



# La preferencia revelada frente al enfoque utilitarista: discusión sobre los fundamentos de la teoría del consumidor

## Revealed preference versus the utilitarian approach: discussion on the foundations of consumer theory

**Carlos Villacís** (cvillaci@ucm.es) Facultad de Filosofía, Universidad Complutense de Madrid (Madrid, España) <https://orcid.org/0000-0003-2339-4463>

### Abstract

This article makes a critical comparison of two approaches to the foundations of the microeconomic theory of consumer behavior, namely: the utilitarian approach and that of the revealed preference theory. An analysis of revealed preference is carried out in the context of introducing restrictions to the choice function in decision theory. Its advantages and limitations are evaluated in contrast to the utility problem proper to the subjectivist point of view of utilitarian theory. And the explanatory capacity of the different models of rational agency is confronted on the bases of the demand curves. Furthermore, a formal critique of the revealed preference method is developed in strategic decision-making environments.

**Key words:** revealed preference, utility, choice, consumer theory, prisoner's dilemma.

### Resumen

En el presente artículo se realiza una comparación crítica de dos enfoques sobre los fundamentos de la teoría microeconómica del comportamiento del consumidor, a saber: el enfoque utilitarista y el de la preferencia revelada. Se lleva a cabo un análisis de la preferencia revelada en el contexto de la introducción de restricciones a la función de elección en la teoría de la decisión. Se evalúan sus ventajas y limitaciones en contraste con el problema de la utilidad propio del punto de vista subjetivista de la teoría utilitarista. Y se confronta la capacidad explicativa de los distintos modelos de agencia racional sobre las bases de las curvas de demanda. Asimismo, se desarrolla una crítica formal al método de la preferencia revelada en entornos estratégicos de decisión.

**Palabras clave:** preferencia revelada, utilidad, elección, teoría del consumidor, dilema del prisionero.



## 1. Introducción

El alcance de la investigación expuesta en el presente artículo abarca el campo de la preferencia revelada como enfoque del programa de investigación de Paul Samuelson en la teoría del comportamiento del consumidor. La naturaleza del problema investigado radica en analizar las restricciones al modelo de agencia racional que la preferencia revelada impone en su estrategia de excluir la noción de utilidad. El método consiste en definir lógico-formalmente la noción de preferencia revelada en el marco de la teoría de la decisión con el enfoque de las funciones de utilidad. El objetivo es evaluar las ventajas y limitaciones de este modelo de agencia racional frente al enfoque utilitarista.

Aunque se mencionan, no se trata de un estudio exhaustivo de las distintas formulaciones y diferentes contribuciones allende las de Samuelson. Antes bien, el análisis se centra en el axioma débil de preferencia revelada como condición de racionalidad de la función de elección en la teoría de la decisión. Por tanto, la investigación se circunscribe a un marco teórico muy preciso que podemos sintetizar del siguiente modo: el concepto de preferencia revelada en relación con el programa de investigación de Samuelson en la teoría del consumidor, con interés en el modelo de agencia racional para la función de elección a partir del condicionamiento del axioma débil de preferencia revelada.

Por consiguiente, la pregunta básica que orienta la investigación se podría formular del siguiente modo: ¿cuál es el rendimiento explicativo del modelo de elección racional condicionado por el axioma débil de preferencia revelada en comparación con las estrategias utilitaristas? El desarrollo se ha estructurado en seis partes. En primer lugar, se ubica conceptualmente la noción de preferencia revelada en el marco de la investigación de Samuelson en la teoría del consumidor. En segundo lugar, se introducen los elementos de la lógica de las preferencias y de la teoría de la decisión. En tercer lugar, se analiza formalmente la preferencia revelada mostrando su innovación teórica para la contrastación empírica. En cuarto lugar, se evalúa su capacidad para explicar los fundamentos racionales de la microeconomía frente al enfoque utilitarista. En quinto lugar, se presenta una relación de críticas y, en sexto lugar, se desarrolla un argumento formal contra la posibilidad del método de la preferencia revelada en dilemas de entornos estratégicos.

## 2. PR como enfoque para la teoría del consumidor

La preferencia revelada se ha comprendido de diversas maneras, de las cuales destacamos las tres siguientes siguiendo a Hands ([Paul Samuelson and revealed preference theory](#)): como un marco conceptual alternativo a la teoría de la utilidad ordinal para el análisis de la teoría de la demanda en la elección del consumidor, como un complemento lógicamente equivalente a la teoría estándar y como una teoría de la elección exhaustiva (a esta última Hands la califica “contemporary revealed preference theory”).

Para acotar el alcance de la investigación sobre la preferencia revelada (en adelante PR), nos centramos en lo que Stanley Wong denomina “el programa de investigación de Paul Samuelson en el enfoque de la preferencia revelada de la teoría del comportamiento del consumidor” (Wong 2006:1). Por mor de la brevedad utilizaremos las siglas SRP (*Samuelson’s research programme*) para referirnos a este enfoque.



Podemos definir el SRP en tres puntos elementales. Tales puntos los presentamos en una relación entre las motivaciones teóricas que Samuelson pretendía con la PR y las publicaciones en las que las desarrollaba, tal como Wong expone. Distinguimos, pues, tres objetivos teóricos sucesivos:

- (i) Excluir cualquier concepto que no pueda corresponderse con datos empíricos, en particular el concepto de utilidad clásicamente definido en términos subjetivistas, mentalistas e internos (por lo tanto, no observables ni comparables). Lo encontramos en Samuelson ([A note on the pure theory of consumer's behaviour](#)).
- (ii) Construir un método mediante el cual poder operar con el concepto de preferencia. Lo encontramos en Samuelson ([Consumption theory in terms of revealed preference](#)).
- (iii) Ofrecer de un modo completo las bases empíricas para sostener el análisis de la utilidad ordinal. Lo encontramos en Samuelson ([The problem of integrability in utility theory](#)).

El O.i. (objetivo i) se desarrolla en una nueva teoría del comportamiento del consumidor definida en términos observacionales. Dicha teoría se fundamenta mediante la suposición de un principio de consistencia de la conducta que permita “dropping off the last vestiges of utility analysis” (Samuelson 1938:62). El O.ii. emplea la nueva teoría (PR) como fundamento que permita producir mapas de indiferencia de individuos a partir de los datos empíricos que el comportamiento de tales individuos manifiesta en el mercado. Un mapa de indiferencia es la representación de varias curvas de indiferencia de un individuo. El O.iii. se consigue mediante la demostración de que la teoría de la utilidad ordinal y la de la PR son lógicamente equivalentes. Fue Houthakker ([Revealed preference and the utility function](#)) quien consiguió demostrarlo.

Sobre este programa podremos evaluar los éxitos y las ventajas de la PR, así como sus críticas y limitaciones. Además, para comprender sus motivaciones y su relevancia teórica es preciso encuadrarla en su marco teórico, a saber, el de la teoría de la elección racional y su vinculación con otras áreas como la microeconomía o la teoría de la elección social (entre otras). No obstante, cabe señalar que la PR ha suscitado multitud de cuestiones colindantes y formulaciones que exceden el marco teórico especificado. En este contexto, precisamos el orden conceptual en el que se desarrolla la presente investigación. Debemos aclarar que en este marco teórico la PR no es *stricto sensu* una teoría. A pesar de las denominaciones habituales en la literatura relacionada, la PR no satisface las condiciones modelo-teóricas de una teoría formal ni las metacientíficas de una teoría positiva. Se trata, antes bien, de un método para dotar de base observacional a la teoría del comportamiento del consumidor. Para presentarlo hacemos uso del sistema axiomático de modo que el método de la PR se introduce como restricciones formales en la función de elección. La teoría es la del comportamiento del consumidor, la cual forma parte de la disciplina microeconómica. Así pues, debemos diferenciar tres niveles conceptuales: el SPR, la teoría del consumidor y los fundamentos de la teoría en términos de agencia racional, los cuales pueden corresponder al enfoque utilitarista o al de la PR.

Al evaluar críticamente el efecto del método de la PR en la teoría antedicha, se incide en los principios que estructuran el razonamiento económico. Por consiguiente, nos encontramos en el plano cognoscitivo para analizar los principios metodológicos que rigen el modelo de agencia racional en el cual se fundamenta la teoría del consumidor. Esto no implica que las cogniciones se constituyan como la base de las preferencias, dado que se trata de principios prácticos, sino que se realiza un análisis filosófico de la teoría económica misma, esto es, en el plano epistemológico. Hay



que distinguir entre el método de la PR y la noción de preferencia. La primera la expresamos en axiomas que condicionan la función de elección y respecto a la segunda realizamos un tratamiento lógico representándola como una relación en términos conjuntistas. Con el método de la PR, en contraste con el enfoque utilitarista, la noción de utilidad subjetiva se convierte en irrelevante para operar con las preferencias.

### 3. Lógica de las preferencias y fundamentos de la teoría de la decisión

Para construir un modelo de agencia racional operativo en la teoría del comportamiento del consumidor necesitamos representar, a través de la modelización lógica, las nociones elementales relacionadas con las preferencias.

#### 3.1. Tipos de relaciones, preferencia estricta e indiferencia

Empezamos por especificar una relación binaria  $R$  en un dominio cualquiera de alternativas  $A$ .  $R$  es un subconjunto del conjunto de todos los pares ordenados  $A \times A$ , es decir, está incluido en el producto cartesiano del conjunto de alternativas consigo mismo. Una relación binaria puede cumplir determinadas características. Destacamos las siguientes:

Reflexividad:  $\forall x \in A: xRx$

Transitividad:  $\forall xyz \in A: (xRy \wedge yRz) \supset xRz$

Completitud:  $\forall xy \in A: (x \neq y) \supset (xRy \vee yRx)$

Simetría:  $\forall xy \in A: xRy \supset yRx$

Asimetría:  $\forall xy \in A: xRy \supset \neg (yRx)$

Antisimetría:  $\forall xy \in A: (xRy \wedge yRx) \supset x = y$

Según las propiedades que la relación cumpla, podemos definir el tipo de relación que es. Existen tipos de relaciones estandarizados:

Un *orden* es reflexivo, transitivo y completo.

Un *cuasi-orden* es reflexivo y transitivo.

Un *orden parcial* es reflexivo, transitivo y antisimétrico.

Una *cadena* es reflexiva, transitiva, completa y antisimétrica.

Un *orden parcial estricto* es transitivo y asimétrico.

Un *orden fuerte* es transitivo, asimétrico y completo.

Con estos elementos podemos definir las dos nociones básicas:

Def.1. La *preferencia estricta*  $P := xPy \leftrightarrow (xRy \wedge \neg (yRx))$

Def.2. La *indiferencia*  $I := xIy \leftrightarrow (xRy \wedge yRx)$

$R$  se podría comprender como una *preferencia débil*:  $xRy \leftrightarrow (xPy \vee xIy)$ . Es frecuente emplear asimismo los símbolos  $\succ$ ,  $\sim$ ,  $\succsim$  en lugar de  $P$ ,  $I$ ,  $R$ , respectivamente.



### 3.2. Conjunto maximal y conjunto de elección

En este punto necesitamos representar de forma unitaria la multiplicidad (puede ser un conjunto unitario o vacío) de aquellos elementos que no son preferidos sobre otros elementos de un conjunto de alternativas. Para esto es preciso definir la noción de elemento maximal.

Def.3. Respecto a la relación binaria  $R$ ,  $x \in A$  es un *elemento maximal* de  $A$  sii  $\neg \exists y (y \in A \wedge yPx)$ .

Dada la definición 3, llamamos *conjunto maximal de  $A$*  a  $M(A, R)$ , el cual no es más que el conjunto de los elementos maximales de  $A$ . Luego, necesitamos representar de forma unitaria la multiplicidad (también puede ser un conjunto unitario o vacío) de aquellos elementos que son al menos tan preferidos como cualquier otro elemento de un conjunto de alternativas. Para esto es preciso definir la noción de mejor-elemento.

Def.4. Respecto a la relación binaria  $R$ ,  $x \in A$  es un *mejor-elemento* de  $A$  sii  $\forall y (y \in A \supset xRy)$

Dada la definición 4, llamamos *conjunto de elección de  $A$*  a  $C(A, R)$ , el cual no es más que el conjunto de los mejores-elementos de  $A$ . Es importante hacer notar que  $C(A, R) \subseteq M(A, R)$ , es decir, que todo mejor-elemento es un elemento maximal, pero puede ser que un elemento maximal no sea mejor-elemento.

### 3.3. Teoría de la decisión: función de elección y propiedades de racionalidad

Steele y Stefánsson ([Decision theory](#)) explican que el objeto de la teoría de la decisión racional reside en las formas del razonamiento que orientan las elecciones de un agente. Dicho razonamiento se define formalmente con base en la lógica de las preferencias. En términos conjuntistas, es común representar la noción de elección (choice) como una función  $f$  tal que, para cualquier subconjunto de un conjunto de alternativas,  $f$  asigna aquellos elementos del subconjunto que un agente ha decidido no descartar en su elección. La intención es llegar a los *elementos elegidos* del conjunto de alternativas a partir de una función.

Def.5.  $f$  es una *función de elección* de  $A$  sii  $f$  es una función tal que para todo subconjunto  $B \subseteq A$ ,  $f(B) \subseteq B$  y si  $B \neq \emptyset$  entonces  $f(B) \neq \emptyset$

La justificación de este concepto desde los elementos de la lógica de las preferencias expuestos es el siguiente: una función de elección de  $A$  es una función tal que  $C(B, R)$  no es el conjunto vacío para cada subconjunto  $B$  no vacío de  $A$ . Esto significa que cada subconjunto no vacío de  $A$  tiene un mejor-elemento. Se están destacando, por tanto, los mejores-elementos de todos los subconjuntos de un conjunto de alternativas dado. De este modo se justifica la noción de *elementos elegidos* de  $A$ , dado que los elementos no descartados son aquellos *al menos tan preferidos como cualquier otro elemento*. Esta noción básica de elección remite, por tanto, a la relación de preferencia débil puesto que con ella se define el concepto de *mejor-elemento* y, consiguientemente, el de *conjunto de elección*. La idea de ser “al menos tan preferido como” alude a un orden reflexivo, asimétrico y transitivo. Por este motivo se define la función de elección en términos de  $R$  (y no de  $P$ ).



La existencia de funciones de elección es un elemento básico para la elección racional estructurada en la teoría de la decisión. No obstante, también se requieren diversas *propiedades de racionalidad* que condicionan la función de elección. Presentamos las tres más destacadas:

Propiedad  $\alpha$ :  $(x \in f(A) \wedge x \in B \subseteq A) \supset x \in f(B)$

Propiedad  $\beta$ :  $(xy \in f(B) \wedge B \subseteq A) \supset (x \in f(A) \leftrightarrow y \in f(A))$

Propiedad  $\gamma$ :  $(x \in \bigcap_i f(A_i) \text{ para todo } A_i \text{ en una clase}) \supset x \in f(\bigcup_i A_i)$

La propiedad  $\alpha$  (también llamada “condición de Chernoff” o “consistencia básica de contracción”) establece que si un elemento es elegido en un conjunto y ese elemento pertenece a un subconjunto del conjunto del que fue elegido, entonces ese elemento será el elegido del subconjunto al que pertenece. Esto quiere decir que la elección no se ve modificada si se eliminan algunos de los elementos que no fueron elegidos. La propiedad  $\beta$  establece que, si dos elementos son elegidos en un subconjunto, entonces ambos también deben ser elegidos en el conjunto en el que está incluido el subconjunto al que pertenecen. La propiedad  $\gamma$  (también llamada “consistencia básica de expansión”) establece que, si un elemento es elegido en todos los conjuntos de una clase, entonces debe ser elegido en el conjunto que resulta de la unión de los conjuntos de tal clase.

Sen ([Collective choice and social welfare](#)) demuestra varias relaciones entre las preferencias y las propiedades de racionalidad de la elección. Entre ellas destacamos las siguientes: una función de elección construida a partir de una relación binaria  $R$  cumple la propiedad  $\alpha$ , pero no necesariamente la propiedad  $\beta$ ; y una función de elección construida a partir de una relación binaria  $R$  cumple la propiedad  $\beta$  sii  $R$  es un orden.

#### 4. Relación entre las preferencias y la elección en la PR

Expuestas (de modo sintético) las funciones de elección y caracterizadas con sus propiedades más destacadas, podemos comprender la operatividad del concepto de PR en su intento de ofrecer bases observacionales a la teoría del consumidor. Dicha operatividad del concepto se vincula con tal objetivo en la particular relación que la PR ofrece entre las preferencias y la elección del consumidor. Los axiomas de la preferencia revelada permiten una relación entre preferencias y elección específica con base observacional.

##### 4.1. Los axiomas de PR

Para poner de relieve este aspecto continuamos con las propiedades de racionalidad de la función de elección en el área de la decisión racional. A las tres anteriores añadimos las dos siguientes: el *axioma débil de preferencia revelada* (en adelante WARP por sus siglas en inglés) y el *axioma fuerte de preferencia revelada* (en adelante SARP por sus siglas en inglés). Es posible expresar ambos axiomas en términos de la función de elección. A continuación, nos centramos en el WARP:

WARP:  $(xy \in A \wedge x \in f(A)) \supset \forall B((xy \in B \wedge y \in f(B)) \supset x \in f(B))$

Lo que el WARP pone como condición a la función de elección es que, si en un conjunto de alternativas un elemento  $x$  es elegido cuando otro elemento  $y$  está disponible, entonces no puede haber otro conjunto de alternativas distinto en el que  $y$  sea elegido y  $x$  no lo sea. Tal como observa



Sen ([Choice functions and revealed preference](#)), el axioma débil de la PR es equivalente a la introducción de las propiedades  $\alpha$  y  $\beta$ . Por otro lado, Sen señala que a partir del WARP podemos definir la relación  $R_w$  de preferencia revelada débil:

Def. 6.  $R_w := xR_wy \leftrightarrow \exists A (x \in f(A) \wedge y \in A)$

Mantenemos el subíndice  $w$  para hacer referencia al WARP. Con esta definición podemos, a su vez, definir la noción de la *binariedad* de la elección:

Def.7. Para cada  $A \neq \emptyset$ ,  $f(A) = \{x \mid x \in A \wedge \forall y \in A: xR_wy\}$

La binariedad de la función de elección es importante puesto que  $R_w$ , en tanto que base de la elección, puede volver a producir la misma función de elección de la cual ha sido producida. Una función de elección es binaria si se cumplen las propiedades  $\alpha$  y  $\gamma$ .

#### 4.2. La medida de las preferencias: la utilidad ordinal

Un tema central de la teoría de la decisión es la representación numérica de los órdenes de preferencias. Tradicionalmente esta cuestión se ha comprendido como una medición de la utilidad. Para representar la mensurabilidad se construyen *funciones de utilidad*. Estas funciones pueden ser de dos tipos según la exactitud de la información que recojan: *funciones de utilidad ordinales* y *funciones de utilidad cardinales*.

Se trata de una aplicación de determinada cantidad de bienes a determinada cantidad de utilidad que le reportan al consumidor que elige. Las funciones ordinales no expresan cuantitativamente las diferencias o proporciones, solo establecen si una alternativa es preferida a otra. Son las escalas informativamente más débiles. No obstante, nos centramos en las ordinales dado que son las pertinentes para la PR. Dada una función de utilidad, la relación de preferencia representa una *maximización de la utilidad*. La relación de preferencia debe ser transitiva y completa para poder ser representada por una función de utilidad ordinal.

Def.8. Una función de utilidad ordinal  $u :=$  para cualquier  $xy \in \text{dom}(u)$ :  $u(x) \geq u(y) \leftrightarrow xRy$

La medición de la utilidad ha suscitado importantes problemas en la justificación epistémica de la teoría que se basan en ella, dado que es difícil (o imposible) acceder y comparar (incluso establecer si existe) algo cuyo estatuto ontológico es de carácter mentalista y, por tanto, interno y no comparable por definición. Podemos referirnos a esta dificultad como el *problema de la utilidad*. El WARP apareció (no con esta denominación y no respecto a las funciones de elección) en Samuelson ([A note on the pure theory of consumer's behaviour](#)) como una condición de consistencia en el contexto de las funciones de demanda del consumidor para eliminar el concepto de utilidad. Sería una reformulación de la utilidad ordinal, lo cual se corresponde con el O.i (supra, § 2). Podemos exponerlo en dos pasos siguiendo el esquema de Hédoin ([Sen's criticism of revealed preference theory and its 'neo-samuelsonian critique'](#)):





(1) Dadas funciones de demanda  $d_i$  continuas para  $n$  bienes, las funciones indican la cantidad  $x_i$  de los  $i$  bienes consumidos según los precios  $p_1, \dots, p_n$  y  $M$ , siendo  $M$  el “ingreso monetario disponible”. Lo representamos del siguiente modo:  $x_i = d_i(p_1, \dots, p_n, M)$ ,  $i = 1, \dots, n$ .

(2) El WARP condiciona de esta manera: si el consumidor elige el conjunto de bienes  $x^0$  a un precio  $p^0$  pudiendo elegir el conjunto de bienes consumibles  $x^1$ , es decir si  $x^0$  se ha revelado preferido a  $x^1$ , entonces  $x^1$  no puede ser elegido al precio  $p^1$ ; y si  $x^1$  es elegido al precio  $p^1$  solo puede ser debido a que  $x^0$  no era asequible al precio  $p^1$ . Asumiendo en el antecedente que  $x^1$  es asequible, esto es, que se considera a un precio  $p^0$ :  $\sum_{i=1}^n p_i^0 x_i^0 \leq \sum_{i=1}^n p_i^0 x_i^1 \Rightarrow \sum_{i=1}^n p_i^1 x_i^0 > \sum_{i=1}^n p_i^1 x_i^1$

Lo que Samuelson consigue es que las funciones de demanda condicionadas por el WARP satisfagan “dos de las tres restricciones estándar de la teoría de la utilidad ordinal: efectos de sustitución negativos y una matriz de sustitución Slutsky semidefinida negativa” (Hands 2014:89). Como muestran Mas-Colell, Dwhinston y Green ([Microeconomic theory](#)) se pueden comprobar los efectos que tiene el WARP (en tanto que condición de consistencia en la *teoría de la decisión* justificada en la elección) en la demanda del consumidor (*microeconomía*). Lo cual pone de relieve el efecto sobre los fundamentos epistemológicos de la preferencia revelada en la teoría del consumidor.

#### 4.3. La posibilidad de inferir la preferencia a partir de la elección

El modo estándar en virtud del cual se relacionan la elección y la preferencia pasa por construir una función de elección a partir de una relación de preferencia. Existen diversos métodos para conectarlos, de los cuales enunciamos dos: la *conexión de la mejor elección* y la *conexión de la no-dominancia*. El primer método establece que  $f(A) := \{x \in A \mid \forall y \in A: xRy\}$ . El segundo método estipula que  $f(A) := \{x \in A \mid \forall y \in A: \neg(yPx)\}$ . Lo que interesa destacar en este punto es la direccionalidad: a partir de las preferencias se construyen elecciones.

Ahora bien, el método de la PR permite definir la relación de preferencia en términos de la función de elección, esto es, se invierte la direccionalidad de la explicación dado que se toma como dato de partida la conducta observable de la elección. Por consiguiente, según las propiedades que cumplan las elecciones se podrán producir las relaciones de preferencia. También existen diversos métodos de llevar a cabo esta operación. A continuación, presentamos tres siguiendo a Hansson y Grüne-Yanoff ([Preferences](#)):

$$\begin{aligned} xRy & \text{ sii } \exists A \text{ tal que } y \in A \wedge x \in f(A) \\ xRy & \text{ sii } x \in f(\{x, y\}) \\ xPy & \text{ sii } \exists A \text{ tal que } x \in f(A) \wedge y \in (A - f(A)) \end{aligned}$$

En los dos primeros métodos podemos obtener P e I aplicando las definiciones 1 y 2, respectivamente. En el tercero podemos obtener I aplicando la definición 2 y R:  $xRy$  sii  $\neg xPy$ . Lo que nos interesa destacar estriba en que si aplicamos el primer método a la restricción del WARP sobre las funciones de demanda (supra, § 4.2.), obtenemos que:  $x^0Rx^1$  sii  $x^0 \times p^0 \geq x^1 \times p^0$ .

El interés teórico radica en que la inversión de la direccionalidad de la carga explicativa permite evitar el *problema de la utilidad*, puesto que de lo que se parte es de una conducta observable (las





elecciones del agente) y no de las supuestas utilidades que orientan las preferencias. En principio, este mecanismo permitiría comprobar la validez empírica de las preferencias.

## 5. PR y modelo de racionalidad económica

La construcción de relaciones de preferencia a partir de elecciones es útil en economía, ya que los datos empíricos de que disponemos son precios y elecciones. Así pues, cabe destacar que los axiomas de la PR tienen un mayor impacto y presencia en los modelos de racionalidad específicamente económica, aun siendo cierto que pueden estar presentes en otras áreas como la Teoría de Juegos, la Teoría de la Elección Social o la Teoría de la Elección Pública. En este punto nos centramos en la vertiente económica con el fin de resaltar los aspectos de interés para el modelo de agencia racional que se necesita.

### 5.1. Supuestos fundamentales de la teoría del comportamiento del consumidor

Las *curvas de indiferencia* nos permiten modelizar en un espacio de combinaciones de bienes aquellos puntos que satisfacen de la misma manera las preferencias de un consumidor. Es decir, que la relación entre los puntos de la curva es de indiferencia, por lo que se pueden trazar infinitas curvas sin que se corten entre ellas. La *función de utilidad* representa la medida de la satisfacción aplicando el orden a las cantidades de los distintos bienes.

Con ciertos supuestos de racionalidad, lo anterior permite obtener una representación de la relación entre la máxima cantidad y el precio de un bien en particular que el consumidor llegaría a comprar. De este modo, llegamos en la teoría del consumidor a la *curva de demanda* de un agente. Esto aporta uno de los elementos fundamentales para la noción de equilibrio, la cual explica la determinación de los precios de una economía (si la teoría se generaliza) en relación con el funcionamiento de la producción y el consumo en condiciones de intercambio.

La indiferencia entre cantidades de diferentes bienes posibilita realizar una predicción sobre la cantidad de cada bien que se adquirirá. La predicción sobre cantidades de mercancías demandadas es central para la microeconomía. Para que puedan cumplirse estos conceptos microeconómicos, el agente racional debe ordenar las alternativas de acuerdo con ciertas propiedades de la relación de preferencia  $R$ . Presentamos los *supuestos estándar de la preferencia* que se exigen racionalmente para la existencia de los conceptos fundamentales de la teoría del consumidor:

- La *completitud* es necesaria para que las cestas de bienes sean comparables entre sí.
- La *transitividad* es necesaria para que no se crucen las curvas de indiferencia y no se incurra en situaciones de pérdida o que no maximicen.
- Se exige una *elección racional* entendida como que, si hay un elemento elegido en un conjunto de alternativas, entonces ese elemento es un *mejor-elemento* (definición 4) de ese conjunto. Dicho de otro modo, un *mejor-elemento* pertenece al *conjunto de elección*. De manera más sencilla: “se elige lo que se prefiere”. Este supuesto indica que el consumidor siempre elegirá la alternativa que maximiza su utilidad, a saber, aquella con la que el valor de la función  $u$  es máximo.
- La *no-saturación* implica que un agente racional prefiere tener más de un bien a tener menos.



-Es necesario suponer que la relación de preferencia debe ser *continua*, tal como establecen Mas-Colell et al. ([Microeconomic theory](#)). Se trata de una *conditio sine qua non* para que una relación de preferencia sea susceptible de ser representada por una función de utilidad.

-Las curvas de indiferencia deben ser *convexas*, lo cual expresa que cuando un bien es más escaso se valora más. Las curvas no llegarán a tocar los ejes.

-Las curvas de indiferencia deben ser *decrecientes* debido a la *relación marginal de sustitución*, la cual marca la pendiente indicando “el número de unidades de *y* al que el sujeto está dispuesto a renunciar para obtener una unidad adicional de *x*, manteniéndose en la misma curva de indiferencia. Es decir, es el precio de *y* en función de *x*” (Torres 2002:33).

## 5.2. Racionalidad y consistencia

Para las teorías económicas la *racionalidad* de la conducta puede definirse, tal como explica Sen ([Sobre ética y economía](#)), a partir de la *consistencia interna de la elección* o de la *maximización del interés propio*. El enfoque utilitarista explica la maximización del interés propio a partir de una noción mental y subjetiva de utilidad. Por su parte, el método de la PR se centra en la consistencia interna de la elección mediante una función de elección, una relación binaria de preferencia y la restricción del WARP. Así pues, atendiendo a las *propiedades de racionalidad de la elección* (supra, § 3.3. y § 4.1.) y a los *supuestos estándar de la preferencia* (supra, § 5.1.), podemos determinar el carácter “racional” de la *elección racional*: la consistencia alude a las propiedades que condicionan la función de elección y la relación de preferencia. Estas condiciones se comprenden como restricciones de racionalidad, de tal modo que una conducta que las cumpla es una conducta considerada racional desde la perspectiva de los fundamentos epistemológicos de la teoría del comportamiento del consumidor. Existe un modelo estandarizado de agencia racional caracterizado por las restricciones formales expuestas sobre la relación de preferencia y la función de elección. Se dice, por tanto, que un consumidor actúa racionalmente si cumple estas propiedades y supuestos.

Obdulia Torres lo expresa de este modo: “si se dan una serie de propiedades en la relación de preferencias podremos explicar y predecir las elecciones de los sujetos. Estas propiedades vienen definidas en forma de axiomas y caracterizan lo que sería una conducta racional por parte de los agentes” (Torres 2002:27). Lo importantes es hacer notar que tales propiedades condicionan las preferencias y la elección de tal modo que permitan construir los conceptos fundamentales de la teoría del consumidor.

Si se lo entiende con carácter descriptivo, el modelo estándar de agencia racional (particularmente económico) ha sido criticado por no corresponderse con la actuación empírica de los consumidores reales. Ahora bien, no en todos los casos son necesarias todas las restricciones. Proponemos el ejemplo de la transitividad en uno de los primeros pasos de la teoría de la decisión, a saber, la construcción de una función de elección. Sen ([Collective choice and social welfare](#)) hace notar que, si una relación de preferencia no es completa y reflexiva, no puede existir una función de elección. Además, señala que, si tenemos una relación de orden en un conjunto finito, entonces existe una función de elección. Es decir, que si a la reflexividad y la completitud le añadimos la transitividad (y obtenemos un orden), esta última sería condición suficiente para que pueda haber una función de elección sobre un conjunto finito. No obstante, Sen demuestra que no es necesaria la transitividad, puesto que basta con una propiedad más débil que la transitividad (e incluso más débil que la cuasi-transitividad) para que se pueda definir una función de elección sobre un conjunto finito. Esta



propiedad es la *aciclicidad* (acyclicity). De modo que, si una relación de preferencia es completa y reflexiva, entonces la aciclicidad es una condición suficiente y necesaria. Asimismo, para la obtención de algunos resultados especialmente relevantes de teoría económica no son necesarios algunas de las restricciones más significativas (y criticadas). En este caso tomamos el ejemplo de la demostración que realiza Mas-Colell ([An equilibrium existence theorem without complete or transitive preferences](#)) del teorema de la existencia del equilibrio walrasiano, sin la necesidad de asumir una relación de preferencia transitiva y completa.

Todo ello muestra que existen distintos modelos de agente racional (no solo el estándar, esto es, el más condicionado), los cuales se pueden construir dependiendo de los resultados que se quieran obtener. Existen modelos mucho más débiles y no tan restringidos. Además, la novedad de la PR es la apertura de la posibilidad de validar los axiomas de la relación de preferencia a partir de las elecciones observables. Es preciso remarcar que en este marco teórico la relación de preferencia restringida de este modo no es una propuesta antropológica general, sino un examen de las condiciones racionales de las distintas versiones de la teoría del consumidor, cada cual con sus correspondientes supuestos. Se deben diferenciar, por consiguiente, el plano ontológico sobre el que tratan las predicciones de la teoría microeconómica positiva y el plano cognoscitivo de la agencia racional sobre el que tratan los fundamentos racionales de la misma teoría. Uno trata sobre situaciones reales para predecir y otro trata sobre teorías y su consistencia lógica.

### 5.3. Utilidad y PR

Para Edgeworth ([Mathematical psychics](#)), obrar por el propio interés es la condición fundamental de la ciencia económica. Se trata de un supuesto de maximización de utilidad en determinadas situaciones, tales como las relaciones contractuales, la guerra o los mercados. Asimismo, es habitual considerar, tal como expresa el propio Samuelson ([Foundations of economic analysis](#)), que el punto de inflexión entre la economía moderna y la clásica se encuentra en la introducción de la teoría subjetiva del valor con la llamada *revolución marginalista* de Jevons, Menger y Walras. Es importante distinguir sus respectivas contribuciones y perspectivas como observa Jaffé ([Menger, Jevons and Walras De-Homogenized](#)); no obstante, para nuestros objetivos es suficiente con señalar que esta perspectiva subjetivista recoge la noción de utilidad de la filosofía moral utilitarista (Bentham, Mill y Sidgwick, entre otros) y la operativiza en la ley de la utilidad marginal decreciente. Esta ley establece una relación de proporción inversa entre la consumición de una unidad adicional de un determinado bien y la utilidad que dicho bien proporciona. El valor subjetivo que el individuo otorga a las unidades que se encuentran en el margen va decreciendo en función del aumento de dichas unidades. Como puede observarse, la noción de utilidad subjetiva es el núcleo conceptual de dicha ley.

Para exponer su *teoría pura del comportamiento del consumidor*, Paul Samuelson examina previamente la evolución del concepto de utilidad para “saber claramente lo que contiene el análisis de utilidad convencional, aunque solo sea para comprender las consecuencias de negar su validez” (Samuelson 1947:90). No podemos detenernos en este recorrido. Sin embargo, nos interesa destacar que él señala una tradición que “para evitar la asimetría de emplear cualquier índice de utilidad favorecido, muchos autores (Pareto, WE Jhonson, Hicks y Allen, *et al.*) han sugerido que se emplee una notación que dependa solo de los elementos invariantes del campo de preferencia



ordinal, a saber, los *loci* de indiferencia” (Samuelson 1947:95). Asimismo, marca la construcción de las funciones de demanda como el objetivo del campo de las preferencias.

Tal como Hands explica ([Paul Samuelson and revealed preference theory](#)), el aspecto fundamental de la teoría de Samuelson es el WARP como condición de consistencia en las funciones de demanda del consumidor. Houthakker definió el SARP ([Revealed preference and the utility function](#)), al cual si se le añaden las condiciones de Samuelson es equivalente a la teoría de la utilidad ordinal. Con el SARP fundamentando el enfoque de las curvas de indiferencia, las funciones de demanda condicionadas por las restricciones empíricas serían las mismas que las que parten, en el enfoque utilitarista, de la maximización que realiza el consumidor de una función de utilidad ordinal ceñida a una restricción presupuestaria y siguiendo ciertas pautas racionales de conducta. Más adelante, Afriat ([The construction of utility functions from expenditure data](#)) elaboró el GARP (Generalized Axiom of Revealed Preference) con el cual vinculó la preferencia revelada con la construcción de la función de utilidad, de tal manera que posibilitaba un análisis empírico de la demanda.

En el contraste entre el enfoque utilitarista y el enfoque de la preferencia revelada, en primer lugar, destacamos la crítica más inmediata al enfoque utilitarista, la cual consiste en una acusación de incurrir en una proposición tautológica, por la cual el pretendido carácter observacional de la conducta en el mercado se perdería. Samuelson expresa que “el comportamiento del consumidor en el mercado se explica en términos de preferencias, que a su vez se definen solo por el comportamiento” (1947:91). De inmediato el propio Samuelson destaca que puede ser una definición circular y, al ser tautológica, carece de base empírica. Sería equivalente a decir que “las personas se comportan como se comportan”. En contraste, la posición más conductista representada por la PR podría solventar la dificultad expresando que, en último término, se llega a la eliminación de las preferencias como estados mentales y nos quedamos exclusivamente con la conducta de la elección. No obstante, esta perspectiva conductista y positivista se ha abandonado en favor de una mera explicación de las preferencias a partir de la conducta de la elección.

En segundo lugar, respecto a las ventajas, es claro que el uso de la PR no aporta ni pretende aportar un rendimiento explicativo mayor que las estrategias utilitaristas, puesto que ofrece un enfoque teórico equivalente al de la utilidad ordinal. Por un lado, el método de la PR no incrementa el poder predictivo de la teoría de la racionalidad y, a su vez, añade elementos explicativos no presentes en el enfoque utilitarista. Por otro lado, hay que destacar que la importancia de eliminar la noción de utilidad no radica en la capacidad explicativa, sino en la ventaja que supone no fundamentar una teoría en conceptos subjetivos internos no susceptibles de ser comparados y medidos, tales como la utilidad. Esto quiere decir que evita lo que hemos denominado *el problema de la utilidad*. Respecto a la metodología la PR, esta no ofrece una explicación más parsimoniosa, sin embargo, respecto a las entidades comprometidas, es un enfoque que elimina causas: la previsión de la utilidad como causa de la elección. Aplicando simplemente la navaja de Ockham, no debemos multiplicar sin necesidad tipos entidades (en este caso mentales subjetivas), dado que los enfoques de la PR y de la utilidad ordinal son equivalentes. Además, la ventaja más clara es la de ofrecer un marco que permite la validación empírica.



## 6. Críticas y objeciones a la PR

En el párrafo 2 se estableció que tomamos la PR en el contexto de un programa de investigación (SRP) que comprende varias teorías y diversas motivaciones. En particular, nos hemos centrado en la PR como un enfoque de la teoría de la decisión y su relación con la utilidad. Existen desarrollos y contribuciones que no corresponden a Samuelson, no obstante, nos hemos ceñido al SRP para acotar el estudio y, dentro de éste, hemos prestado particular atención a Samuelson ([A note on the pure theory of consumer's behaviour](#)) donde se formula por primera vez el WARP.

Las objeciones han sido muy diversas. Por lo tanto, es necesario tener en cuenta este panorama para comprender los puntos en los que las distintas críticas inciden. A continuación, recogemos de forma sintética las críticas más destacadas, que son pertinentes para los objetivos de nuestra investigación. No se trata de todas las críticas posibles o existentes, sino de las pertinentes.

Hausman ([Revealed preference, belief and game theory](#)) distingue entre “preferencia\*” interpretada como relación binaria formal y “preferencia” en tanto que contenidos mentales subjetivos de la *folk psychology* para explicar que (i) la elección no revela la “preferencia” y (ii) la economía necesita preferencias subjetivas. La “preferencia” incluye a la “preferencia\*” y a las creencias, de tal modo que de la elección no se puede inferir una “preferencia\*” pura. La crítica de Hausman se puede reconstruir del siguiente modo: en el método de la PR, dado WARP y el conjunto de elección, se sigue la revelación de las preferencias. Puesto que la elección no revela la “preferencia”, la estrategia de Hausman es constatar que R no puede interpretarse como una “preferencia” en el método de la PR. La réplica en favor de la PR es inmediata: la relación de preferencia empleado en el análisis económico basado en la PR no es la noción de “preferencia”, sino el concepto técnico de “preferencia\*”. El sentido lógico técnico-económico no es el mismo que el subjetivo de la *folk psychology*. La contrarréplica señala que la predicción del comportamiento de los agentes económicos debe tener en cuenta los motivos de sus actos, de tal modo que el análisis debe involucrar tanto las expectativas como las probabilidades subjetivas y las creencias, por lo cual la “preferencia\*” es insuficiente.

Torres ([El problema de las reglas de elección social en la teoría de la acción colectiva](#)) llama la atención sobre tres objeciones: (i) la violación del WARP puede no deberse a una conducta irracional, sino a aspectos como cambios en los gustos y se necesitaría una cantidad ingente de observaciones para contrastar las hipótesis; (ii) existen problemas en dilemas de elecciones estratégicas, como el dilema del prisionero, puesto que supone una escisión entre la elección racional y la situación máximamente útil; y (iii) se dejan al margen cuestiones externas a la propia consistencia de la elección, lo cual no revelaría las preferencias de los agentes reales que incluyen normas morales o sociales.

Sen ([Internal consistency of choice](#)) argumenta que no es posible determinar si una función de elección es consistente o no sin referir a algo externo, al conjunto de elección tales como objetivos, valores o normas. Esta es la idea que está de fondo en la crítica 2.(iii) y de modo general en la crítica 1. No obstante, Sen no impugna con esta objeción en particular la noción de PR, sino que objeta contra el enfoque que caracteriza la racionalidad del agente económico a través de la mera consistencia interna de la elección, el cual es el propio de la PR.



Sen distingue entre el nivel de los fundamentos y el nivel práctico para exponer sus objeciones a la consistencia interna de la elección. En el nivel fundamental señala la dificultad de la comprensión de los actos de elección como declaraciones o sentencias (*statements*) susceptibles de ser lógicamente consistentes. En el nivel práctico Sen explica que, por ejemplo, la siguiente situación no cumpliría el WARP, el SARP, la binariedad, ni la propiedad  $\alpha$ :  $f(\{x,y\}) = \{x\}$  y  $f(\{x,y,z\}) = \{y\}$ . Sin embargo, no sería una conducta irracional si es consistente respecto a propiedades externas. Para hacerlo habría que tener en cuenta el contexto particular y al agente real. Es decir, no se puede saber si una conducta es consistente o no sin tener en cuenta propiedades externas. En particular, el incumplimiento racional de las propiedades de racionalidad puede deberse a las *elecciones posicionales*, al *valor epistémico del menú* y a la *libertad de rechazar*. De este modo, es posible que una elección sea internamente consistente pero irracional, en cuanto a la adecuación entre la finalidad que el agente pretende obtener y las acciones tomadas.

Sen ([Rational fools](#)) muestra la necesidad de incluir los compromisos personales dentro de los distintos aspectos del comportamiento. En este caso (a diferencia del punto anterior) no se trata de determinar si la conducta de las personas es racional o no, sino en poner de relieve una dimensión que no tiene por qué asumir o excluir los condicionamientos racionales al ordenar las preferencias en rankings. De lo contrario, cabe la posibilidad de que el agente sea un “tonto racional” (*rational fool*) por no tomar en consideración las estructuras sociales reales. En este caso se trata de una crítica a una teoría de la racionalidad en general y no de la PR en particular.

Ahora bien, lo pertinente para una crítica de la PR es la forma en la que evita la impugnación general. Reconstruimos en tres puntos la observación crítica de Sen ([Rational fools](#)): a) el compromiso de los agentes es una parte inevitable de su comportamiento, b) los compromisos no son necesariamente racionales (aunque tampoco excluyen la racionalidad) y c) el orden de los rankings de las preferencias se establece por conceptos previos tales como los compromisos. Por este motivo, la atención se debe dirigir a los meta-rankings, en lugar de al mero orden de preferencias. Se trata de una crítica dirigida contra cualquier forma de teoría de la racionalidad que se elabore a partir del principio de maximización del interés propio en cada acto económico. Sin embargo, lo específico de la PR reside en que es un método que pretende eludir el egoísmo racional puesto que la función de utilidad se construye asignando valores más elevados a las alternativas de *facto* elegidas. No se basa, por consiguiente, en una utilidad mental y subjetiva orientada por propio interés, sino en la consistencia lógica de las preferencias reveladas. Lo fundamental estriba en que el comportamiento no sea lógicamente inconsistente y no en estados mentales egoístas. El comportamiento de un altruista racional también cumple con dicha condición. La contrarréplica específica frente a la PR la podemos encontrar en que existen vías de información no relativas al comportamiento de la elección sobre las preferencias del consumidor. Y tal información es determinante para los actos en contextos económicos, ya que afectan a la dimensión conceptual en virtud de la cual se estructuran los órdenes de preferencias, donde se encuentran los compromisos personales.

Sen examina “la filosofía detrás del enfoque de la preferencia revelada” (1973:241) y concluye críticamente: (i) el uso de que el comportamiento de la elección revele la preferencia no quiere decir que el comportamiento se pueda explicar haciendo referencia exclusivamente al propio comportamiento; (ii) si este es el propósito, entonces no se necesitarían propiedades de racionalidad añadidas, pues serían redundantes; (iii) es posible que una ‘comparación sistemática de alternativas’ no caracterice el comportamiento de los consumidores y la falta de un patrón de





preferencia puede confundirse con la indiferencia; y (iv) los cálculos racionales individuales pueden no funcionar en entornos estratégicos de decisiones interdependientes entre varios agentes, dado que no considera los compromisos sociales propios de la acción humana real.

Hansson y Grüne-Yanoff ([Preferences](#)) señalan cinco dificultades que afectan a la posibilidad de inferir la preferencia de la elección. Destacamos dos de ellas: (i) existen preferencias que no pueden ser elegidas aun perteneciendo a un conjunto de alternativas; y (ii) existen elecciones no vinculadas a preferencias estables y que conducen a acciones circunstanciales. Para dar cuenta de este tipo de elecciones se necesitarían tener en consideración los estados mentales subjetivos del agente.

Grüne-Yanoff ([The problems of testing preference axioms with revealed preference theory](#)) arguye contra la posibilidad de comprobar tanto la validez de los supuestos de racionalidad de la relación de preferencia como de la hipótesis de la maximización de consumidor, a través de testear las propiedades de la elección. La argumentación se basa en observar que los resultados de las comprobaciones empíricas son ambiguos, puesto que aquello que se testea no es claro.

Respecto al SRP, Wong argumenta que las tres motivaciones teóricas no son consistentes y que es incompleto en tanto que no consigue resolver los problemas: “el problema de derivar los principales resultados de la teoría de la utilidad ordinal sin el uso de utilidad o cualquier otro concepto no observacional; el problema de construir el mapa de indiferencia de un individuo a partir de observaciones del comportamiento del mercado; o el problema de encontrar el equivalente observacional de la teoría de la utilidad ordinal” (Wong 2006:3).

## 7. PR en dilemas estratégicos

Lo compartido por gran parte de las críticas expuestas consiste en mostrar que el comportamiento de los agentes en la realidad económica depende de elementos externos a la mera consistencia interna. Así pues, para el análisis económico sería inevitable tomar en consideración aspectos tales como las creencias, los compromisos o la expectativa de utilidad involucrada en el comportamiento del consumidor. Las críticas habituales se dirigen a estos aspectos y desarrollan alternativas basadas en la epistemología y la psicología del consumidor tales como Kahneman y Tversky ([Prospect theory](#)). No obstante, el método de la PR permite no pronunciarse sobre la naturaleza de las preferencias. Lo que este método elimina es la relevancia en el análisis económico de las nociones mentales subjetivas como la de utilidad. Las críticas se orientan a señalar que sí se trata de nociones analíticamente relevantes. Esta discusión lleva a un impasse, dado que se pretende criticar la PR desde el punto que -precisamente- se elimina con su metodología, en tanto que esta evita que el análisis dependa de las interpretaciones sustantivas que se hagan de los deseos de los consumidores. Por este motivo nos interesa desarrollar otro tipo de crítica de carácter formal, la cual se pone de manifiesto en la relación del método de la PR con los dilemas estratégicos. Otras objeciones relacionadas con entornos estratégicos (supra, § 6: críticas 2 y 5) apuntan hacia la insuficiencia de la consistencia interna. La particularidad de nuestra crítica reside en el efecto de una imposibilidad formal.





## 7.1. Contextos estratégicos

En el contexto de la decisión racional es determinante distinguir entre dos tipos de elecciones, dependiendo del entorno de decisión en función de la probabilidad de que se hagan efectivos los estados de cosas. En primer lugar, las elecciones en entornos *paramétricos* son aquellas en las que la probabilidad de que sucedan ciertas situaciones o estados de cosas es independiente de la acción del agente. Por su parte, las elecciones en entornos *estratégicos* se caracterizan por la dependencia de la probabilidad de que se dé determinado estado de cosas en función de la acción de un segundo agente. Debemos destacar dos rasgos de los entornos estratégicos o *juegos*: la dependencia del conjunto de oportunidades según las acciones y la participación de un mínimo de dos agentes, por lo que las elecciones son interdependientes. Lo que pretendemos exponer consiste en que la definición de la noción de decisión racional basada en la consistencia interna condicionada por el método de la PR se ve afectada por dilemas estratégicos en cuanto a su capacidad de expresar cabalmente el razonamiento económico.

Existen dos métodos principales de representación formal de las elecciones posibles de los agentes involucrados en un juego con las cuales se pueden expresar las estrategias alternativas. Por un lado, la *forma extendida* es un grafo que representa la sucesión de elecciones de un agente según las disyuntivas de cada nodo del árbol. Por otro lado, la *forma normal* es una matriz que presenta las recompensas o pagos según la combinación de las diferentes estrategias de los agentes o *jugadores*. Empleamos la forma normal y nos atenemos a la situación más simple, es decir, la de dos jugadores con dos estrategias posibles:

$$\begin{array}{c|cc}
 & \text{Q} & \\
 & \phi & \psi \\
 \text{S} & \phi & \begin{array}{|c|c|} \hline x,y & x,y \\ \hline \end{array} \\
 & \psi & \begin{array}{|c|c|} \hline x,y & x,y \\ \hline \end{array}
 \end{array}$$

Donde  $\{S,Q\}$  es el conjunto de jugadores,  $\{\phi,\psi\}$  es el conjunto de estrategias y  $\langle x,y \rangle$  son pagos (nótese que es un par ordenado de modo que  $x$  corresponde a un jugador e  $y$  corresponde al otro). Los pagos son las consecuencias de la intersección de las estrategias de los jugadores. Así pues, convencionalmente la expresión “ $\{S_\phi, Q_\phi\} \rightarrow \langle x,y \rangle$ ” denota que el conjunto de estrategias  $\{S_\phi, Q_\phi\}$  produce como consecuencia el par  $\langle x,y \rangle$ , tal que  $x$  es el pago para S e  $y$  es el pago para Q. Los pagos son los estados de cosas que dependen de las acciones o estrategias de los jugadores. Este esquema permite formalizar diferentes situaciones o tipos de juegos en función de la relación de los pagos establecidos para cada par de estrategias. Es particularmente relevante observar que el pago ya está establecido en las propias reglas del juego. No se trata de contenidos mentales de los agentes, sino de un orden de preferencias ya establecido en la propia dinámica de un juego en cuestión. Si simplificamos los resultados de modo simétrico entre los pagos para cada jugador, entonces se generan veinticuatro modelos. Por lo tanto, en la mitad de los modelos  $\{S_\phi, Q_\phi\}P\{S_\phi, Q_\psi\}$  y en la otra mitad  $\{S_\psi, Q_\phi\}P\{S_\psi, Q_\psi\}$ . Los modelos más destacados son:



El juego del punto muerto	$\{S_\psi, Q_\phi\}P\{S_\psi, Q_\psi\}, \{S_\psi, Q_\psi\}P\{S_\phi, Q_\phi\}, \{S_\phi, Q_\phi\}P\{S_\phi, Q_\psi\}$
Ejemplo	$\{S_\phi, Q_\phi\} \rightarrow \langle 1, 1 \rangle, \{S_\phi, Q_\psi\} \rightarrow \langle 0, 3 \rangle, \{S_\psi, Q_\phi\} \rightarrow \langle 3, 0 \rangle, \{S_\psi, Q_\psi\} \rightarrow \langle 2, 2 \rangle$
El juego de las garantías mutuas	$\{S_\phi, Q_\phi\}P\{S_\psi, Q_\phi\}, \{S_\psi, Q_\phi\}P\{S_\psi, Q_\psi\}, \{S_\psi, Q_\psi\}P\{S_\phi, Q_\psi\}$
Ejemplo	$\{S_\phi, Q_\phi\} \rightarrow \langle 3, 3 \rangle, \{S_\phi, Q_\psi\} \rightarrow \langle 0, 2 \rangle, \{S_\psi, Q_\phi\} \rightarrow \langle 2, 0 \rangle, \{S_\psi, Q_\psi\} \rightarrow \langle 1, 1 \rangle$
El juego de la gallina	$\{S_\psi, Q_\phi\}P\{S_\phi, Q_\phi\}, \{S_\phi, Q_\phi\}P\{S_\phi, Q_\psi\}, \{S_\phi, Q_\psi\}P\{S_\psi, Q_\psi\}$
Ejemplo	$\{S_\phi, Q_\phi\} \rightarrow \langle 2, 2 \rangle, \{S_\phi, Q_\psi\} \rightarrow \langle 1, 3 \rangle, \{S_\psi, Q_\phi\} \rightarrow \langle 3, 1 \rangle, \{S_\psi, Q_\psi\} \rightarrow \langle 0, 0 \rangle$
El dilema del prisionero	$\{S_\psi, Q_\phi\}P\{S_\phi, Q_\phi\}, \{S_\phi, Q_\phi\}P\{S_\psi, Q_\psi\}, \{S_\psi, Q_\psi\}P\{S_\phi, Q_\psi\}$
Ejemplo	$\{S_\phi, Q_\phi\} \rightarrow \langle 2, 2 \rangle, \{S_\phi, Q_\psi\} \rightarrow \langle 0, 3 \rangle, \{S_\psi, Q_\phi\} \rightarrow \langle 3, 0 \rangle, \{S_\psi, Q_\psi\} \rightarrow \langle 1, 1 \rangle$

## 7.2. Imposibilidad de la PR

Las consecuencias se pueden comprender desde dos ópticas: desde la combinación de estrategias y desde el pago que reciben los jugadores. Así pues, se dice que la consecuencia de una determinada combinación de estrategias es un *equilibrio de Nash* si cada estrategia no tiene una alternativa que mejore el estado de cosas resultante para el agente teniendo en cuenta las demás estrategias. Dada la estrategia del otro agente, no es posible obtener una mejora si se modifica la estrategia tomada. Se dice que una consecuencia es un *óptimo paretiano* si no existe otro pago que mejore el estado de cosas para un jugador sin que empeore el estado de cosas para el otro jugador. No existe un pago alternativo que sea mejor para los dos jugadores a la vez. Por consiguiente, no son lo mismo el *equilibrio* y la *optimalidad* o una *estrategia óptima* y un *resultado óptimo*. El punto relevante reside en que, en los dilemas estratégicos, la intersección del conjunto de los resultados que son óptimos paretianos y el conjunto de los resultados que son equilibrios de Nash, es el conjunto vacío. Por consiguiente, no hay un resultado que sea simultáneamente un óptimo paretiano y un equilibrio de Nash.

Esta condición la encontramos en el dilema del prisionero. En el contexto narrativo ideado por Tucker  $\phi$  es “callar” y  $\psi$  es “confesar” sobre un delito, las cuales son las estrategias a las que pueden optar dos presos cuyas recompensas en años de prisión siguen la relación  $\{S_\psi, Q_\phi\}P\{S_\phi, Q_\phi\}, \{S_\phi, Q_\phi\}P\{S_\psi, Q_\psi\}, \{S_\psi, Q_\psi\}P\{S_\phi, Q_\psi\}$ . En este esquema la *estrategia dominante* es confesar. Para ser dominante una estrategia, debe satisfacer la condición en virtud de la cual el pago que genera es igual o mayor que el de la otra estrategia, cualquiera que sea la estrategia del segundo agente. Tanto el juego del punto muerto como el dilema del prisionero tienen una estrategia dominante. Sin embargo, el juego de las garantías mutuas y el juego de la gallina no tienen estrategia dominante, puesto que depende de la estrategia del segundo jugador. En el juego del punto muerto la estrategia dominante conduce a un óptimo paretiano. La especificidad del dilema del prisionero, la razón por la cual es un dilema, consiste en que los resultados óptimos paretianos y las estrategias en equilibrio de Nash son excluyentes. Se trata de un teorema de imposibilidad que impide obtener simultáneamente un óptimo paretiano y un equilibrio de Nash. Considerándolo desde los pagos, que un jugador confiese y el otro se calle (y viceversa), son las combinaciones de estrategias que producen óptimos paretianos. Considerándolo desde las estrategias, la confesión de ambos jugadores constituye un equilibrio de Nash.



Dicho dilema supone la imposibilidad formal de una racionalidad estratégica completa. Es una imposibilidad si se entienden los entornos estratégicos en términos de maximización de la utilidad subjetiva de los agentes en los que se infiere la elección a partir de las preferencias de los pagos. Ahora bien, lo que nos interesa poner de manifiesto es que también imposibilita la metodología de la PR en entornos estratégicos: este método debe ser capaz de explicar el orden de preferencias establecido por el propio juego a partir de las estrategias óptimas. El dilema del prisionero genera una escisión excluyente entre la optimalidad y la racionalidad estratégica, de tal modo que la estrategia elegida no revela la preferencia. Para que funcione, las estrategias óptimas deberían producir los resultados óptimos. Cabe diferenciar dos aspectos:

- (i) No es posible establecer cuál es la preferencia que revela la estrategia elegida: al ser conjuntos excluyentes, la estrategia dominante (a saber, confesar) puede indicar una preferencia por el equilibrio de Nash o por el óptimo paretiano.
- (ii) La elección de la estrategia no explica la preferencia ya dada por las reglas del juego: cada jugador sabe que la estrategia dominante del otro jugador también es confesar, de modo que su elección conduce a un estado de cosas que no prefiere:  $\{S_\psi, Q_\psi\}$ . De la elección no se infiere la preferencia.

Hausman ([Revealed preference, belief and game theory](#)) argumenta en contra de la PR, afirmando que un juego no se ha definido hasta que no se determinan las estrategias de los agentes, de modo que se requieren aspectos como la utilidad esperada subjetiva para optar por una estrategia en lugar de otra. Sin embargo, desde la PR se puede replicar señalando que se debe distinguir la definición de un juego y el modo en el que se debe jugar, esto es, no son lo mismo las reglas del juego y las estrategias que se deben adoptar. De este modo, los pagos forman parte de la definición del propio juego y el orden de preferencias entre los resultados viene dado por las reglas. Si se sigue otro orden, entonces no se está jugando a ese juego. El mejor estado de cosas es una situación objetiva en tanto que condición del juego y no un deseo subjetivo mental. Por consiguiente, el método de la PR permite explicar las preferencias desde las estrategias a las que se pueden optar y no son las utilidades las que explican las elecciones. Ahora bien, lo que nos interesa destacar es el contraste entre nuestro argumento y la crítica de Hausman. Mientras que la crítica de Hausman se mantiene en consonancia con el conjunto de críticas que aluden a la necesidad de apelar a elementos ajenos a la mera consistencia interna, nuestro argumento alude a una imposibilidad formal que afecta al método de la PR. El argumento formal contra la PR desde dilemas en entornos estratégicos tiene la ventaja de no recurrir a elementos extraños a su mismo marco de comprensión.

## 8. Conclusión

La conclusión es una respuesta directa a la pregunta de la que parte la investigación: el enfoque de la preferencia revelada no supone una ventaja en su rendimiento explicativo respecto a las estrategias utilitaristas. Las ventajas y limitaciones quedan puestas de manifiesto en la discusión con las objeciones, de modo que la validez de la conclusión no debe buscarse en los motivos teóricos o la corrección formal de las formulaciones pertinentes de la PR. Ahora bien, se trata de una respuesta parcial respecto al método de la PR, puesto que consiste en una conclusión válida dentro de los límites del marco teórico de la presente investigación, a saber: el WARP como restricción de la función de elección para delinear un modelo de agencia racional respecto a las estrategias utilitaristas.



Por un lado, la gran ventaja consiste en ofrecer bases empíricas a la teoría de la decisión mediante la introducción de propiedades de racionalidad (los axiomas de PR) a la función de elección que permiten inferir las preferencias a partir de la conducta observable de la elección. Hay que destacar que la intención no es conseguir explicar más fenómenos, sino elaborar un aparato conceptual equivalente al de la teoría de la utilidad ordinal, pero capaz de prescindir de conceptos problemáticos ajenos a la consistencia interna, tales como la utilidad que alude a contenidos mentales no contrastables ni comparables. Por otro lado, sintetizamos las dificultades en los núcleos en los que inciden las críticas:

- (1) El SRP no consiguió sus objetivos.
- (2) No se puede determinar si una función de elección es consistente sin considerar aspectos ajenos a la mera consistencia interna.
- (3) Existe una imposibilidad formal en los dilemas en entornos estratégicos que impiden la inferencia de las preferencias a partir de las elecciones.

Lo anterior no debe suponer una cancelación del programa de investigación, la eliminación del concepto de PR, ni un freno a los desarrollos y formulaciones actuales con sus distintas motivaciones teóricas. La discusión sigue abierta, como puede verse en la polémica sobre las críticas de Sen y la versión neo-samuelsoniana que recoge Hédoïn ([Sen's criticism of revealed preference theory and its 'neo-samuelsonian critique'](#)). Asimismo, Hands ([Foundations of contemporary revealed preference theory](#)) expone los fundamentos de la denominada "Teoría de la Preferencia Revelada Contemporánea" y tiene en cuenta los desarrollos recientes de los axiomas. Allende los desarrollos contemporáneos de la PR, la actualidad de tal enfoque se hace patente en la discusión de los efectos que tiene sobre la *teoría descriptiva de la decisión*, la comparación crítica con la *Expected Utility* y la *Cumulative Prospect Theory*, así como la posibilidad de su contraste en las investigaciones propias de la *economía experimental*.

## Bibliografía

- Hands, D.W. (2014). Paul Samuelson and revealed preference theory. *History of Political Economy*, 46(1), 85-116. <https://doi.org/10.1215/00182702-2398939>
- Samuelson, P.A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, 5(17), 61-71. <https://doi.org/10.2307/2548836>
- Samuelson, P.A. (1947). *Foundations of economic analysis*. Harvard University Press.
- Sen, A. (1973). Behaviour and the concept of preference. *Económica*, 40(159), 241-259. <https://doi.org/10.2307/2552796>
- Torres, O. (2002). *El problema de las reglas de elección social en la teoría de la acción colectiva. Retribución, justicia y bienestar*. Tesis doctoral, Universidad de La Laguna, Tenerife.
- Wong, S. (2006). *Foundations of Paul Samuelson's revealed preference theory: a study by the method of rational reconstruction*. Routledge.

Recibido el 2 Jun 2021

Acceptado el 14 Sep 2021