

Economía, evolución y teoría de juegos: redefiniendo al homo economicus

Laureano Castro Nogueira / Miguel Ángel Toro Ibáñez

ECONOMÍA Y EVOLUCIÓN: ANTECEDENTES HISTÓRICOS E INTERACCIONES MUTUAS

La influencia del pensamiento económico sobre el evolucionista se remonta a la propia génesis de la teoría darwinista. Sin duda, un factor clave en el desarrollo del concepto de selección natural fue la obra de Thomas Robert Malthus *Essay on the Principle of Population*, de donde extrajo Darwin la idea de la lucha por la existencia. La tesis de Malthus afirmaba que mientras las poblaciones crecen siguiendo una progresión geométrica, los alimentos lo hacen según una progresión aritmética. Esto provoca una lucha intrapoblacional por los recursos. Los historiadores han discutido hasta la saciedad qué fue exactamente lo que conmocionó a Darwin de la lectura de Malthus. Parece que uno de los elementos importantes fue el cambio de perspectiva que supone dar prioridad a la competición entre los individuos de una especie frente a la de la competición interespecífica. Cuando una población posee recursos abundantes crece rápidamente, pero, a medida que se expande, los recursos se convierten en factores limitantes y el censo de la población se estabiliza. Se alcanza así un equilibrio que surge como una consecuencia no intencionada del conflicto de intereses entre unidades individuales.

Lo que Darwin descubrió leyendo a Malthus es que este equilibrio es más aparente que real. En la *teología malthusiana*, el equilibrio no da lugar a una situación óptima, sino que continúa la lucha, la maldad y el sufrimiento. La moral no sólo sigue siendo necesaria, sino que adquiere un significado pleno como el instrumento que nos fortalece en la lucha por el bien en esa disputa permanente a la que conduce el equilibrio. De todas formas, no fueron los aspectos morales los que impresionaron a Darwin, sino los que podríamos llamar aspectos ecológicos aplicables a la población. Se percató de que en todas las fases de la curva de crecimiento de una población, incluido el equilibrio, los individuos no utilizan los recursos de forma simétrica. Existe siempre una competición que se traduce en que unos individuos contribuyen más que otros a la siguiente generación. Esta idea, que hoy nos parece de ecología elemental, no se percibía así en los tiempos de Darwin. La idea que prevalecía en economía sostenía que se alcanzaba un equilibrio entre los recursos y el censo poblacional y que entonces nada cambiaba; esto es, la competencia no conducía a un cambio continuo, sino a una situación estática.

Las ideas de Malthus sobre la importancia de los conflictos y la competencia individual en las poblaciones permitieron a Darwin encontrar una respuesta completamente diferente para explicar la presencia de diseño en los seres vivos: la lucha por la supervivencia –la selección natural darwiniana– podía generar la adaptación. Las ideas de Darwin encontraron una enorme acogida en el campo de las ciencias sociales británico, que se plasmó con el desarrollo del denominado darwinismo social de

Herbert Spencer. Este movimiento postulaba como algo natural de nuestra especie una lucha por la existencia que conduciría al triunfo de los más aptos, definidos tautológicamente como aquellos que alcanzan un mayor éxito social y económico. Las diferencias de clase entre individuos y grupos se explicaban como el producto de sus diferencias en aptitud, apoyando y justificando así las ideas del liberalismo económico; desde entonces, este movimiento ha sido inspiración de aquellos que ansían encontrar la superioridad biológica de unos grupos sociales sobre otros. El justo rechazo que este tipo de teorías ha merecido por parte de los sectores más progresistas de la comunidad científica y de la sociedad en general ha provocado también, como un efecto no deseado que se extiende en el tiempo hasta la actualidad, un distanciamiento de la cultura humanista con respecto a la biología evolutiva. Por otra parte, la aguda crisis que experimentó el concepto de selección natural tras la muerte de Darwin atenuó la influencia del liberalismo económico en la evolución biológica e impidió a su vez que el darwinismo ejerciera un impacto importante sobre el ámbito cultural y económico.

A partir de 1930, cuando se produjo la síntesis neodarwinista, la selección natural volvió al primer plano de la teoría evolutiva, pero se mantuvo una nítida separación entre lo biológico y lo cultural. El neodarwinismo asume que el extraordinario potencial del cerebro humano ha permitido que nuestra especie haya alcanzado un grado de desarrollo cultural que nos independiza en gran medida de nuestra biología. Los cambios culturales son tan rápidos y sus efectos sobre la conducta tan poderosos que, en la práctica, anulan la variabilidad genética subyacente. Sólo en los últimos años, este planteamiento ha sido cuestionado por la sociobiología y, más recientemente, por la psicología evolucionista. Estas nuevas disciplinas intentan conciliar la autonomía de los procesos culturales con la existencia de importantes condicionantes biológicos.

Paralelamente, el pensamiento económico liberal comenzó a gestarse en los siglos XVIII y XIX a partir de las ideas de Adam Smith, David Ricardo y Stuart Mill, que dieron lugar a la llamada teoría económica clásica. Más tarde fue consolidando sus rasgos a lo largo de casi un siglo, entre 1860 y 1950. La corriente principal de la economía adoptó los modelos de la mecánica newtoniana porque esta disciplina se consideraba el arquetipo de lo científico y trató de alcanzar un estatus epistemológico entre las ciencias sociales similar al que ocupaba la física en las naturales. Surgió así la concepción neoclásica de la economía que ha sido el paradigma ortodoxo dominante, fuera del ámbito marxista, durante todo el siglo XX en las universidades y organismos económicos internacionales. Los presupuestos básicos de dicha teoría se traducen en lo que se denomina individualismo metodológico, que lleva a interpretar los fenómenos económicos a partir del comportamiento individual de los agentes implicados. En otras palabras, la idea de que el orden de la naturaleza y de la sociedad puede surgir como una consecuencia, no intencionada, de las interacciones y del conflicto de intereses entre unidades individuales.

El énfasis por explicar los sucesos económicos a partir de la conducta individual hizo que, a mediados del siglo pasado, se aplicasen a la conducta económica los principios de la teoría de juegos, en la cual cada individuo implicado se considera que actúa como un ente racional, un *Homo economicus*, dotado de una ilimitada capacidad para el análisis de las situaciones y empeñado en maximizar los beneficios de su conducta, en un ambiente en que los parámetros básicos se supone que son lo suficientemente conocidos como para evitar que los sucesos imprevisibles desempeñen un papel

relevante. Este enfoque de la teoría de juegos ha tenido también una aplicación muy valiosa en otras disciplinas, sobre todo, en el estudio de la conducta animal desde una perspectiva evolutiva. Obviamente, los organismos no se comportan como seres capaces de elegir la opción más racional, pero la variabilidad genética que subyace detrás de un determinado grupo de conductas puede permitir que la selección natural *descubra* la opción más favorable en términos de eficacia reproductiva. Los hallazgos teóricos en este campo han sido importantes y han contribuido de manera decisiva a la formación del concepto moderno de evolución y a la consolidación de la teoría evolutiva como el paradigma central de la biología.

Curiosamente, este pensamiento evolutivo moderno viene siendo utilizado desde 1980 por un grupo minoritario, pero significativo, de economistas para introducir un nuevo enfoque en el análisis de los procesos económicos que se conoce con el nombre de economía evolucionista. Los libros de Geoffrey Hodgson, uno de los pioneros de este movimiento, y de Ulrich Witt, citados en la selección bibliográfica que complementa este trabajo, son especialmente adecuados para un recorrido a fondo por los principales desarrollos teóricos de esta joven disciplina[1]. La nueva perspectiva cuestiona el concepto de equilibrio estático y el uso de la metodología individualista basada en ese *Homo economicus* racional que caracteriza el paradigma de la teoría neoclásica. Según Hodgson la economía y la biología tienen puntos de coincidencia en torno a tres grandes cuestiones: el papel de la historia en el devenir de los procesos evolutivos, la similitud metodológica en el análisis de los cambios evolutivos y, por último, la propia teoría de la evolución o conjunto de leyes y principios que explican la dinámica evolutiva. Por ello, no sería extraño que en un futuro próximo el papel que ejerció la física newtoniana como soporte de la teoría económica sea ocupado por la biología darwinista.

A continuación centraremos nuestro comentario en el análisis detallado de lo que probablemente ha resultado ser la interacción más fructífera entre el pensamiento económico y el darwinista en los últimos años: el estudio de la conducta cooperativa a partir de modelos extraídos de la teoría de juegos y la caracterización psicológica de ese *Homo economicus*, que ha acabado siendo bien distinta de la que predecían los modelos clásicos.

LA EVOLUCIÓN DE LA COOPERACIÓN: EL DILEMA DEL PRISIONERO

La teoría de juegos ha demostrado ser un marco teórico adecuado para la interpretación de la evolución del comportamiento, sobre todo el cooperativo, desde una perspectiva darwinista. Esta disciplina matemática fue desarrollada por John von Neumann en los años veinte del siglo pasado, pero no adquirió importancia hasta la publicación en 1944 de su libro *Teoría de los juegos y comportamiento económico*, del que fue coautor el economista austríaco Oskar Morgenstern. En su formulación inicial la teoría de juegos surge como un intento de definir cómo debemos comportarnos los seres humanos en los conflictos humanos interpersonales adoptando un punto de vista racional. Pero, ¿qué significa en realidad un comportamiento racional? Se trata de escoger, entre las distintas estrategias posibles, aquella cuyo resultado nos sea más favorable. Sin embargo, la idea de racionalidad ha ido cambiando con el desarrollo de la propia teoría de los juegos. Primero se pensó que el comportamiento racional consistía en actuar de acuerdo con el denominado principio minimax: es decir, actuar

de forma que uno obtenga el menor coste posible (o, en su caso, el mayor beneficio) asumiendo que el contrincante va a tratar de seguir la estrategia que más nos perjudica. Ahora bien, esta regla es en sí misma inestable, ya que si alguno de los jugadores utiliza la estrategia minimax, al contrincante le interesa cambiar de estrategia.

Esto llevó al también matemático John Nash, muy conocido popularmente desde que se le concediera el Premio Nobel y, todavía más, desde que Ron Howard rodase la película *Una mente maravillosa*, basada en una biografía de este científico, a proponer un nuevo criterio de racionalidad: el llamado equilibrio de Nash. Se define como una combinación de estrategias, una para cada jugador, tales que ningún jugador puede mejorar su beneficio si cambia unilateralmente de estrategia. Este concepto de equilibrio se considera hoy central en la resolución de problemas o juegos desde una perspectiva no-cooperativa. Veamos en concreto su aplicación al conocido juego del «dilema del prisionero», el cual permite una formulación sencilla y elegante de los problemas que plantea la cooperación.

La estructura del juego es conocida: la policía detiene a dos conspiradores; interrogados por separado, cada uno de ellos debe escoger entre permanecer callado (cooperar con el otro prisionero) o confesar inculpando al otro (no cooperar). Si ambos confiesan (no cooperan entre sí), ambos serán condenados a cinco años de cárcel, pero si ambos guardan silencio (cooperan), el máximo castigo que pueden imponerles es el de un año en prisión. El problema surge cuando uno confiesa y el otro no. En este caso, el primero quedará libre mientras que el segundo será condenado a diez años de cárcel. Parece claro que lo mejor que puede hacer cada reo, cuando no sabe qué va a hacer el otro, es confesar inculpando al otro, ya que si éste permanece callado él saldrá libre y, si su compañero también confiesa, serán cinco años de cárcel cada uno.

La solución racional –que ambos jugadores elijan no cooperar– es precisamente la solución que corresponde al equilibrio de Nash. La paradoja es que, aunque lo racional es confesar, lo que más les beneficiaría a ambos presos en conjunto sería permanecer callados. La solución cooperativa del dilema del prisionero –esto es, que ninguno confiese a la policía– se conoce en teoría de juegos con el nombre de óptimo de Pareto en honor al economista italiano decimonónico Vilfredo Pareto. Se dice que el resultado de un juego es un óptimo de Pareto cuando ningún jugador puede mejorar su resultado sin que empeore el de algún contrincante. La solución de Pareto es central en la rama de la teoría de juegos que estudia los juegos cooperativos y, aunque la solución no es racional en sentido estricto, podría justificarse su adopción si suponemos que los dos prisioneros pueden negociar y firmar un acuerdo ante una instancia superior que vele por su cumplimiento. De alguna forma, lo que se está suponiendo es que sería racional aceptar una autoridad superior que trabajara en pro del bien común. Rescataríamos así la idea hobbesiana del Estado como garante de la estabilidad social.

En los últimos años ha habido una tendencia a resolver el dilema de la cooperación desde una perspectiva no-cooperativa que asuma como punto de partida la imposibilidad de controlar el cumplimiento de un posible pacto previo entre los jugadores. A ello han contribuido enormemente los modelos que la teoría neodarwinista ha elaborado para explicar la evolución de la cooperación y del altruismo en la naturaleza. La aplicación de la teoría de juegos al estudio del comportamiento

animal se debe, en buena medida, al eminente evolucionista John Maynard Smith, recientemente fallecido, y constituye en la actualidad uno de los pilares teóricos de la sociobiología. Este investigador propuso la idea de que, en los conflictos animales, la selección natural debe conducir a que se imponga como solución ganadora una estrategia que sea evolutivamente estable (EEE). Se define como EEE a cualquier estrategia tal que si todos los individuos de una población la practican, ningún mutante con una estrategia diferente obtiene un beneficio mayor. Desde el punto de vista formal, la EEE equivale a un equilibrio de Nash (aunque no es cierto que todo equilibrio de Nash tenga que ser una EEE).

La existencia de comportamientos altruistas plantea un reto a la interpretación neodarwinista del comportamiento: ¿de qué forma la selección natural ha podido favorecer conductas que son perjudiciales, en términos de disminución de su eficacia biológica, para los individuos que las practican? Una primera alternativa, propuesta ya por Darwin, es la denominada selección entre grupos. Se trata, en cierta forma, de una solución cooperativa del tipo óptimo de Pareto. Su lógica es bien sencilla. Cuando se estudia un comportamiento no sólo deben examinarse sus consecuencias con respecto al individuo que lo lleva a cabo, sino también las correspondientes a los demás individuos de su entorno. Si un comportamiento beneficia a todos, la selección natural lo favorecerá, mientras que si es perjudicial para todos, desaparecerá. Ahora, si tiene un impacto negativo en el individuo pero positivo para el grupo, la respuesta dependerá de la relación entre costes y beneficios. Hasta los años sesenta fue habitual pensar que existían características de los seres vivos que habían surgido para favorecer no la supervivencia del individuo, sino, en la terminología de entonces, *el bien de la especie*. Puede decirse que hoy en día la mayor parte de los biólogos evolutivos dudan de que este proceso pueda ser eficaz y piensan que la selección natural actúa favoreciendo a unos individuos frente a otros y no a unos grupos frente a otros. Piénsese que, además, el mantenimiento de un comportamiento altruista mediante selección entre grupos es esencialmente inestable, ya que un grupo altruista siempre puede ser invadido por individuos egoístas que aparecieran por mutación o migración y que estarían favorecidos por la selección natural, puesto que recibirían beneficios sin coste. Para contrarrestar este efecto, sería necesario una tasa muy alta de extinción y formación de nuevos grupos, lo cual no parece ser una situación habitual en la mayoría de las especies.

En un artículo ya clásico aparecido en 1964, el entonces joven biólogo británico William Hamilton ofreció una explicación de los comportamientos altruistas alternativa a la selección de grupos que se conoce como selección de parientes (*kin selection*). Este autor señaló que si un gen determinase a un individuo a sacrificar su vida salvando las de varios parientes, el número de copias de ese gen en las generaciones siguientes podría aumentar más rápidamente que si el sacrificio no se hubiera realizado, ya que los parientes tienen una probabilidad más alta de ser portadores de los mismos genes que el resto de los individuos de la población y esa probabilidad aumenta a medida que el parentesco se estrecha. En definitiva, el comportamiento altruista supone un coste para el individuo que lo practica pero comporta un beneficio para los que interactúan con él y, si estos individuos son sus parientes, dicho beneficio revertirá indirectamente en el abnegado. Richard Dawkins ha popularizado este mecanismo, para el que ha utilizado la atinada expresión de gen egoísta, que concibe a las adaptaciones como beneficiosas no para el grupo, ni para el individuo, sino para el propio gen que las

condiciona.

Una tercera forma de que el altruismo cooperativo pueda evolucionar surge si hay reciprocidad de manera que las ventajas y las cargas del comportamiento altruista se equilibran a lo largo del tiempo entre las parejas de individuos que interactúan. Si los individuos se turnan como autores y receptores de las acciones altruistas, los beneficios del altruismo pueden a largo plazo compensar los costes. Esta teoría fue propuesta por primera vez por Robert Trivers, pero correspondió de nuevo al biólogo Hamilton –y al economista Robert Axelrod– la formulación matemática de la misma a partir de un modelo enmarcado en el juego del dilema del prisionero.

El factor que introdujeron Axelrod y Hamilton para resolver el dilema fue la posibilidad de reconocer a los contrincantes con los que ya se ha jugado y de recordar algunos resultados obtenidos en esos encuentros. Es interesante detallar el método que emplearon estos autores para encontrar la solución óptima. Invitaron a catorce científicos de disciplinas varias (psicología, economía, política, matemáticas y sociología) a escribir un programa de ordenador con la estrategia que cada uno de ellos consideraba mejor para ganar. Por ejemplo, un programa podría especificar cooperar en todos los movimientos, otro cooperar los movimientos pares y no en los impares, etc. Los programas compitieron por parejas y, en cada caso, se calculó la puntuación obtenida. Posteriormente se llevaron a cabo sucesivas rondas, de manera que cada programa estaba representado proporcionalmente a la puntuación obtenida en la ronda anterior, tratando de simular la acción de la selección natural hasta que sólo un programa sobrevivió. El programa ganador fue el denominado Tit For Tat (TFT) enviado por el profesor de psicología Anatol Rapoport, de la Universidad de Toronto.

Sorprendentemente, la estrategia vencedora era extremadamente simple: consistía en empezar el juego siempre cooperando y en los sucesivos movimientos reproducir la actitud adoptada por el contrincante en el movimiento previo. Las características de esta estrategia que la hacen enormemente estable y eficaz son tres. En primer lugar, se trata de una estrategia *amable*, que nunca es la primera en dejar de cooperar o, en otras palabras, que no abandona la colaboración sin ser incitada a ello. En segundo lugar, responde inmediatamente al cese de la cooperación. Por último, no es una estrategia *rencorosa*, sino que vuelve a cooperar tan pronto como el oponente está dispuesto a hacerlo. No es fácil encontrar una expresión en castellano que transmita la idea rectora de TFT. Una posibilidad es considerarla equivalente a la ley del talión *ojo por ojo y diente por diente*. Sin embargo, como ha señalado Luigi Luca CavalliSforza, TFT no tiene sólo un sentido negativo: hay que devolver mal por mal, pero también bien por bien. Sería más razonable utilizar la frase *pagar con la misma moneda*, más acorde con el origen de la sentencia inglesa que, de acuerdo con el diccionario Webster's, deriva de *plus tip for plus tap*, algo así como más propina por más cerveza (de barril).

En realidad, tanto TFT como la estrategia egoísta no cooperativa son ambas EEE en una situación como la que plantea el dilema del prisionero cuando se juega de forma repetida y, por tanto, una vez mayoritarias en una población no pueden ser desalojadas por otra estrategia. La diferencia entre ambas es que la eficacia biológica de los individuos es mayor en una población de individuos TFT que en una de egoístas. Evidentemente, esta forma de estudiar la cooperación es simplista, pero permite establecer claramente las condiciones que pueden facilitar o no su evolución: los

individuos deben tener oportunidad de interactuar a menudo, deben ser capaces de recordar los apoyos que han dado y recibido, y ofrecer asistencia sólo a aquellos que les han auxiliado. Esto último es necesario para que las interacciones a medio y largo plazo no ocurran aleatoriamente. De esta forma, los beneficios de la cooperación no se repartirán al azar, sino que se transmitirán mayoritariamente a los que muestran una disposición colaboradora. El libro de Brian Skyrms, que se menciona en la selección bibliográfica final, recoge con inteligencia y claridad la aproximación evolucionista no sólo al problema de la cooperación, sino también a otros relacionados con el compromiso, la agresividad, el sentido de la justicia o la comunicación, incluidos en lo que se ha denominado el contrato social. Los libros de Fernando Vega Redondo y Herbert Gintis, también recogidos en la bibliografía, contienen una información más técnica y completa sobre estos temas desde un punto de vista eminentemente teórico.

REDEFINIENDO AL *HOMO ECONOMICUS*

El éxito de la teoría de juegos en el análisis de la evolución del comportamiento supuso un acicate para que los economistas trataran de encontrar analogías biológicas aplicables a la evolución de los procesos económicos. De hecho, parece claro que existe una afinidad entre las preguntas y los métodos que utilizan los economistas y los evolucionistas. Los evolucionistas interpretan los comportamientos animales como resultado de la interacción entre individuos o genes *egoístas* que tratan de dejar el máximo de descendientes. Por su parte, los economistas tratan de interpretar los fenómenos económicos como el resultado de las interacciones entre individuos inteligentes e interesados en su propio beneficio. En la teoría económica, los actores son individuos, no clases sociales (obviamente, esta afirmación no sería compartida por economistas de orientación marxista), que tratan de maximizar de manera racional su beneficio. La similitud entre los análisis económicos y los evolutivos se manifiesta precisamente de forma más clara en la resolución de problemas relacionados con esa maximización, bien de los beneficios en un caso, bien de la eficacia biológica en otro, y con la definición del equilibrio alcanzado. Paul Krugman, un conocido economista, ha señalado con cierta maldad que existe otra analogía entre economía y evolución. Mientras que ambos campos se fundamentan en modelos matemáticos cuyo análisis requiere cierta destreza algebraica, sus más famosos representantes son personas que, aunque no entienden bien de lo que hablan, son magníficos escritores que suplen el rigor con la literatura (o que sustituyen la jerga matemática por la literaria). Se refiere obviamente a Stephen Jay Gould y John Kenneth Galbraith, respectivamente. El primero porque, como dice Maynard Smith: «Los biólogos evolutivos tienden a verlo como una persona cuyas ideas son tan confusas que apenas merece la pena tenerlas en cuenta»; y el segundo porque, en palabras del propio Krugman: «Es para los economistas más serios una especie de intelectual diletante al que le falta la paciencia para elaborar un pensamiento profundo»[2].

No obstante, hay también importantes diferencias entre ambos planteamientos. En primer lugar, en la teoría de juegos económica los costes-beneficios dependen de una función de utilidad determinada por valoraciones subjetivas, mientras que en los juegos evolutivos se trata de una utilidad objetiva que corresponde al intercambio en eficacia biológica que tiene lugar como consecuencia de la interacción. La segunda diferencia está en el proceso de decisión que, en el primer caso, la toman individuos que se supone que son racionales, mientras que en la segunda es el proceso de la selección

natural. Por último, mientras que la teoría de juegos evolutiva es fundamentalmente descriptiva y trata de interpretar el comportamiento social en el marco de la selección natural, la teoría de juegos económica tiene un aspecto normativo, ya que trata de proponer cómo debe comportarse el jugador racional.

Precisamente una de las contribuciones más sorprendentes que ha aportado la analogía evolucionista a los análisis económicos ha sido la constatación de que los seres humanos no nos comportamos tan racionalmente como predecía la teoría neoclásica, entendiendo por racional el comportamiento que pretende hacer máximo el beneficio esperado. La evidencia en este sentido proviene del trabajo de dos grupos de investigadores: los economistas experimentales, que estudian el comportamiento real de los individuos implicados en juegos sociales, y los psicólogos evolucionistas, que muestran cómo los seres humanos se guían por sentimientos de justicia, culpabilidad y altruismo al relacionarse con sus congéneres. Veamos, pues, cómo se comporta en realidad el *Homo economicus* cuando se le insta a participar en dos conocidos juegos que permiten actitudes cooperativas.

El dilema del bien común o la tragedia de los comunes : uno de los supuestos de partida presente en muchos modelos cooperativos es el de que las interacciones ocurren en parejas y, por lo tanto, el beneficio del altruista se dirige a otro individuo. Si este último no devuelve la cooperación, el altruista puede buscar otro compañero o cambiar su comportamiento hacia él. Sin embargo, si el beneficio va dirigido a un grupo, el problema de la reciprocidad se vuelve complicado. Si sólo parte del grupo devuelve la cooperación, el altruista se encuentra ante un verdadero dilema: si suspende la cooperación, no está siendo justo con aquellos miembros del grupo que cumplieron el trato, pero, si continúa colaborando, no está castigando a los individuos del grupo que no lo hicieron. Este problema constituye la esencia del juego denominado dilema del bien común o tragedia de los comunes.

Se trata de un dilema social propuesto por Garrett Hardin que ha sido muy estudiado por sociólogos, politólogos, economistas y evolucionistas, y ha sido objeto de numerosas evaluaciones empíricas. En su forma más sencilla, podría plantearse como el siguiente juego: a cada uno de, por ejemplo, cuatro estudiantes, se les proporciona una cantidad, digamos cinco euros. Se les indica que pueden invertir parte de esta cantidad en un proyecto del grupo introduciendo una cantidad (entre cero y cinco euros) en un sobre. El experimentador recoge los sobres, suma la cantidad invertida por los estudiantes, la duplica y la reparte de nuevo entre ellos.

La predicción de la teoría económica clásica es que ninguno de los estudiantes contribuirá al proyecto común, ya que cada uno de ellos razonará del siguiente modo: si yo soy el único que contribuyo con, por ejemplo, dos euros, esta cantidad, una vez duplicada y repartida entre los cuatro, resultará en un beneficio de un euro. En ausencia de información sobre lo que los otros estudiantes harán, lo racional es ponerse en el peor escenario, esto es, que no aportarán nada, en cuyo caso lo mejor es que uno tampoco lo haga. El dilema surge porque el máximo beneficio lo obtendrían todos si invirtieran los cinco euros, ya que conseguirían otros cinco. Los intereses individuales y los del grupo son contradictorios.

Sin embargo, los resultados empíricos indican que los individuos cooperan más de lo

que predice la teoría clásica. Un aspecto interesante es que, si se modifica el juego de forma que se permita el castigo a los no cooperadores, el nivel de cooperación (la aportación de cada uno al bien común) aumenta sustancialmente. En un experimento muy conocido de Ernst Fehr y Simon Gächter, cada uno de los miembros de un grupo de cuatro estudiantes recibe veinte euros, parte de los cuales pueden invertirlos en un proyecto común, quedándose con el resto. Por cada euro invertido en la comunidad cada miembro del grupo recibe 0,40 euros con independencia de cuál haya sido su aportación concreta. Según la hipótesis clásica, los individuos persiguen únicamente su propio interés y, por ello, deberían quedarse con los veinte euros sin aportar nada al proyecto en común, a pesar de que si todos ellos aportaran sus veinte euros obtendrían $0,40 \times 80 = 32$ euros cada uno. La hipótesis no se cumplió, puesto que un 75% de los individuos contribuyó con cinco o más euros. Pero lo más interesante es lo que ocurre cuando el juego se modifica de forma que pueda castigarse a los no cooperadores. Para introducir esta modificación se informa a los jugadores sobre la aportación que ha hecho cada uno y se les permite castigar a los otros miembros del grupo asignándoles entre cero y diez puntos. Cada punto cuesta un euro al castigador e implica una pérdida de tres euros para el castigado. Aunque de nuevo la predicción es que no deberían imponerse castigos, ya que no rinde ningún beneficio, casi un 85% de los jugadores castigó en alguna ocasión. La mayor parte de las veces el castigo era impuesto a los que contribuían menos que la media por parte de los que contribuían por encima de la misma. Además, el nivel de cooperación aumentó considerablemente: un 75% de los individuos contribuyó con quince o más euros.

Parece razonable que los individuos traten de evitar el castigo y, por lo tanto, tiendan a comportarse de acuerdo con lo que consideran la norma del grupo. Lo que no parece tan claro es la razón de castigar a los que no colaboran (*free riders*) cuando esta acción es costosa. Los investigadores creen que el mecanismo psicológico responsable de estos comportamientos se basa en las emociones negativas que despiertan en los seres humanos los individuos insolidarios que se aprovechan del esfuerzo ajeno, emociones que nos impulsan a imponer un castigo aunque nos resulte costoso.

El juego del ultimátum (lo tomaso lo dejas) consiste en lo siguiente: hay un bien a repartir (por ejemplo, una cantidad sustancial de dinero). El jugador 1 debe hacer una oferta de reparto, mientras que el jugador 2 puede aceptarla o rechazarla. Si la acepta, recibe lo que se le ofreció, mientras que el otro obtiene la diferencia; si no la acepta, ambos jugadores se quedan sin nada. Según la predicción del modelo clásico, en el que los individuos actúan únicamente en función de su propio interés, el jugador 1 considerará que el jugador 2 debe aceptar cualquier propuesta, puesto que para dicho jugador representa siempre un beneficio neto y, por lo tanto, propondrá la cantidad más pequeña posible.

Se ha demostrado repetidamente en experimentos realizados con estudiantes universitarios que estas predicciones no se cumplen. En general, la oferta que propone el jugador 1 suele ser próxima a un 44% del recurso, mientras que el jugador 2 rechaza las ofertas inferiores a un 20%, con probabilidad 0,50. Incluso en una variante del juego en que el individuo que recibe la oferta está obligado a aceptarla, conocida como el juego del dictador, las ofertas que se efectúan, aunque más bajas que en el juego normal del ultimátum, en lugar de ser nulas, que sería el resultado esperado, son casi siempre mayores que cero.

Este tipo de experimentos se han criticado en el sentido de que pudieran ser reflejo de la cultura occidental y no un comportamiento universal. Sin embargo, un estudio coordinado por Joseph Henrich en el que dicho experimento se efectuó en quince sociedades con economías de *pequeña escala* (hablar de sociedades *primitivas* no es políticamente correcto) indicaba un incumplimiento sistemático de las predicciones racionales, aunque exhibía una considerable variabilidad entre culturas. Mientras en los Machiguenga de Perú la oferta de reparto fue en promedio del 26%, en los Orma de Kenia fue del 58%. La explicación de estas diferencias parece radicar en el modo de vida y la estructura social de estas sociedades. Así, las descripciones etnográficas de los Machiguenga reflejan poca cooperación, intercambio o compartimiento de alimentos que no sea estrictamente dentro de la unidad familiar, mientras que los Orma parecen estar siempre dispuestos a contribuir a actividades comunitarias como escuelas o caminos. El estudio completo y una exhaustiva interpretación de los mismos se recogen en el libro de Henrich y colaboradores mencionado en la ficha bibliográfica.

Los resultados obtenidos en ambos juegos parecen concluyentes: el *Homo economicus* neoclásico tiene que dejar paso a un ser humano racional, pero impregnado de emociones y sentimientos morales que condicionan de forma nítida sus decisiones y le hacen ser extraordinariamente cooperativo. Tendemos a incrementar nuestro beneficio, pero menos de lo previsto. En cierto modo, la pelota está de nuevo en el terreno evolucionista: la tarea pendiente consiste en investigar qué significado evolutivo poseen esos comportamientos y los sentimientos y rasgos psicológicos que los hacen posibles.

[1] Véase también el número de la revista *Anthropos* dedicado a «Economía y evolución»: *Anthropos*, núm. 182 (1999).

[2] Maldades aparte, es cierto que Stephen Jay Gould es el evolucionista más leído y citado por los economistas evolutivos. Una visión más positiva de este autor puede encontrarse en las reseñas de algunos de sus libros escritas por Carlos López-Fanjul (*Revista de libros*, números 17, 34 y 86) y de Steven Rose (*Revista de libros*, núm. 86).