que una permutación del vector de utilidad individual entre las personas implicadas. En este contexto, probablemente vale la pena señalar que esta repulsión es también difícil de acomodar dentro de un marco utilitarista. El utilitarismo no recomendará el rechazo definitivo del modelo LH que conduce a una permutación de las utilidades individuales, sino que meramente recomendará una *indiferencia* sobre si deberíamos aceptarlo o no. Por lo tanto, el ejemplo LH, tal y cómo lo caracteriza Harsanyi, puede resultar

* Sen llama al argumento de Harsanyi del intercambio de recién nacidos por una lotería con las siglas "HELL" de "Harsanyi's exchange-baby lottery legend". "Hell" es una de las palabras inglesas habituales para "infierno", connotación que se pierde en la traducción que hacemos (N. del. E.).

‘excesivo’. De hecho, yo me inclinaría a argumentar que los enfoques ‘consecuencialistas’ orientados al bienestar (incluyendo al utilitarismo entre otros) con funciones lineales o no lineales del bienestar social, no pueden realmente captar los matices de esta repulsión innata ante la idea de redistribuir recién nacidos por medio de una lotería.¹

El segundo punto de Harsanyi se refiere al caso en el que se aceptan tanto el axioma de independencia fuerte como el conjunto completo de axiomas de von Neumann-Morgenstern para las elecciones *sociales* que implican riesgo. Yo había argumentado que incluso entonces el utilitarismo podría no ser aceptable excepto en un sentido trivial. Harsanyi parece ignorar la mayoría de los argumentos presentados y malinterpreta el que él elige comentar.

Empezando con las utilidades *individuales*, éstas no tienen por qué tener siempre representaciones cardinales como tampoco tienen que ser cardinalmente comparables entre las personas. Harsanyi no considera este caso, a pesar de la evidencia de que con la mera ordinalidad y la comparabilidad de niveles la regla rawlsiana funciona completamente —como muchas otras reglas, aunque no el utilitarismo. La axiomática considerada en mi conferencia se basaba en no suponer nada más que esto (véanse, por ejemplo, los teoremas (T.1) — (T.4)). Dado que Harsanyi pone todos estos procedimientos de agregación no utilitaristas bajo el ancho sombrero de “no linealidad” —de manera engañosa, en mi opinión— no es absurdo plantear la cuestión de la mensurabilidad y la comparabilidad al contrastar los méritos de tales reglas no utilitaristas con el utilitarismo.

Cuando el comportamiento individual satisfaga los axiomas de von Neumann-Morgenstern, existirá, por supuesto, una representación cardinal de la función de utilidad individual; a estos valores podemos llamarlos valores *V*. Estos [...](**) representación puede, no obstante, coexistir con la representación de von Neumann-Morgenstern. Aunque estos valores *V* *predecirán* el comportamiento de los individuos bajo riesgo, la cuestión de la mayor relevancia de los valores *V* para la elección social, por encima de otras representaciones, es un asunto completamente diferente. Yo he hecho la modesta afirmación de que “cuando alguien se refiere al bienestar

¹ Para una crítica al utilitarismo, véase Williams (1973).

** El espacio punteado entre corchetes indica una laguna en el texto original que tiene la extensión de una línea; dicha línea en el texto original ha sido suplantada por error tipográfico por otra que aparece unas cuatro líneas más abajo (N. del E.).

social como una función no lineal de bienestar individuales, no tiene por qué referirse necesariamente a los valores V en absoluto” y “no hay ninguna obligación de referirse sólo a los valores V cuando se habla del bienestar social o individual” (Sen, 1977, p. 277; p. 86 de esta ed.). Por razones algo oscuras, Harsanyi lo interpreta como que yo recomendaba que “deberíamos usar una terminología que consistiera en *disfrazar* cuidadosamente el hecho de que somos utilitaristas” y que “deberíamos negarnos a *identificar* una función de utilidad de vNM de un individuo con nuestra medida para su bienestar personal” (Harsanyi, 1977, p. 294; pp. 104-5 de esta ed.). Nada de eso. El hecho es que, al igual que en otras representaciones cardinales, hay elementos de arbitrariedad en el sistema von Neumann-Morgenstern —que está diseñado específicamente para *predecir* el comportamiento bajo incertidumbre— y el asunto no es de ‘disfrazar’ nada, sino el de advertir que cuando decimos que el bienestar social es una función no lineal de los bienestar individuales, *de hecho* nos estamos refiriendo típicamente a representaciones que no son las que surgen del sistema von Neumann-Morgenstern.²

Harsanyi tiene dos justificaciones esencialmente diferentes para emplear la forma lineal para las funciones de bienestar social. Una se basa en su teorema de que si tanto el bienestar social como el individual satisfacen los axiomas de von Neumann-Morgenstern y si la indiferencia de Pareto se refleja en la indiferencia social, entonces el bienestar social (medido en valores V) debe ser una combinación lineal de los bienestar individuales (medidos en valores V). A esto lo llamaré la ‘justificación de consistencia’. La otra se basa en su concepción de una posición *como si* de igualdad primordial, en la que las personas tienen que elegir entre estados sociales sin conocer sus propias posiciones en los mismos, *suponiendo* que tienen la misma probabilidad de ser cualquier persona. Si en esta posición *como si* se supone que las personas se comportan de acuerdo con los axiomas de von Neumann-Morgenstern, entonces debe suponerse que acaban maximizando las utilidades esperadas de las diferentes loterías; y dado que los ‘premios’ son el ser individuos concretos, esto equivaldría a maximizar la utilidad media de todos los individuos juntos (medidos en valores V). A esto lo llamaré la “justificación *como si*”. Ambos enfo-

² Se analizan enfoques alternativos a la medición y comparación de los bienestar individuales en mi “Interpersonal Comparisons of Welfare”, artículo preparado para un homenaje a Tibor Scitovsky [en: M. Boskin (ed.), *Economics and Human Welfare. Essays in Honor of Tibor Scitovsky*, New York: Academic Press, 1979].

ques son merecidamente célebres. De hecho, Harsanyi tiene derecho a sentirse algo agraviado porque su concepción de la incertidumbre *como si* ha sido a veces olvidada debido al entusiasmo que siguió a la publicación de *A Theory of Justice* de Rawls con su concepción de la 'posición original'.

La 'justificación de consistencia' incorpora más estructura y conlleva dificultades sobre las representaciones del bienestar individual discutidas anteriormente. Además, no nos da el utilitarismo como tal, sólo la linealidad. La 'justificación *como si*' es más directa. Es, de hecho, 'ineludible' con tal de que aceptemos las premisas, a saber, la equi-probabilidad *como si*, el comportamiento *de vNM* e intentemos relacionar los valores V del bienestar social según los considera una persona *determinada* con los valores V de las posiciones individuales según los considere la *misma persona* bajo incertidumbre *como si*. La utilidad de la lotería debe necesariamente ser igual a la suma de la utilidad —ponderada por la probabilidad— de los 'premios' de ser individuos diferentes.

Por supuesto, el modelo 'como si' es extremadamente interesante e importante desde el punto de vista de establecer un requisito ético de 'impersonalidad', pero la forma lineal que produce es esencialmente trivial. $W = \sum_i (1/n) W_i$ se obtiene del hecho de que W es la utilidad de la lotería y los W_i son los premios con igual probabilidad. Las actitudes frente al riesgo ya se han incluido en las medidas W_i . Habría que interpretar de un modo completamente erróneo la estructura para proponer que bajo los axiomas de von Neumann-Morgenstern la utilidad de la lotería difiriese en algo de la utilidad esperada de ella. El resultado es muy poco controvertido, ya que es un "placer que cuesta tan poco" (Sen, 1977, p. 279; p. 88 de esta ed.). A Harsanyi le podría divertir el hecho de que *en este sentido* "el planteamiento de Sen demuestra (su) argumento" y que "no podríamos evitar el seguir el enfoque utilitarista" (Harsanyi, 1977, p. 295; p. 105 de esta ed.); pero el que no sea controvertido surge esencialmente de que no afirma nada.

Debo, por lo tanto, atenerme a mi conclusión de la conferencia de que "una justificación axiomática del utilitarismo tendría más contenido si arrancara de un lugar algo más distante del destino último" (p. 279; p. 89 de esta ed.). Por fortuna, recientemente se han propuesto axiomatizaciones bastante más reveladoras del utilitarismo por diversos autores —por ejemplo, d'Aspremont y Gevers (1975), Hammond (1975), Maskin (1975), Deschamps y Gevers (1976)— que nos permiten debatir los asuntos subyacentes en vez de hacer que el utilitarismo sea 'imposible de evitar' en un

sentido trivial. La contribución de Harsanyi (1955) al proponer su modelo *como si* fue extremadamente valiosa por lo que nos enseñó sobre el requisito de 'impersonalidad' de los juicios éticos. Como alguien cuya comprensión de estos asuntos estaba muy influida por el enfoque de Harsanyi, me entristece que Harsanyi continúe creyendo que su contribución consistía en proporcionar una justificación axiomática del utilitarismo con contenido real.

BIBLIOGRAFÍA

- D'ASPREMONT, C. and GEVERS, L.: 1975, "Equity and the Informational Basis of Collective Choice", presentado en el Tercer Congreso Mundial de Econometría, mecanografiado, en prensa en *Review of Economic Studies* [44 (1977), 199-210].
- D'ASPREMONT, C. and GEVERS, L.: 1976, "Leximin and Utilitarian Rules", mecanografiado, Faculté des Sciences Economiques et Sociales, Namur, 1976 [pub. como "Leximin and Utilitarian Rules: A Joint Characterization", *Journal of Economic Theory*, 17 (1978), 143-163].
- DIAMOND, P.: 1967, "Cardinal Welfare, Individualistic Ethics and Interpersonal Comparisons of Utility: A Comment", *Journal of Political Economy*, 61, 765-766.
- HAMMOND, P. J.: 1975, "Dual Interpersonal Comparisons of Utility and Welfare Economics of Income Distribution", mecanografiado, Essex University [*Journal of Public Economics*, 6 (1977), 51-71].
- HARSANYI, J. C.: 1955, "Cardinal Welfare, Individualistic Ethics and Interpersonal Comparisons of Utility", *Journal of Political Economy*, 63, 309-321.
- HARSANYI, J. C.: 1975, "Nonlinear Social Welfare Functions", *Theory and Decision*, 6, 311-332 [trd. cast. en este volumen como "Funciones no lineales de bienestar social: ¿tienen los economistas del bienestar una exención especial de la racionalidad bayesiana?"].
- MASKIN, E.: 1975, "A Theorem on Utilitarianism", mecanografiado, Cambridge University [*Review of Economic Studies*, 45 (1978)].
- SEN, A. K.: 1977, "Welfare Inequalities and Rawlsian Axiomatics" [trd. cast. en este volumen como "Desigualdades de bienestar y axiomática rawlsiana"].
- SMART, J. J. C. and WILLIAMS, B.: 1973, *Utilitarianism. For and Against*, Cambridge University Press [trd. cast. *Utilitarismo: pro y contra*, Madrid: Técnos, 1981].
- WILLIAMS, B.: 1973, "A Critique of Utilitarianism", en Smart and Williams (1973).