

Análisis filosófico y epistemológico de la Teoría de la Decisión

El caso de Prospect Theory

Autor:

Weisman, Diego

Tutor:

Marqués, Gustavo

2011

Tesis presentada con el fin de cumplimentar con los requisitos finales para la obtención del título Doctor de la Universidad de Buenos Aires en Filosofía.

Posgrado

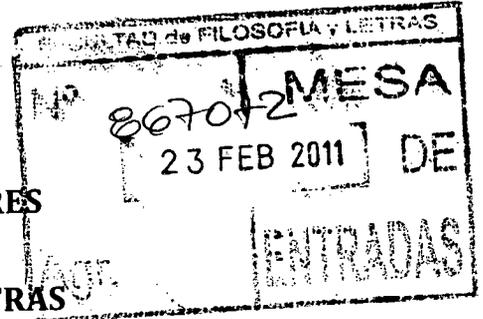
TESIS 16-2-4

Tesis
16-2-4

TESIS DOCTORAL

UNIVERSIDAD DE BUENOS AIRES

FACULTAD DE FILOSOFIA Y LETRAS



**Análisis filosófico y epistemológico de la Teoría de
la Decisión: El caso de *Prospect Theory***

UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE FILOSOFIA Y LETRAS
Dirección de Bibliotecas

Doctorando: Lic. Diego WEISMAN

Director: Dr. Gustavo MARQUÉS

TEST 16-2-4

INDICE

1-Introducción General y Tesis a Sostener

PRIMERA PARTE: Teorías de la Decisión. *Expected Utility Theory*, *Behavioral Model* y *Prospect Theory*.

Capítulo I: Antecedentes: El origen de la hipótesis de utilidad esperada

- 1.1-Valor esperado
- 1.2-Límites del valor esperado: La Paradoja de San Petersburgo
- 1.3-El surgimiento de la utilidad esperada
- 1.4-Herencia de la hipótesis de Bernoulli

Capítulo II. Teoría de la decisión convencional (EUT). Anomalías.

- 2.1-Introducción
- 2.2-Descripción de EUT
- 2.3-El surgimiento del *Golden Dream*
- 2.4-Impacto de EUT
- 2.5-Funciones de EUT
- 2.6-Evaluación empírica de EUT: La *Paradoja de Allais*
- 2.7-Dos clases de contraejemplos: anomalías y paradojas
- 2.8-Evaluación metodológica de los contraejemplos
- 2.9-Critica de EUT como teoría *descriptiva* de la decisión. El problema de la flecha del *Modus Tollens*
- 2.10-Descubrimiento de *preference reversals*

Apéndice I: Procedimiento de Medición de la Utilidad

Apéndice II-Características Generales de los Modelos de Decisión Racional

Capítulo III. *Bounded Rationality*. El modelo de Simon, 1955

- 3.1-La crítica del “irrealismo” de los supuestos
- 3.2-*A Behavioral Model of Rational Choice*
- 3.3-Naturaleza del proyecto
- 3.4-El proceso racional en Simon (1955)
- 3.5-Racionalidad de la Decisión
- 3.6-Incompatibilidad Psicológica
- 3.7-Derivaciones ulteriores

Capítulo IV: *Prospect Theory*

- 4.1-Introducción
- 4.2-Intuición y razonamiento
- 4.3-Sistema I y Sistema II
- 4.4-El *Error de Bernoulli*
- 4.5-Función valor y función de utilidad
- 4.6-Función Ponderación
- 4.7-Proceso de toma de decisión en *Prospect Theory*
- 4.8-La predicción de *Preference Reversals*
- 4.9-Decisores Extensionales
- 4.10-Principio de Extensionalidad
- 4.11-La estrategia convencional
- 4.12-La estrategia heterodoxa: *Prospect Theory* como teoría puramente *descriptiva* de la decisión
- 4.13-Predicción convencional y predicción en *Prospect Theory*

SEGUNDA PARTE: Análisis filosófico y epistemológico de *Prospect Theory*

Capítulo V. Características del cambio teórico de EUT a *Prospect Theory* - Acumulación vs. Revolución

- 5.1-Introducción
- 5.2-El acercamiento entre Psicología Cognitiva y Teoría de la Decisión Racional
- 5.3-EUT como caso especial de *Prospect Theory*
- 5.4-Inconmensurabilidad y Cambio Teórico
- 5.5-La reacción disciplinar
- 5.6-Cambio de lógica subyacente
- 5.7-Quine y el cambio teórico
- 5.8-Ontología
- 5.9- Clases de Individuos y Racionalidad

*Apéndice I: Un sistema de lógica para *Prospect Theory**

Capítulo VI: Racionalidad

- 6.1 Racionalidad en condicione de *certeza*: Teoría de la utilidad ordinal

- 6.2-Certeza e Incertidumbre
- 6.3-Racionalidad en *Expected Utility Theory*
- 6.4-Racionalidad Económica vs Racionalidad Clásica
- 6.5-Racionalidad Económica ¿Formal o Sustantiva?
- 6.6-*Prospect Theory* como teoría de la decisión no racional
- 6.7-Racionalidad como maximización en Economía *Mainstream*
- 6.8-Racionalidad en *Prospect Theory*
- 6.9-Posible carácter normativo de *Prospect Theory*

Apéndice: Las teorías de la decisión como instanciaciones de la Lógica de la Situación popperiana.

- A- Meta Modelos (MM) y Modelos Empíricos (ME)
- B- Meta Modelos. El caso de La Lógica de la Situación
 - B.1. Modelo Empírico I: Teoría de la Utilidad Esperada (EUT)
 - B.2. Modelo Empírico II: *Prospect Theory* (PT)
- C-Comentario sobre los Modelos Empíricos descriptos
- C- Deshaciendo un enredo

Capítulo VII. Manipulación de las decisiones individuales y Paternalismo

- 7.1-El problema de las Encuestas
 - 7.1.1-El mito de las preferencias libres y autónomas
- 7.2-Preferencias Democráticas e Incompetencia de la Ciudadanía
- 7.3-Del elector económico al elector político
- 7.4-Paternalismo libertario
 - 7.4.1-La Utilidad Experimentada
 - 7.4.2-Cambio en el concepto de Utilidad
- 7.5-De la Teoría de la Decisión a la Teoría Económica: *Behavioral Economics*
- 7.6- La necesidad normativa del concepto convencional de utilidad
- 7.7-El Paternalismo Libertario y las elites
- 7.8-*Libertad* en el Paternalismo Libertario

VIII - Tesis defendidas y nuevas líneas de investigación

IX- Fuentes utilizadas

**Análisis filosófico y epistemológico de la Teoría de la
Decisión: El caso de *Prospect Theory***

1-Introducción General y Tesis a Sostener

La teoría de la decisión ha experimentado notables cambios en las últimas décadas (Starmer, 2000; Peterson, 2009) en particular en las teorías referentes a la toma de decisión *individual*. La teoría convencional de decisión racional bajo riesgo, conocida como *Expected Utility Theory* (EUT) ha sido ampliamente desafiada, y un número creciente de modelos económicos utiliza formulaciones alternativas. Una de ellas, *Prospect Theory*, ha ganado impulso y actualmente forma parte central de una serie creciente de modelos agrupados bajo el rótulo de *Behavioral Economics* (Camerer, Loewenstein y Rabin, 2004). A pesar de presentárselas como teorías alternativas, ni la naturaleza del cambio epistemológico involucrado en el pasaje de EUT a *Prospect Theory*, ni las implicaciones filosóficas de esta última han sido tematizada desde el punto de vista epistemológico.

Durante las primeras décadas del siglo XX la teoría de la decisión existente en Economía se limitaba al dominio de las elecciones bajo condiciones de certeza. No existía entonces una visión disciplinar consensuada acerca de las elecciones bajo riesgo e incertidumbre, pese a que el interrogante acerca del curso de acción racional en contextos de incertidumbre había recibido respuestas de creciente sofisticación ya a partir de los trabajos seminales de Pascal (1670) y Bernoulli (1738). Esta situación se revirtió en 1944 con la publicación de *Teoría de los Juegos y Comportamiento Económico*, de Morgenstern y Von Neumann. Allí se propone una teoría axiomática de la elección bajo incertidumbre (*Expected Utility Theory*) en la cual un agente racional – el *homo economicus*- decide sobre la base de un complejo cálculo maximizador. El trabajo fue recibido con entusiasmo por la profesión, porque por primera vez podían subsumirse bajo un cuerpo de hipótesis unificados fenómenos que hasta entonces se explicaban por razones psicológicas, con lo cual la economía ganaba homogeneidad (Friedman y Savage, 1953). Por otra parte, la nueva teoría incorporaba un concepto de

utilidad *medible*, lo que fue percibido como un avance en dirección a la cientificidad de la disciplina (Friedman, 1954).

De manera semejante a la teoría de la decisión racional bajo condiciones de certeza, *Expected Utility Theory* identifica *racionalidad* con *consistencia*: la racionalidad de los individuos se funda exclusivamente en las propiedades estructurales del ordenamiento de las preferencias y en las propiedades formales del conjunto de creencias con que se los modela. En particular, ninguna restricción es impuesta por la teoría al *contenido* de las preferencias individuales (Hausman, 2001). Por otra parte, como los individuos postulados son optimizadores, EUT permite conservar la maximización dentro del núcleo duro del concepto económico de racionalidad, integrando esta perspectiva con las intuiciones básicas que los economistas habían venido desarrollando desde Adam Smith en adelante (Boland, 1981; Arrow, 1987, Weisman, 2009). Una parte significativa de los modelos económicos *mainstream*¹ actuales continúan incorporando agentes racionales en este sentido. Lo que es más importante, a diferencia de otros análisis de la decisión, EUT, considerada como teoría descriptiva, *posee contenido empírico*, lo cual ha permitido testarla de manera sistemática y controlada. Los experimentos, de manera típica, se llevan a cabo sometiendo a voluntarios a elecciones entre “premios” probables y observando sus elecciones.

Históricamente, la posibilidad de testear la teoría abrió un panorama singularmente atractivo para la profesión, por las obvias connotaciones vinculadas a la posibilidad de avance científico y la perspectiva de contar con una pieza teórica confirmada empíricamente. Los frutos del testeo, sin embargo, fueron diversos a los esperados, y ya en la década del 70 comenzaron a hacerse

¹ El término *mainstream*, así como el adjetivo *convencional* o *tradicional*, designan en lo que sigue a la economía neoclásica y a los modelos agrupados bajo la llamada “síntesis neoclásica” (Samuelson, 1970, p.922).

públicos resultados experimentales anómalos (Allais, 1953; Grether y Plott, 1979)

La reacción de la mayoría de los economistas ante las anomalías fue defensiva: inicialmente los resultados adversos fueron atribuidos a falta de incentivos o pobre diseño experimental, pero ni la presencia de incentivos monetarios ni diseños experimentales más complejos lograron hacerlos desaparecer. Esta situación persistente de desajuste empírico condujo con el tiempo a cierto malestar teórico y a sucesivos intentos de dejarla de lado a favor de teorías alternativas (Machina, 1982; Loomes y Sugden, 1982; Levy, 1980) Una de ellas, conocida como *Prospect Theory* (Khaneman y Tversky, 1979) ha mostrado un sugestivo poder predictivo y es comparativamente exitosa desde un punto de vista empírico. El agente racional que se encuentra en la base de EUT es reemplazado en *Prospect Theory* por un sujeto con características psicológicas particulares y poseedor de una racionalidad limitada (*bounded rationality*), lo que incidentalmente hace justicia a un viejo reclamo vinculado a la falta de realismo de los modelos estándar (Simon, 1955)

Uno de los puntos de quiebre entre *Prospect Theory* y EUT es que ésta última es una teoría de la decisión *racional*. Esto explica la función *normativa* que le conceden muchos economistas, en tanto establece cuál es la decisión óptima a tomar en contextos de incertidumbre (Starmer, 2000). Por otra parte, los agentes postulados por la teoría convencional poseen características altamente “irrealistas”, tales como conocimiento perfecto, capacidad de cálculo instantánea, y acceso a la información ilimitado (Simon, 1955). Por el contrario, es común asumir que *Prospect Theory* es una teoría de la decisión *no racional*, porque representa decisores con racionalidad limitada y carentes de una condición necesaria para ser representados por cualquier teoría de la decisión *racional*: el principio de extensionalidad (Kahneman, 2003) La ambigüedad del término “racionalidad” dificulta sin embargo un debate claro, y pueden encontrarse elementos, incluso dentro de sus mismos creadores, que operan a favor de una

racionalidad implícita en las decisiones representadas por *Prospect Theory* (Tversky y Kahneman, 1981; Kahneman, Wakker y Sarin, 1997).

Por otra parte, suele afirmarse que *Prospect Theory* guarda una relación particular con EUT, en el sentido de que ésta sirve como *benchmark* para contrastar los modelos psicológicos de decisión económica (Camerer *et. al.*, 2003) Lo anterior sugiere una relación de *continuidad* entre ambas teorías. Coincidiendo con esta interpretación, Starmer (2003) establece que *Expected Utility Theory* es un caso límite de *Prospect Theory*, a la cual podría ser reducida bajo ciertas circunstancias. Schoemaker (2000) señala una divergencia más profunda, al apuntar a la pérdida de transitividad como característica diferencial de *Prospect Theory*; pero su distinción juega un rol meramente taxonómico y no se involucra con la naturaleza epistemológica de la ruptura entre ambas teorías. Marqués y Weisman (2008) sugieren que la relación conflictiva entre ambas teorías es mejor comprendida como un *cambio revolucionario*, antes que acumulativo, debido a las mutaciones en la referencia experimentadas por ciertas variables fundamentales, como el término “utilidad”, que posee características semánticas muy distintas en ambas teorías. Si esto fuera cierto, cabría esperar una relación de inconmensurabilidad del tipo Kuhniano en el pasaje de *Expected Utility Theory* a *Prospect Theory*.

El reemplazo de la visión acumulativa por la revolucionaria permite además iluminar zonas disciplinares un tanto menos obvias donde también impactan los cambios conceptuales que resultan del pasaje de EUT a *Prospect Theory*. El sujeto racional, que expresa preferencias bien formadas en sus decisiones, está en la base no sólo de EUT y de numerosos modelos de la economía neoclásica, sino que también desempeña un papel relevante en la defensa del liberalismo económico no intervencionista, e incluso del liberalismo político. La racionalidad de los individuos y la inexistencia de sujetos con conocimiento privilegiado hacen *necesaria* (además de deseable) la libertad de elección, volviendo la manipulación improbable a largo plazo. Por el contrario, el sujeto modelado por

Prospect Theory es incapaz de tomar decisiones óptimas en el sentido tradicional, y un tipo de intervencionismo paternalista, apoyado en asimetrías de conocimiento, se perfila como necesario. Lo que es peor, sus preferencias pueden en algunos casos resultar inexistentes o contradictorias, lo que implica desafíos de fuste para la Teoría Política de la democracia.

Sobre la base de lo anterior defenderemos las siguientes tesis:

1-Los modelos de racionalidad limitada (*bounded rationality*) son considerados alternativas más *realistas* a los modelos tradicionales que, basadas en *Expected Utility Theory*, asumen agentes completamente racionales. El artículo de *Bounded Rationality* de Simon (1955) se convirtió en un punto basal para la nueva corriente porque incorpora agentes con capacidades computacionales finitas y acceso a la información limitada, aproximándose a las condiciones en que adoptan decisiones los agentes “reales”. Brevemente expresada, la tesis de Simon (1955) es que los agentes reales no pueden maximizar, dado que su racionalidad y conocimiento es limitado; en lugar de eso, satisfacen. Sin duda, esta visión seduce por lo familiar y hasta parece ser descriptivamente precisa. Sin embargo, si es examinada cuidadosamente carece de contenido empírico –o, al menos, tiene *menor contenido informativo* que los modelos basados en EUT.

2-*Prospect Theory* es una teoría empírica capaz de resolver anomalías que ni la teoría convencional (EUT) ni sus variantes “convencionales” pueden acomodar (Starmer, 2000). Para ello modela individuos “irracionales” en un sentido particular: son agentes que violan el *Principio de Extensionalidad* de la lógica estándar en sus elecciones. Tradicionalmente se ha considerado que la Extensionalidad es una nota “fuerte” de la racionalidad, y por lo tanto un requisito necesario para considerar a un agente como racional (Kahneman y Tversky, 1979). Se mostrará que los agentes de *Prospect Theory* violan extensionalidad y son sin embargo *racionales* en un sentido relevante para la Economía *mainstream*.

3-El patrón de cambio teórico que conduce de *Expected Utility Theory* a *Prospect Theory* no ha sido explícitamente tematizado, y suele ser presentado de manera un tanto vaga como el pasaje de una teoría a otra teoría alternativa con supuestos más “realistas”. (i.e., con agentes más semejantes a los individuos “reales”). Prevalece una visión *acumulativa* del cambio teórico (Schomaker, 2003; Starmer, 2000). Se discutirá esta interpretación, argumentando que el cambio teórico es posible de ser modelado siguiendo un patrón *revolucionario* en el sentido de Thomas Kuhn (1970)

4-En relación con el punto anterior, se pondrá en evidencia posibles fenómenos de inconmensurabilidad concomitantes al cambio teórico en el campo de la decisión humana: específicamente se argumentará que el término “utilidad” posee significados claramente distintos (e incompatibles) en *Expected Utility Theory* y *Prospect Theory*.

5-Desde una perspectiva un tanto diferente, Quine (1951) postula que una anomalía recalcitrante puede obligar a realizar cambios profundos en la red teórica. Si bien no ofrece ningún ejemplo concreto, sugiere un esbozo de ilustración para el caso de la mecánica clásica/ mecánica cuántica. Se mostrará que el patrón que identifica Quine posee semejanzas con en el caso del par EUT/ *Prospect Theory*, y que el cambio teórico realizado, lejos de ser periférico, afecta la *ontología* asumida por los agentes en una y otra teoría.

6-El cambio ontológico relativo al sujeto que toma decisiones alcanza la trascendencia suficiente como para afectar a los resultados de la teoría económica *mainstream* y su influencia se extiende desde la economía a la *teoría política* que está en la base de los sistemas democráticos, basados en la competencia electoral de los tomadores de decisión.

7-El agente racional de EUT es heredero de la tradición iluminista, en el sentido de que es autónomo y sus preferencias no dependen del contexto de elección ni de la descripción con que las opciones le son presentadas. *Prospect Theory*

rompe con esta tradición, y modela individuos cuyas preferencias son contexto-dependientes, y en la medida en que dicho contexto les es proporcionado externamente, dichas preferencias son *manipulables* de una manera no evitable.

8-La posibilidad de manipular las decisiones individuales está en la base del éxito predictivo de *Prospect Theory*, lo que incidentalmente obliga a reconceptualizar el sentido epistemológico tradicional de *predicción* para incorporar esta faceta “tecnológica”.

9-La teoría económica basada en *Prospect Theory* incorpora además un intervencionismo de carácter muy diferente al tradicional, ejercido sobre las variables agregadas de la economía, de carácter paternalista. Tal paternalismo es *necesario*, y está basado en la posibilidad técnica de manipular las elecciones individuales.

SECCIÓN I

Teorías de la decisión bajo riesgo e incertidumbre

Capítulo I

Antecedentes: El origen de la hipótesis de utilidad esperada

1.1-Valor esperado

Los orígenes de la Teoría de la Decisión bajo riesgo e incertidumbre se remontan al siglo XVII, y en particular al progresivo interés despertado por el juego de azar en las cortes europeas. Las apuestas involucran *premios* y cierta *probabilidad* de obtenerlos, y es indudable que el atractivo de un juego aumenta con ellos; sin embargo, no es sencillo determinar la manera precisa de operar con ambas magnitudes. ¿Cómo se decide cuál es el precio “correcto” (razonable) que debe ser pagado por una lotería o, alternativamente, cómo se elige racionalmente entre loterías alternativas? Durante mucho tiempo se pensó que la clave para responder esta pregunta era la noción de *valor esperado*. A comienzos del siglo XVII se estimaba que un juego era “justo” cuando todos los jugadores tenían idéntica esperanza matemática. (Landro, 1987; Weisman, 2007)

La noción de valor esperado puede ser ilustrada mediante algunos ejemplos. Tomemos el juego consistente en arrojar una moneda al aire y perder un peso si sale ceca y ganar un peso si sale cara. En el caso de una moneda normal las probabilidades de que se dé uno u otro resultado son 50-50. El juego puede entonces escribirse sintéticamente así: [\$1, 0.5; -\$1, 0.5], colocando cada resultado con su probabilidad asociada. El valor esperado de este juego se obtiene multiplicando el *valor monetario* de cada resultado por su *probabilidad*, y sumando los productos.

$$\text{Valor esperado: } (\$1 \times 0.5) + (-\$1 \times 0.5) = 0$$

En este caso no sería justo que se exigiera pagar una suma positiva por jugar a este juego y no sería racional hacerlo.

Este concepto puede extenderse a juegos en que los resultados no se cancelan mutuamente. Supongamos que el juego consiste en arrojar una moneda al aire y ganar un peso si sale cara o perder 50 centavos si sale cruz. En el caso de una moneda normal las probabilidades de que se dé uno u otro resultado son 50-50.

El juego puede entonces escribirse así [$\$1, 0.5$; $-\$ 0.5, 0.5$]. Su valor esperado es 25 centavos, lo que indica que es racional pagar hasta esa suma para jugar el juego (pero no más).

Veamos un tercer caso. Tirada de un dado (normal). Se recompensa a quien lo arroja con el valor que aparece en la cara superior del dado (la que queda para arriba). Si sale un 1 gana \$1, Si sale el 5 gana \$5, etc. El valor esperado de este juego es \$3.5. Ello significa que el agente racional debería pagar hasta esa suma (pero no más) para jugar ese juego.

En general, el valor esperado de una juego ($R_1, p_1; R_2, p_2; \dots; R_n, p_n$), puede ser representado de esta manera:

$$\sum R_i \cdot p_i \quad [1]$$

Supongamos que a un individuo se le brinda la oportunidad de participar gratuitamente en sólo uno de los juegos mencionados. ¿Por cuál debería decidirse? La teoría del valor esperado sugiere una regla sencilla para elegir entre ellos: deben ser ordenados (rankeados) según su valor esperado y se debe elegir aquél que obtiene el valor mayor (o, dicho con mayor precisión, no se debe elegir un juego existiendo otro de valor esperado mayor). Sintéticamente, los individuos deben elegir entre loterías de manera de *maximizar el valor esperado*.

Cabe remarcar algunas de las características salientes de la elección entre juegos sobre la base del concepto de valor esperado:

- 1) La suma de las probabilidades asociadas a los resultados debe ser igual a la unidad. Más formalmente, $\sum p_i = 1$. Las probabilidades asociadas a los resultados son probabilidades en el sentido técnico del término: deben satisfacer los axiomas del cálculo de probabilidades.
- 2) Las probabilidades están dadas y son conocidas por quien toma la decisión (suele llamárselas probabilidades *objetivas*).

3) Los resultados (en este caso, monetarios) también son conocidos sin distorsión (el sujeto valúa una magnitud x como x : no la sobrevalúa ni subvalúa). Sumado al punto anterior esto determina que el *valor esperado* de un juego sea el mismo para cualquier sujeto, y pueda ser representado con un único número.

1.2-Límites del valor esperado: La *Paradoja de San Petersburgo*

No todos los juegos son como los que acabamos de mencionar. Nicolas Bernoulli planteó a comienzos del siglo XVIII este juego²:

Se arroja al aire una moneda normal,. Si sale cara en la primera instancia la “banca” paga \$1, si sale cara recién en la segunda, paga \$2, si sale cara recién en la tercera, paga \$4, etc, sin límite (el juego recién termina cuando sale la primer cara). Como la probabilidad de que salga cara la primera vez es $\frac{1}{2}$, que recién lo haga en la segunda vez es $\frac{1}{4}$, en la tercera, $\frac{1}{8}$, etc., el *valor esperado* de este juego $[(\frac{1}{2} \times \$1) + (\frac{1}{4} \times \$2) + (\frac{1}{8} \times \$4) \dots (\frac{1}{2^n} \times \$n) + \dots]$. Es decir, $\sum p_i x_i = \infty$.

Esta es la llamada Paradoja de San Petersburgo³. Como en este caso el valor esperado es infinito⁴, ya no es racional que un individuo esté dispuesto a pagar esa suma (lo que tampoco sería posible) y ni siquiera una muchísimo menor (digamos 1 millón de pesos) para jugar el juego. Daniel Bernoulli, primo del anterior, lo expresa de la siguiente manera:

² Bernoulli, D. 1954 (1738), p. 29.

³La formulación original es la siguiente: “*Peter tosses a coin and continues to do so until it should land “heads” when it comes to the ground. He agrees to give Paul one ducat if he gets “heads” on the very first throw, two ducats if he gets it on the second, four in the third, eight in the fourth, and so on, so that with each additional throw the number of ducats he must pay is doubled. Suppose we seek to determine the value of Paul’s expectation.*” (Bernoulli, D. 1954 (1736), p.31)

⁴Landro,(1987), p. 117. Existe consenso en la actualidad respecto a que no existe el valor esperado de este juego, porque conforma una serie divergente.

“Although the standard calculation shows that the value of [player’s] expectation is infinitely great, it has (...) to be admitted that any fairly reasonable man would sell this chance, with great pleasure, for twenty ducats. The accepted method of calculation does, indeed, value [player’s] prospects at infinitely though no one would be willing to purchase it at a moderately high price.” (1954, p. 32)

Se presenta aquí una paradoja lógica y práctica para la teoría del valor esperado:

- a) la noción de *infinito* torna a la regla de maximización no operativa;
- b) aunque esa suma pudiera pagarse, no parece ser razonable jugar semejante juego. La maximización del valor esperado no es en este caso una respuesta que pueda ser considerada *racional*.
- c) personas razonables no pagarían la cantidad sugerida por el valor esperado, sino una menor (“20 ducados”, sugiere Bernoulli). Desde el punto de vista *descriptivo* la teoría del valor esperado enfrenta un contraejemplo.

1.3-El surgimiento de la utilidad esperada

Para resolver la paradoja de San Petersburgo, Daniel Bernoulli⁵ (1738) trazó una distinción entre un monto cualquiera de dinero y la *utilidad* que le reporta a una persona. Es la consideración *subjetiva* del monto monetario, y no el monto en sí lo que debe ser incorporado en la regla para determinar el valor de un juego:

“(...) the determination of the value of an item must not be based on its price, but rather on the utility it yields. The price of the item is dependent only in the thing itself, and is equal for everyone; the utility, however, is dependent of the particular circumstances of the person making the estimate. Thus there is no

⁵ Daniel Cramer había arribado a una solución semejante a la de Bernoulli en 1728. (Bernoulli, *Op. cit.*, p. 33)

*doubt that a gain of one thousand ducats is more significant to a pauper than a rich man though both gain the same amount”.*⁶

Ganar cierta cantidad de dinero no proporciona la misma utilidad a un rico que a un pobre., o alternativamente, a medida que un individuo se enriquece obtiene menos utilidad de unidades adicionales de riqueza.

Sobre esta base, Daniel Bernoulli introdujo la *Hipótesis de la Utilidad Esperada* (HUE). Desde este nuevo enfoque, lo que los individuos procuran maximizar al elegir entre loterías es la utilidad esperada. Aunque se mantiene la idea de que las probabilidades involucradas en el cálculo son objetivas, las “utilidades” que se obtienen de los montos numéricos dependen de dichos montos pero también de la riqueza previa poseída por el decisor (además de otros factores).

La hipótesis de Bernoulli de que los individuos maximizan su utilidad esperada puede ser expresada de esta manera:

$$\sum U(R_i) \cdot p_i \quad [2]$$

donde $U(R_i)$ es la utilidad asignada al resultado i -ésimo y p_i es su probabilidad asociada. El rasgo distintivo de [2] es que “ $U(R_i)$ ” sustituye a “ (R_i) ” (en otros términos, [2] es una flexibilización o generalización de [1]). Pese a que resolvía la paradoja para la que fue propuesta, no fue tomada con seriedad, por varias razones:

a) presupone utilidad cardinal, que parecía de difícil o imposible medición. Ello motivó que hacia 1936 la teoría económica sustituyera esta noción por la de utilidad ordinal.

b) asumía que el dinero tenía utilidad marginal decreciente: a medida que un sujeto se enriquece, unidades adicionales idénticas de riqueza le proporcionan

⁶ *Ibidem*, pág. 25.

cada vez menos utilidad. El supuesto de concavidad para la utilidad del dinero sugirió dos objeciones particularmente relevantes para la teoría económica. En primer lugar, se cuestionó la hipótesis de que la curva de utilidad es en general cóncava (es decir, salvo excepciones que Bernoulli desecha por hipótesis, siempre es cóncava en todos sus tramos), por resultar incompatible con los hechos: la gente apuesta y toma riesgos. En segundo lugar, aunque esta fuera la forma de la curva, ¿cuál sería exactamente su pendiente? Según Stigler, Cramer consideraba que la utilidad proporcionada por $\$x$ descendía de manera proporcional a la raíz cuadrada de dicho ingreso. Edgeworth era menos preciso, pero también creía que la utilidad generada por el ingreso descendía de manera más rápida que lo que creía Bernoulli⁷. Estas discrepancias motivaron que Stigler considerara que las distintas hipótesis que se adelantaron por ese entonces acerca de la forma particular de la función de utilidad eran completamente arbitrarias y construidas a partir de simple introspección. Esa arbitrariedad nos remite nuevamente a la objeción inicial: no se disponía por ese entonces de una manera objetiva de *medir* la utilidad y, en consecuencia, se recurría a la imaginación descontrolada⁸.

c) suponía, además, que los aumentos en utilidad asociados a una cierta ganancia X son menores que las desutilidades asociadas a la pérdida de ese monto. Estas características implicaban que cualquier juego (aún uno “justo”) tenga utilidad esperada negativa. No es racional que un maximizador apueste⁹. Esto tiene la

⁷ Véase Stigler (1950)

⁸ Un reclamo recurrente de la Escuela de Chicago, existiendo ya un procedimiento de medición de la utilidad adecuado, es la necesidad de establecer *empíricamente* la forma de la función. Esto explica en parte sus continuas incursiones en el debate metodológico y la utilización, especialmente de parte de Friedman, de una argumentación casi popperiana en defensa del contenido empírico de la teoría de la utilidad y de la necesidad del test empírico. Véase Friedman y Savage (1948).

⁹ Consideremos una variante del juego mencionado anteriormente: $[\$1, 0.5 ; -\$1, 0.5]$, cuyo valor esperado es cero. El sujeto pierde $\$1$ y gana $\$1$ si sale cara. Si se acepta la doctrina de la utilidad marginal decreciente, $u(-\$1) > u(\$1)$. En estas condiciones, un maximizador de la utilidad esperada nunca tomaría riesgos si no recibe un pago adicional para ello. Contra esta

desagradable consecuencia de que coloca al ámbito de las decisiones riesgosas fuera del campo de la teoría económica: de hecho, Smith y Marshall explicaban las decisiones de riesgo mediante factores psicológicos o sociológicos.

d) no existía razón para pensar que la hipótesis de la utilidad esperada fuera *normativamente* adecuada. La *racionalidad* de la conducta consistente en maximizar la utilidad esperada no tenía fundamento.

d) Finalmente, pese a su plausibilidad intuitiva se trataba de una hipótesis “suelta”, no conectada de manera sistemática con un cuerpo de conocimiento más amplio que le brindara apoyo adicional. Parece tratarse de una hipótesis ad-hoc introducida de apuro para ajustar un hecho anómalo, pero carente de fundamento en el cuerpo de conocimiento “básico” de la disciplina¹⁰.

1.4-Herencia de la hipótesis de Bernoulli

El enfoque de Bernoulli forjó una larga tradición en teoría de la decisión, especialmente en economía. De hecho, existe una serie de intuiciones que se preservan en la forma moderna de pensar las decisiones en circunstancias de incertidumbre, y que son herencia más o menos directa de Bernoulli. En primer lugar, se identifican sólo dos variables relevantes para explicar las decisiones: los *juicios de probabilidad* y las *preferencias* (utilidades) de los agentes. En segundo lugar, los dos factores seleccionados recortan, por decirlo así, el ámbito de la racionalidad: la hipótesis de la utilidad esperada es una hipótesis para explicar las decisiones *racionales* en condiciones de riesgo. Es además, una versión de la idea de sentido común, pertenecientes a la *folk psychology*, de que las *acciones* (entendidas como comportamientos deliberados dirigidos a alcanzar un cierto

teoría, se observa que muchos sujetos apuestan en juegos “justos”, como el mencionado, e incluso en juegos “injustos”, como las loterías corrientes.

¹⁰ Si se busca una analogía sugerente, piénsese en la hipótesis kepleriana del movimiento elíptico de los planetas antes que este movimiento pudiera ser deducido de la teoría de Newton.

objetivo) pueden ser explicadas únicamente en términos de *creencias* y *deseos*, reflejados en las probabilidades y las preferencias respectivamente. Como se mostrará en el Capítulo 2, EUT podría ser considerada una formulación *precisa* de esta idea para el caso particular de las elecciones riesgosas. En tercer lugar, se mantiene la idea básica, presente en el cálculo del valor esperado, de que las probabilidades y los resultados (en este caso, las utilidades de los mismos) pueden ser adicionados y multiplicados entre sí, y que este procedimiento matemático brindaría una representación adecuada de las preferencias de los individuos en condiciones de riesgo o incertidumbre, a la vez que permitiría predecir sus elecciones. Sobre estas mismas bases se desarrollará posteriormente las distintas versiones de la teoría de la decisión.

Capítulo II

Teoría de la decisión convencional (EUT). Anomalías.

2.1-Introducción

En 1944 John von Neumann y Oskar Morgenstern publicaron *Theory of Games an Economic Behavior*, donde expusieron una teoría axiomática de la elección bajo riesgo –conocida como *Expected Utility Theory*- la cual implicaba la hipótesis bernoulliana de utilidad esperada. Su propuesta atrajo el interés inmediato de los economistas, porque zanjaba el aparentemente irresoluble problema de la medibilidad de la utilidad y, bajo ciertas condiciones bien especificadas, daba lugar a predicciones relativas a la conducta de individuos enfrentados a situaciones de elección entre loterías¹¹ alternativas. En el pasado, la teoría de la utilidad había sido exitosamente aplicada a situaciones de certeza dentro de la economía (típicamente, situaciones de consumo), pero era considerada inútil bajo condiciones de riesgo e incertidumbre. La teoría de Von Neumann y Morgenstern, así como otras teorías equivalentes desarrolladas poco después de su publicación, parecía abrir la puerta para la formidable tarea de unificar el campo de las decisiones en economía. Se esperaba que, adecuadamente desarrollada, pudiera ayudar a explicar aquellas conductas que involucran riesgo, como el juego o los seguros, las cuales hasta el momento estaban consideradas fuera del rango de alcance de la teoría económica, y originadas en reacciones arbitrarias o explicables por factores psicológicos.

Pocos años después de que fuera propuesta, el status de EUT comenzó a ser objeto de debate. Un conjunto de economistas, siguiendo a Milton Friedman, asumió la distinción positivo-normativo, tradicional en Economía, y consideró que la hipótesis de la utilidad esperada era meramente una hipótesis positiva (descriptiva), al tiempo que aceptó la adecuación normativa de los *axiomas* que en EUT la implican¹². Otros *papers* (entre ellos Savage, 1954 ; Marshak, 1950;

¹¹ Una lotería es la representación de una circunstancia de riesgo o incertidumbre, La notación típica suele ser $L: (R_1, R_2; p_1)$, donde $R_1...R_n$ son los resultados posibles y p_n la probabilidad correspondiente al n-ésimo resultado.

¹² Ver Friedman and Savage (1948, 1952), donde prevalece claramente la orientación metodológica de Friedman. Friedman (1953) contiene la exposición más completa de su visión metodológica de las teorías económicas.

Strotz, 1953; Ellsberg, 1961) encuentran los axiomas de EUT persuasivos y plausibles a nivel normativo¹³. Si fueran correctos, la economía podría contar al fin con una teoría *general* de la decisión *racional*, extendiendo los instrumentos que habían probado ser útiles en condiciones ciertas a situaciones inciertas.

Los axiomas a satisfacer por las decisiones son los de ordenamiento (transitividad, completitud y consistencia) más un *principio de independencia*, que garantiza la linealidad de las probabilidades. Estos cuatro axiomas permiten introducir la hipótesis de utilidad esperada, que en términos analíticos resulta formalmente idéntica a la de Bernoulli¹⁴.

La capacidad de EUT de realizar predicciones precisas, una nota inusual para la disciplina económica, permitió a las mentalidades críticas discutirla abiertamente con argumentos empíricos. Allais (1953), Markowitz (1952) y Ellsberg (1961) constituyen trabajos pioneros en señalar algunas de sus dificultades empíricas, conocidas indistintamente como *anomalías* o *paradojas* (ya veremos que en rigor existe una diferencia entre ambas categorías). Los defensores de la estrategia convencional acometieron diversos intentos de disolver los contraejemplos mediante los argumentos tradicionales –falta de incentivos, o mal diseño experimental-, los cuales condujeron a nuevos experimentos, que no hicieron sino revelar la robustez de los resultados anómalos. A medida que ganaban prestigio como contraejemplos empíricos serios, las anomalías y paradojas dieron lugar a un movimiento creciente de cambio teórico (Starmer, 2000).

¹³ “These axioms have strong intuitive appeal. It would seem that every normal person would clearly accept them as precepts of behaviour” (Strotz, 1953, p. 391).

¹⁴ Con el tiempo se desarrollaron algunas diferencias significativas. El concepto de *utilidad* introducido por Morgenstern y Von Neumann en 1944 es concebido como *satisfacción* en el caso de los consumidores y *beneficio* monetario para empresarios. (1955, p. 18) A partir de la Teoría de las Preferencias Reveladas, sin embargo, toda alusión al contenido psicológico desaparece, y la utilidad pasa a ser simplemente un “índice de preferencias” (Hausman, 1992, p. 19)

Los nuevos modelos eran concebidos como *generalizaciones* de EUT, surgidos a partir de debilitar o directamente abandonar al axiomas identificado como el culpable del desajuste con la realidad. Por un tiempo la estrategia de abandonar las virtudes normativas de EUT para recuperar su eficacia descriptiva se presentaba como un camino prometedor. Sin embargo, el surgimiento de un nuevo tipo de contraejemplo, conocido como *preference reversals*, puso sobre la mesa sus límites. En la medida que el nuevo contraejemplo se revele como robusto, la posibilidad misma del proyecto convencional parece enfrentar serios problemas.

2.2-Descripción de EUT

Desde una perspectiva simple, EUT es un modelo que responde a la pregunta por la decisión racional entre acciones (o loterías) que conducen a resultados probables. Para ello parte de un conjunto de cuatro *axiomas* fundamentales.

Axiomas de EUT

Completitud: para todas las loterías X e Y, o bien X es preferida a Y, o bien Y es preferida a X o bien el sujeto se muestra indiferente entre ambas;

Transitividad: para todas las loterías X, Y y Z, si Z es preferida a X y X es preferida a Y, Z será preferida a Y;

Continuidad: para todas las loterías X, Y y Z, tales que X es preferida a Y, e Y a Z, entonces existe una probabilidad p tal que Y es indiferente respecto a una lotería con probabilidad p de obtener X y probabilidad (1-p) de obtener Z.

Completitud, Transitividad y Continuidad constituyen un núcleo duro que garantiza la representación de las preferencias en una *función de utilidad esperada*. Esto quiere decir sencillamente que, si las preferencias de un individuo satisfacen tales axiomas, es posible construir un orden bien formado (*well behaved order*) de loterías.

El último¹⁵ axioma es periférico:

Independencia: para todas las loterías X, Y y Z, X se prefiere a Y si y sólo si la preferencia no se altera ante cambios de igual magnitud en las probabilidades y ante otras loterías.

El axioma de independencia, también llamado *Axioma de Von Neumann y Morgenstern*¹⁶, es específico de EUT y establece que la elección de una lotería sobre otra es *independiente* de la presencia de otras loterías, o de una modificación idéntica en las probabilidades de los resultados. En términos analíticos:

$$X > Y \text{ sii } aX + b > aY + b$$

Si las preferencias de una agente respetan estos axiomas –y vale el llamado *Postulado de Reducción*, (Schoemaker, 1982, p. 532)- entonces pueden ser representadas por una función de utilidad que posee la propiedad de la *utilidad esperada*. Tal propiedad establece que la utilidad esperada de una lotería L es igual a la utilidad de sus resultados multiplicada por la probabilidad asociada a cada uno de ellos¹⁷. Es decir:

$$UE(L) = \sum U(R_i) \cdot P_i$$

Morgenstern y Von Neumann *no inventan* la hipótesis de utilidad esperada, sino que la retoman de la tradición inaugurada por Daniel Bernoulli (ver Cap. I). Hasta entonces la hipótesis era utilizada de manera circunstancial y esporádica en la literatura económica.

¹⁵ Otras presentaciones agregan un quinto axioma, conocido como *Postulado de Reducción* (ver Hausman, 1992, p.23)

¹⁶ Paradójicamente, el axioma de independencia está ausente de la formulación original de Morgenstern y Von Neumann. Ver Fishburn y Wakker (1995) y Mongin (1997).

¹⁷ (Hausman, 1992, p.24).

La *forma* de la función de utilidad esperada suele ser interpretada conductualmente como dando información acerca de las *actitudes frente al riesgo* de los individuos. Una curva cóncava¹⁸ implica *aversión al riesgo*, mientras que una figura convexa está asociada con una conducta *favorable al riesgo*.

Es importante señalar uno de los méritos más relevantes del descubrimiento de Morgenstern y Von Neumann. ¿Por qué es *racional* el criterio de maximizar la utilidad esperada? Hasta entonces, tal pregunta carecía de respuesta formal. Al probar que la hipótesis de la utilidad esperada es *equivalente* con un conjunto de axiomas¹⁹ normativamente atractivos (Friedman y Savage, 1952; Schoemaker, 1982, Hausman, 1992), los autores dieron inicio a una suerte de *Golden Dream* en el campo de la economía: la inesperada (y ampliamente bienvenida) posibilidad de contar con una Teoría de la Decisión *Racional*, es decir una teoría de la decisión *descriptivamente* adecuada y *normativamente* defendible.

2.3-El surgimiento del *Golden Dream*

En rigor, la ventaja de la derivación de Morgenstern y Von Neumann es doble: los axiomas son razonables, por una parte, y susceptibles de comprobación empírica casual, por la otra (Friedman y Savage, 1952). La posibilidad de defender EUT como una teoría normativamente correcta y al mismo tiempo descriptivamente adecuada explica tanto el fuerte impacto ejercido por su aparición, como la amplitud de alcance a la que dio lugar su desarrollo:

“It is no exaggeration to consider expected utility theory the major paradigm in decision making since the Second World War. It has been used prescriptively in management science (...), predictively in finance and economics, descriptively by

¹⁸ La concavidad de la función en condiciones de certeza modela la utilidad marginal decreciente del dinero, algo que no debería ser confundido con la aversión al riesgo.

¹⁹ En rigor, la axiomatización en sentido estricto fue acometida por Marshak (1950).

psychologists, and has played a central role in the theories of measurable utility”
(Schoemaker, 1982, p. 529)

Tanto Savage (1954), como Marschack (1950) y Strotz (1953), señalan explícitamente la importancia normativa de la teoría, como una faceta adicional a su adecuación descriptiva. Existía un amplio consenso en relación a que EUT no era únicamente una teoría descriptiva acerca de cómo la gente decide de hecho, sino también una teoría normativa capaz de instruir acerca de cómo *deben* tomarse las decisiones.

“Consider any person not deemed insane who holds contradictory preferences.... Imagine that we explain to this person the nature of the contradiction, pointing out clearly how his preferences violate our axioms. Will he in consequence of understanding the nature of the contradiction decide that his preferences are ill-founded and proceed to change them, or will he persist in his original preferences even though it is entirely clear to him exactly what precepts his preferences violate. If for nearly every person holding contradictory preferences an understanding of the character of the contradiction induces him to straighten out his preferences, then the Neumann-Morgenstern axioms may properly be regarded as precepts of rational choice. My own feeling is that it would be a strange man indeed who would persist in violating these precepts once he understood clearly in what way he was violating them” (Strotz, 1953, p. 393).

Lo que es más importante, ambas propiedades –el carácter tanto *normativo* como *descriptivo* de EUT- eran consideradas no sólo compatibles, sino además íntimamente relacionadas una con la otra. Las esperanzas que Savage, Marschack y Strotz depositaban en su adecuación empírica estaban basadas probablemente en lo que se creía era su valor normativo: por lo regular, los individuos normales serían capaces de aprender de sus errores. Como Strotz (1953, p. 393) afirma, *“my own feeling is that it would be a strange man indeed who would persist in violating these precepts once he understood clearly in what way he was violating them”*. En la medida que EUT inyecte racionalidad en la conducta de los

individuos, los eventuales desvios de las elecciones correctas podrían ser gradualmente eliminados. Aun cuando EUT fuera inicialmente inadecuada desde un punto de vista descriptivo, su influencia sobre las decisiones de los agentes resultaría en un progresivo ajuste entre la conducta anticipada y la observada. La adecuación normativa garantizaría su autovalidación a nivel empírico.

2.4-Impacto de EUT

Resulta difícil sobrestimar el impacto causado por EUT sobre la profesión económica, por las siguientes razones:

1) resolvía el espinoso problema de la *medibilidad* de la utilidad²⁰, algo que había resultado imposible hasta la fecha. Al respecto, cabe destacar que la originalidad de idear un método aceptable para “medir” preferencias y el hecho de que ello se conformaba a la concepción de ciencia por ese entonces imperante, fue considerado un logro significativo para la disciplina. Bajo la influencia del empirismo, se consideraba que la ciencia fáctica debía estar conformada por enunciados sintéticos, con contenido empírico, y sus definiciones debían ser operacionalmente significativas. Los axiomas de Neumann y Morgenstern lograban ambas cosas²¹.

2) en el pasado, la teoría de la utilidad había sido aplicada con éxito en economía a situaciones de certeza (consumo), pero se la consideraba inaplicable a situaciones de riesgo o incertidumbre. La teoría axiomática de Neumann y Morgenstern de la utilidad esperada, así como otras versiones posteriores, propuestas bajo su influencia, como la de Marschack (1950), parecían cambiar esta situación y abrir la posibilidad de unificar la explicación del campo de los fenómenos económicos. Se esperaba que, adecuadamente desarrollada, fuera posible explicar cierto tipo de decisiones que involucran riesgo y que anteriormente quedaban fuera del alcance de la teoría y eran atribuidas a la

²⁰ En el *Apéndice* se mostrará un procedimiento sencillo para medir la utilidad.

²¹ Véase Friedman y Savage (1948 y 1952).

arbitrariedad o la preeminencia de factores psicológicos. El artículo de Friedman²² y Savage (1948) inicia este “programa” de expansión de las ideas contenidas en EUT.

3) Los axiomas de EUT podían ser interpretadas de manera normativa y descriptiva. En primer lugar, podía sostenerse que EUT especificaba la manera en que elegiría entre loterías un decisor racional ideal (dotado de *well behaved preferences*). En segundo lugar, podía plantearse la hipótesis empírica de que aplicada al campo de las decisiones humanas, la teoría también resultaba compatible con el comportamiento registrado en situaciones de riesgo. Las actitudes entre los economistas acerca de cuál interpretación adoptar no eran unánimes. Algunos parecen dar preeminencia a la interpretación normativa (Raiffa, 1961), en tanto que otros se inclinan por la interpretación descriptiva (Friedman y Savage, 1948, 1952). Otros autores señalan la importancia normativa y descriptiva de EUT (Strotz, 1953).

4) Pero la virtud sobresaliente de la teoría que es relevante para este trabajo se vincula a su interpretación descriptiva. Asumiendo que los individuos se comportan consistentemente en el sentido definido en la teoría, es posible “deducir” (y *predecir*) las elecciones que los agentes realizarán entre loterías alternativas, una vez que algunas de sus elecciones son conocidas. Esta también era considerada una propiedad *necesaria* de las teorías científicas, aunque más no fuere por su vinculación con el testeo.

²² “The hypothesis, if reasonably valid in a sufficiently wide domain, has far-reaching implications for economic theory. It provides a unified interpretation of two kinds of economic behavior that have traditionally been rationalized on divergent, and largely inconsistent, lines – first, choices among alternatives regarded as certain, rationalized in terms of consistent preferences for the goods in question and deliberative selection of the alternative highest in the scale of preference; second, choices among alternatives involving risk, rationalized in terms of ill-defined preferences for “risk” or “uncertainty”, generally regarded as “irrational” in the sense that they admit of no simple and reasonably universal description but can only be determined *ad hoc* in each individual case. Furthermore, the expected utility hypothesis has potentially rich empirical content” (Friedman y Savage, 1952, pág. 463)

Este es un punto que merece ser destacado. La hipótesis de la utilidad esperada *posee* contenido empírico: no toda acción es permitida como racional en el entramado de EUT.

2.5-Funciones de EUT

Se ha empleado la Teoría de la Utilidad Esperada para resolver cinco²³ clases de problemas diferentes. Esto muestra la flexibilidad y fecundidad de la teoría, pero también hace necesario especificar qué uso específico afirman los críticos que la teoría falla en resolver, algo que muchas veces es pasado por alto. La evaluación de la aceptabilidad de cualquier teoría no puede hacerse de manera independiente a sus objetivos, salvo por supuesto que se trate de la evaluación lógica, porque los errores formales pueden detectarse por medios puramente lógicos. Pero no hay ninguna manera semejante de responder *a priori* si una teoría cumple o no con determinados objetivos no formales; de ahí la importancia de centrar la crítica en alguno de los siguientes usos.

1-Se ha utilizado a EUT como una teoría normativa o prescriptiva²⁴ de la decisión racional. Cuando es así, suele asumirse que los agentes toman decisiones sub óptimas, y EUT proporciona el *benchmark* o norma que establece el resultado óptimo, por lo que se pueden prescribir conductas con la mirada puesta allí. Para la visión prescriptiva, el hecho de que los agentes *puedan* seguir los axiomas (es decir, posean las capacidades computacionales y de acceso a la información necesarias para realizar los cálculos de EUT) constituye evidencia empírica relevante. Nada de esto aplica al uso predictivo e interpretativo de EUT.

²³ Una gran parte de la literatura pasa por alto la distinción entre las funciones de EUT, lo cual constituye un problema en términos metodológicos. Schoemaker (1982) menciona 4 usos, todos vinculados con la decisión individual. El quinto uso subyace a la práctica de los economistas, tal como es presentada en Friedman (1955).

²⁴ La dicotomía habitual en la literatura filosófica tradicional es entre juicios normativos y juicios descriptivos. En teoría de la decisión sin embargo, existe una tricotomía: hay enunciados normativos (o teorías normativas), descriptivos y prescriptivos. Los enunciados prescriptivos son aquellos que recomiendan conductas. EUT podría ser una teoría normativa pero no prescribir nada, algo que podría resultar paradójico para la mentalidad filosófica.

2-Se ha utilizado a EUT como teoría *descriptiva* o *predictiva*, para conocer de antemano cuál va a ser la elección de los agentes en situaciones de riesgo (Fredman y Savage, 1954; Machlup, 1967). El uso descriptivo no se compromete con la existencia del mecanismo de elección propuesto por la teoría. En otros términos, el mecanismo de elección podría ser radicalmente diferente al supuesto en EUT, y sin embargo la teoría podría predecir de manera suficientemente adecuada dentro de un dominio relevante. Para este uso la violación flagrante de los axiomas es *irrelevante*, porque siempre se puede asumir que los agentes se comportan *como si* hicieran los cálculos previstos en EUT.

Estas dos formas de interpretar a EUT pueden sugerir una distinción entre lo descriptivo y lo normativo desorientadora respecto a la forma en la que muchos economistas entienden el tema. En la tradición de Hume se suele distinguir tajantemente entre ambos planos, el normativo y el descriptivo, acentuando la ausencia de conexión lógica entre enunciados de hecho y enunciados de valor. Pero que no exista conexión lógica no implica que no exista conexión alguna. El denominado *Golden Dream* establece un vínculo preciso y con derivaciones empíricas, como veremos. (Marqués y Weisman, 2009)

Existen otros fines de la teoría, menos habituales. En particular:

3-Se ha utilizado EUT como una teoría *procedural*, para modelar el *proceso* de decisión que subyace a la elección en condiciones de riesgo (Simon, 1955; Bettam, 1979). A diferencia del uso descriptivo, hay aquí un compromiso con la existencia del mecanismo propuesto por EUT. Quienes lo entienden así reclaman que la teoría no sólo sea compatible con las elecciones observables de (al menos la mayoría de) los agentes, sino que el *proceso* por el cual los agentes arriban a su decisión sea también adecuado desde el punto de vista descriptivo. La evidencia empírica buscada es por lo tanto de naturaleza psicológica: cómo se almacena y codifica la información, cómo se la recupera de la memoria, como se la evalúa, etc. *Bounded Rationality* surge como corriente a partir de interpretar a

EUT como una teoría fallida desde el punto de vista descriptivo de los procesos humanos (Marqués y Weisman, 2008; Weisman, 2009)

4-Se ha utilizado a EUT como teoría *interpretativa*, para darle sentido a la conducta de los agentes (Becker, 1976) . En este uso EUT funciona como teoría infalsable, y afirma que cualquier conducta observable es maximizadora de la utilidad esperada *si se la modela de manera adecuada*. Los aparentes contraejemplos son salvados introduciendo una serie de modificaciones (costos, restricciones, acceso limitado a la información, etc.)

5-Un quinto uso, independiente de los anteriores, consiste en emplear EUT como teoría *idealizante*. En este caso, se la asume como una *descripción* de un estado *ideal* particular. De acuerdo al uso idealizante, si la decisión de los agentes es óptima y prevalecen condiciones de competencia, entonces los agentes que sobrevivan al final del día de mercado serán aquellos que implícitamente se hayan comportado *como si* hubieran tomado sus decisiones siguiendo los dictámenes de EUT. Esta es una característica particular de la economía, que comparte quizás con otras disciplinas como la biología: la presencia de *modelos de estado final* (Weisman, 2010)

En el uso idealizante, ni las violaciones de los axiomas ni las violaciones descriptivas de EUT cuentan como anomalías o contraejemplos. Esto es así porque el uso del modelo refiere a un contexto competitivo que asegura *a priori* la eliminación de los casos individuales anómalos. La conexión empírica no puede venir del lado de las predicciones. Sin embargo, el modelo no es completamente vacío: antes bien, posee éxito empírico garantizado. Hay garantías lógicas que aseguran que, en caso de que logren exportarse las condiciones presupuestas por el modelo a la realidad, se verificarán allí dichos resultados. De manera más débil, suele afirmarse que a medida que las condiciones de la realidad se vayan aproximando asintóticamente a las planteadas por el modelo, los resultados serán progresivamente semejantes a los que se derivan de él. Bajo esta interpretación, un modelo no intenta describir la

conducta de todos los individuos, ni siquiera la conducta de individuos reales: sino de aquellos que resultan ser “exitosos”. El conocido “ejemplo del billarista”²⁵ de Friedman ilustra este caso.

2.6-Evaluación empírica de EUT: La *Paradoja de Allais*

La virtud sobresaliente de la teoría que es relevante para este trabajo se vincula a su interpretación descriptiva. EUT hace posible “deducir” (y *predecir*) las elecciones de los agentes ante loterías alternativas, una vez que algunas de sus elecciones son conocidas. La posibilidad de contar con una teoría interpretable empíricamente y de la cual surgen predicciones no es común en economía. Históricamente, la posibilidad ha generado un vigoroso movimiento de interés experimental, y como resultado las contrastaciones empíricas no se hicieron esperar. Y con ellas comenzaron a surgir las primeras dificultades, cuando el comportamiento observado en ciertas situaciones bien definidas resultó ser incompatible con el que EUT había anticipado. Consideremos el siguiente ejemplo:

*Experimento de Allais*²⁶

“Se extrae una bolilla de un recipiente que contiene 100 bolillas, de las cuales 1 bolilla es roja, 89 blancas y 10 son azules. Dependiendo del color de la bolilla extraída, y de la elección entre A y B, y luego C y D, se reciben los siguientes premios²⁷:”

²⁵ Friedman (1953)

²⁶ Allais (1953). La intención original de Allais fue, sin embargo, criticar EUT como teoría normativa de la decisión, mostrando que existían elecciones que violaban EUT y sin embargo eran defendibles racionalmente. Ver Hausman, *loc. cit.*

²⁷ Extraído de Hausman, 1992, p. 213.

Problema	Elección	Premios		
		Roja (1)	Blanca (89)	Azul (10)
I	A	\$ 1.000	\$ 1.000	\$ 1.000
	B	\$ 0	\$ 1.000	\$ 5.000
II	C	\$ 1.000	\$ 0	\$ 1.000
	D	\$ 0	\$ 0	\$ 5.000

El juego está dividido en dos pasos. En el primero es menester seleccionar entre la opción A y la B. Una vez hecho esto, se pasa al segundo problema y allí se decide entre las opciones C y D.

Cuando se reproduce este juego, se observa que mayoría de la gente elije la opción A del problema I, y a continuación seleccionan la opción D del problema II. Este patrón viola la predicción que emerge de EUT. La teoría da plena libertad para elegir entre A y B en el primer caso: quien elija A (B) revela que (no) prefiere obtener \$ 1000 con certeza a entrar a un juego donde tiene un 10% de probabilidades de ganar \$ 5000 y 89% de probabilidades de ganar \$ 1000. Elegir la opción A sobre la B implica que:

$$U(A) > U(B)$$

$$U(\$ 1000) > U(\$1000) \cdot 0.89 + U(\$ 5000) \cdot 0.1, \text{ de donde sale que}$$

$$U(\$ 1000) - U(\$1000) \cdot 0.89 > U(\$ 5000) \cdot 0.1,$$

$$\mathbf{I: U(\$ 1000) \cdot 0.11 > U(\$ 5000) \cdot 0.1}$$

Pero una vez tomada la decisión en el Problema I, la elección racional en el Problema II está determinada. Para ser consistentes, quienes se deciden por A en el primer paso están obligados a elegir C en el segundo. Esto es así porque la elección de D revela que:

$$U(D) > U(C)$$

$$U(\$ 5000) \cdot 0.1 > U(\$ 1000) \cdot 0.01 + U(\$ 1000) \cdot 0.1, \text{ es decir}$$

$$\text{II: } U(\$ 5000) \cdot 0.1 > U(\$ 1000) \cdot 0.11$$

Lo cual contradice el resultado anterior.

La misma violación de EUT se verifica empíricamente en el siguiente experimento:

<i>Elija, en el problema I entre A y B, y en el problema II entre C y D.</i>			
<i>I</i>	<i>A: (\$4000, 0.8)</i>		<i>B: (\$3000)</i>
<i>II</i>	<i>C: (\$4000, 0.2)</i>		<i>D: (\$3000, 0.25)</i>

En la primer elección los sujetos deben elegir entre obtener \$4000 con una probabilidad de 0.8 (o, en caso contrario, nada) u obtener \$3000 con certeza. Las dos opciones siguientes deben entenderse de la misma manera. Nótese que II se obtiene a partir de multiplicar las dos opciones del problema I por 0.25. El *Axioma de Independencia* asegura en tales casos que las preferencias se mantendrán. De acuerdo con la teoría, si un individuo racional elige A por sobre B, debe elegir C en vez de D. La primera elección es completamente libre, pero una vez llevada a cabo el requisito de consistencia exige seguir un patrón específico en la segunda elección.

Tal patrón específico no es, sin embargo, el que se ha observado cada vez que se realiza el experimento. El fracaso evidenciado en éste y otros²⁸ diseños experimentales parecen indicar que las limitaciones *predictivas* de la teoría son insoslayables.

²⁸ Ver Starmer (2000), pág. 337 y ss

2.7-Dos clases de contraejemplos: anomalías y paradojas

No todos los contraejemplos empíricos son iguales. De acuerdo con Marqués (2008) resulta conveniente distinguir anomalías de paradojas. Las *anomalías* son contraejemplos invocados para señalar fallas predictivas de una teoría. En el terreno de la teoría de la decisión racional, si se predice como *racional* la elección de la canasta A sobre la B y los individuos seleccionan en su lugar la canasta B sobre la A, estamos en presencia de una *anomalía*. En este caso, se asume que los individuos eligen de manera irracional, aceptando de manera implícita la autoridad de la teoría para señalar los cursos de acción racionales. Una *paradoja*, por su parte, es un contraejemplo en el cual los individuos se desvían de la norma, pero además están dispuestos a afirmar –y defender– que su elección es *racional*. A diferencia del caso anterior se cuestiona la teoría, indicando la existencia de conductas racionales no contempladas u opuestas a las señaladas por ella. Es lo que sucede cuando además de elegir la canasta B sobre la A, se defiende la decisión como racional, a pesar de ser contraria a la recomendada por la teoría principal²⁹. Evidentemente existen dos nociones de racionalidad distintas en pugna aquí, lo cual se revela en la disputa por adjudicarle o no el adjetivo racional a cierto patrón de elecciones. La tensión surge a la hora de aplicar EUT para recomendar cursos de acción supuestamente racionales a individuos que enfrentan situaciones de riesgo.

¿En qué consiste la relevancia de los experimentos para evaluar el valor *normativo* de EUT? Podría ocurrir que no todos los individuos lograran seguir la instrucción que proporciona EUT o no la aceptaran (implícitamente) al decidir. Podría ser que en un porcentaje de casos (que puede ser alto) los agentes elijan

²⁹ Incidentalmente, esto es lo que ocurre en el caso de Allais (1953), quien pensó el caso como un experimento mental destinado a señalar las falencias del uso *normativo* de EUT. No es un mero contraejemplo empírico. Allais quiere sostener que el patrón de elección anómalo (A sobre B y luego D sobre C) es defendible *racionalmente*. Esto explicaría el carácter de *paradoja* que posee el contraejemplo.

sistemáticamente de manera inconsistente con EUT. Sin embargo, la mera incompatibilidad empírica con las decisiones de algunos agentes, no descalifica *ipso facto* a la teoría (salvo en el caso en que todos los agentes fallen, lo cual sugiere que la instrucción provista por algunos de sus axiomas podría no ser factible). Las llamadas “paradojas” (como las de Allais) son relevantes para la evaluación empírica de EUT porque en estos casos las elecciones que realizan los sujetos no parecen ser irracionales.

La presencia de una paradoja no indica que los sujetos no pueden ajustarse a los axiomas de EUT (por sufrir limitaciones de registro, memoria o computación), sino más bien que aunque dispusieran de las capacidades sobre-humanas necesarias y fueran capaces de comportarse como los axiomas instruyen, algunos de los resultados alcanzados de esta manera pueden ser inaceptables de todos modos. La relevancia de las *paradojas* no es el hecho de que los sujetos se desvíen de lo prescrito por EUT, sino que el comportamiento que manifiestan puede ser (razonablemente) interpretado como preferible racionalmente al aconsejado por EUT.

Esto da la pauta de que hay dos nociones de racionalidad diferentes en juego. La racionalidad definida por EUT, que autoriza a calificar de manera inmediata como irracional a cualquier desviación, y una idea *intuitiva* de racionalidad, que autoriza a defender a algunos comportamientos desviados. Siempre que se presenta una *paradoja* se admite por implicación dos cosas: a) existe una “teoría” intuitiva de racionalidad que rivaliza con EUT; b) en el caso particular considerado tal teoría intuitiva tiene preeminencia por sobre EUT.

Pero además, la carga normativa asociada al término *racionalidad* ocasiona que se confundan factores empíricos y evaluativos, en mayor medida quizás que en otras teorías. Un ejemplo de este enredo lo constituye la reacción de Baumol frente a la capacidad predictiva de EUT:

“The point is simply that the assumptions of the system, in Samuelson’s happy phrase, ‘put a straitjacket on the person’s preferences’. Once he has made up his

mind on a few things, the rest is decided by him for the rules. From his choices among some limited sets of items we know how he will react to the rest, and there is no apparent reason why choice should be so circumscribed in fact” (Baumol, 1951, p. 64)

La respuesta de Friedman parece a simple vista, plenamente justificada: la posición de Baumol involucra un desconocimiento profundo de la naturaleza de las hipótesis empíricas. En la medida en la que refieren a la realidad, las predicciones colocan una camisa de fuerza sobre los fenómenos, en el sentido de limitar el universo de lo posible, y esto es una virtud más que un defecto. No todo puede pasar, y tal es el sentido empirista (notablemente mencionado por Popper, por ejemplo) del término “ley” como prohibición o restricción de la naturaleza.

“Baumol is saying (...) that the hypothesis is not obviously and inevitably true, that behavior is conceivable that would contradict it, rather than that such behavior has been observed, however crudely. (...) This feature is clearly a virtue of a scientific hypothesis, not a defect -it is a valid objection only to claims that the hypothesis must be true (i.e., is a tautology). The possibility of specifying behavior that would contradict the hypothesis means that the hypothesis is not empty; that such behavior is not clearly "pathological" or "introspectively obvious" means that the hypothesis, if valid, would enable us to make more precise predictions than we can without it.” (Friedman y Savage, 1952, p. 464 y s.)

La situación, sin embargo, podría resultar más compleja. La crítica de Baumol podría apuntar al uso *normativo* de EUT: señalaría una *paradoja*, más que una *anomalía*. Si esto es así, Friedman confunde dos críticas diferentes, por no diferenciar de manera adecuada entre anomalías y paradojas.

2.8-Evaluación metodológica de los contraejemplos

¿Cómo afecta a EUT los problemas formulados por Allais? ¿Comprometen de manera necesaria a la teoría? Un inconveniente fundamental al plantear tal pregunta es que existen al menos cinco problemas diferentes a los cuales responde EUT. Esto torna necesario determinar en cada caso qué uso se está vulnerando con los contraejemplos; de otro modo se corre el riesgo de pensar, equivocadamente, que se ha desterrado la teoría por completo cuando sólo se ha determinado que falla en relación a uno de los problemas que, se creía, podía resolver. Empecemos con los menos problemáticos:

1-Los contraejemplos no afectan al uso *interpretativo*, en el cual existen garantías de que cualquier conducta puede interpretarse como un caso de conducta racional. En el Capítulo VI veremos que este uso está más extendido de lo que pareciera a primera vista. En cualquier caso, siempre es posible salvar la refutación reinterpretando con algún ingenio los términos del contraejemplo. En particular, los defensores del uso interpretativo pueden argumentar que las preferencias de los sujetos del experimento de Allais han *cambiado* en el pasaje del problema I al problema II. De manera menos evidentemente *ad hoc*, pueden sostener que el supuesto de que la utilidad es una función monótona de los resultados monetarios es falso, y por lo tanto la utilidad de obtener 0 en el problema I cuando se elige la opción B no es nula, como implícitamente suponemos, sino un intenso arrepentimiento (Hausman, 1992, p. 214). De ambas formas se mantiene a EUT a resguardo de la evidencia empírica contraria (y de cualquier contenido empírico, por supuesto)

2-El uso *idealizante* tampoco está afectado por contraejemplos puntuales. Independientemente de cuáles sean las decisiones tomadas por los individuos reales, el mecanismo de mercado garantiza que al final sobrevivan o prosperen aquellos que se han comportado *como si* hubieran tomado sus decisiones a la manera propuesta por EUT. Tomando una analogía con la biología evolucionista, la existencia de un individuo pésimamente dotado de habilidades y con rasgos

físicos o conductuales marcadamente desmañados no tienen impacto sobre la teoría evolutiva, porque se asume que al final del proceso de lucha darwiniana nuestro ejemplar o bien será eliminado a favor de individuos mejor ajustados o bien existirá un nicho donde sus características le reporten alguna ventaja competitiva. Sea como sea, al final del proceso los individuos que sobrevivan mostrarán una adaptación con el entorno, de tal manera que el estado final será aquel donde prime una perfecta racionalidad ecológica (Gigerenzer, 2002)

3-El uso *normativo* parece estar a salvo, de acuerdo a la venerable tradición humeana, de los contraejemplos. Una teoría normativa de la decisión afirma como *deben* comportarse los individuos, no de qué manera se comportan de hecho. La distinción de Hume suele ser aceptable en una amplia variedad de contextos, pero las cosas no son tan fáciles cuando se trata de una teoría de la decisión. En particular, hay vínculos intrincados entre lo normativo y lo predictivo cuando la predicación se realiza sobre una teoría de la decisión *racional*.

En particular, hemos visto que las *paradojas*, a diferencia de las anomalías, disputan la adecuación normativa de la teoría. Una situación es elegir de manera contraria a EUT y aceptar que se ha cometido un error, y otra defender la propia decisión como razonable. En el primer caso se admite la preponderancia normativa de la teoría, y el patrón de elección se identifica como erróneo a la luz de ella. En el segundo, se discute que la teoría contenga o encierre en ella todo el comportamiento *racional*: existen elecciones que son racionales, a pesar de que son “irracionales” bajo EUT. En tales casos la discusión se traslada al plano normativo. Los contraejemplos *pueden* afectar a EUT en su uso normativo.

4-Claramente pareciera que es el uso *descriptivo* (predictivo) el que enfrenta problemas. De alguna manera, los contraejemplos hacen necesario modificar la teoría o postular al menos cierta influencia distorsiva para explicar la desviación entre la conducta observada y la predicha. Históricamente esta línea de argumentación ha sido la más aceptada, y dio lugar a una serie de modelos

alternativos a EUT con pretensiones de adecuación descriptiva. Starmer (2000) denomina sugestivamente a este movimiento *The Hunt for a Descriptive Theory of Choice under Risk*.

Sin embargo, las cosas no son tan sencillas tampoco en este caso. Los contraejemplos no tienen por qué afectar necesariamente el uso descriptivo o predictivo de EUT, al menos para quienes están dispuestos a aceptar la adecuación normativa de la teoría.

En efecto, una teoría *normativa* puede dar lugar a una función descriptiva exitosa en la medida que se autovalide. Los individuos irían modificando su conducta a medida que incorporaran las normas de comportamiento provistas por EUT. Esta es, después de todo, una promesa implícita en el *Golden Dream*. De tal manera la brecha entre la conducta observada y la predicha se iría cerrando progresivamente. Es lo que Hausman (1992) llama “función educadora” de EUT:

“(...) one might reasonably hold that, when people behave irrationally, the theorist’s response should not be to revise utility theory, but to encourage agents to change their behavior. Those who are unclear on what rationality requires or who are lazy or ineffective in their efforts to conform need reeducation” (Hausman, *op,cit*, pág.219)

5- El uso *procedural* también puede enfrentar problemas. Las fallas empíricas siembran dudas entre quienes interpretan la teoría como la descripción del *proceso* realmente llevado a cabo por los individuos para tomar elecciones. Sin embargo, dadas las dificultades intrínseca de rastrear un proceso cuya naturaleza es la de ser un estado interno, estas dudas no tuvieron fruto visible³⁰.

En la sección siguiente veremos la manera en la que afrontan el problema quienes consideran que los contraejemplos han vulnerado el uso *descriptivo* de EUT, y reaccionan en consecuencia.

³⁰Si bien Simon desconfía del proceso postulado por EUT, lo hace en virtud de otras razones, muy lejanas de los contraejemplos empíricos revistados en este capítulo.

2.9-Critica de EUT como teoría *descriptiva* de la decisión. El problema de la flecha del *Modus Tollens*

Hemos visto que EUT posee un conjunto de cuatro axiomas. En rigor, una presentación más elaborada reconocería que la teoría no sólo tiene axiomas postulados explícitamente, sino que también involucra *supuestos*. Una descripción comprehensiva del contenido de EUT puede encontrarse en Starmer (2000, p. 2 y 3). Esquemáticamente los elementos son:

Axiomas y Supuestos de EUT

Axiomas	Supuestos
•Transitividad	• Invariancia Procedural
•Completitud	• Invariancia Descriptiva
•Continuidad	
•Independencia	

Quienes interpretan los contraejemplos experimentales como fallas descriptivas de EUT se enfrentan a un viejo problema metodológico. Al menos uno de estos elementos debe ser modificado para restaurar la conexión con la experiencia. ¿Hacia dónde dirigir la flecha del Modus Tollens? ¿Qué parte o partes de la estructura teórica deben cargar con el peso de los contraejemplos?

Los supuestos de Invariancia Procedural y Descriptiva yacen en un nivel profundo de la estructura teórica. El primero afirma que la estructura de preferencias permanece invariante ante procedimientos equivalentes utilizados para sacarlas a la luz. La Invariancia Descriptiva, también llamada *Extensionalidad*, implica que descripciones equivalentes de la misma situación riesgosa no afectan las preferencias. Ambos supuestos parecen intachables.

Entre los axiomas, los de transitividad, completitud y continuidad definen preferencias bien formadas (*well behaved preferences*) Constituyen el núcleo central de la estrategia convencional, al que es difícil renunciar por cuanto garantiza la presencia de una función de utilidad esperada.

En la periferia se ubicaba el axioma de independencia, al que por otra parte siempre se vio con suspicacia. Su posición marginal en la red teórica lo ubicaba como el blanco natural del *modus tollens*:

“The Independence axiom of EUT places quite strong restrictions on the precise form of preferences: it is this axiom which gives the standard theory most of its empirical content (and it is the axiom which most alternatives of EUT will relax)” (Starmer, 2000, p. 334)

Se pensó que era posible construir una teoría más abarcativa que prescindiera o debilitara el axioma de independencia y sea por lo tanto compatible con la evidencia empírica disponible. Esto dio lugar a una plétora de modelos en los años '70 guiados por el ambicioso objetivo de conseguir una explicación unificada de las decisiones humanas en condiciones de riesgo. Entre ellas, teorías como la de Machina (1982), Chew y Mc Crimmon (1979) y Fishburn (1983), que conservan un núcleo básico compartido con EUT –la idea avanzar conservando las preferencias bien formadas y debilitando o directamente dejando caer el axioma de independencia- forman parte de lo que se ha llamado “estrategia convencional” (Starmer, *Op. cit.*)

Sin embargo, existe una clase de contraejemplos de EUT, de naturaleza distinta a los ya presentados, que constituye una amenaza de peso para las posibilidades de la estrategia convencional.

2.10-Descubrimiento de *preference reversals*

Supongamos que un individuo prefiere la lotería A a la B. Tal preferencia puede expresarse al menos de dos maneras:

- 1-Seleccionando la lotería A cuando se le ofrecen A y B al mismo tiempo, o
- 2-Exigiendo un precio más alto por desprenderse de A en caso de que ya posea ambas

1 y 2 son procedimientos equivalentes para sacar a la luz la preferencia del individuo. El supuesto de Invariancia Procedural asegura que tal preferencias es independiente del procedimiento utilizado para elucidarla.

En 1971 dos psicólogos, Paul Slovic y Sarah Lichtenstein descubrieron que los seres humanos violan este requisito. Enfrentados a dos loterías, elijen una de ellas, pero cuando se les pregunta a cuánto venderían cada una le asignan el precio más alto a la *otra*. Lo sorprendente de este fenómeno es que *elegir* y *valorar* suelen asumirse como formas alternativas pero equivalentes de sacar a la luz las preferencias de los agentes. En palabras de Starmer:

“It presents a puzzle for economics because, viewed from the standard theoretical perspective, both tasks constitute ways of asking essentially the same question, that is ‘which of these two prospects do you prefer?’ In [preference reversals’ experiments] however, the ordering revealed appears to depend upon the elicitation procedure” (loc.cit., p. 338 y s.)

Los experimentos diseñados por Slovic y Lichtenstein consisten en pares de loterías, una de las cuales otorga una alta probabilidad de ganar un premio modesto, y la otra una probabilidad baja de ganar un premio sustancioso. Por ejemplo, sean las loterías A y B las siguientes³¹:

Lotería A: P-bet*

En una urna con 36 bolillas, de las cuales 1 es roja y las restantes azules, se pierde 1\$ si se obtiene la bolilla roja, y se ganan 4\$ en caso de obtener una azul.

³¹ Extraído de Hausman, 1992, p.228 y ss, con modificaciones.

El valor esperado del juego es de $3,86\$$ ³². Como hay altas probabilidades de ganar (35 sobre 36), Slovic y Lichtenstein llaman a esta lotería “P*-bet”. Compárese con la siguiente opción:

Lotería B: \$-bet

En una urna con 35 bolillas, de las cuales 11 son blancas y 25 verdes, se pierde 1,5 \$ si la bolilla extraída resulta verde, y se ganan 16 \$ si es blanca.

El valor esperado es en este caso de $3,85\$$ ³³. Nótese que los premios en esta ocasión son sensiblemente mayores, y estas loterías son denominadas “\$-bet” por tal razón.

La predicción condicional de Slovic y Lichtenstein es que si los individuos eligen la lotería A sobre la B, estarían dispuestos a *pagar más* por la lotería B que por la A.

En *Reversals of Preference between bids and choices in gambling decisions*, se registran y describen los resultados de los experimentos llevados a cabo. El procedimiento involucra dos fases. En la primera, los sujetos son enfrentados a pares de juegos como el A y B, caracterizados por un valor esperado similar, y deben elegir alguna de las loterías involucradas. El segundo paso está precedido por tareas aleatorias destinadas a distraer la atención de los sujetos. Por fin se les presenta A y B en sucesión y se les solicita ahora que fijen un *precio* para cada uno. Los experimentos arrojan el siguiente resultado: la mayoría de participantes (cerca del 75%) que eligieron la lotería con mayor probabilidad de ganar (B) pusieron un precio más alto a la *otra* lotería (A).

³² Es decir, $35/36 \cdot 4\$ - 1/36 \cdot 1\$$

³³ El resultado surge de la siguiente operación: $11/36 \cdot 16\$ - 25/36 \cdot 1.5\$$

El significado de este fenómeno para la economía es profundo. En palabras de los economistas David Grether y Charles Plott:

“Taken at face value the data are simply inconsistent with preference theory and have broad implications about research priorities within economics. It suggests that no optimization principles of any sort lie behind even the simplest of human choices and that the uniformities in human choice behavior which lie behind market behavior may result from principles which are of a completely different sort from those generally accepted.

Notice this behavior is not simply a violation of some type of expected utility hypothesis. The preference measured one way is the reverse of preference measured another and seemingly theoretically compatible way. If indeed preferences exist and if the principle of optimization is applicable, then an individual should place a higher reservation price on the object he prefers. “
(Grether y Plott, 1979, p. 623, cit. por Hausman, *op. cit.*, p. 230)

Las implicaciones del descubrimiento de Slovic y Lichtenstein explican en parte la actitud cautelosa de los economistas. Quienes tomaron el fenómeno seriamente, como Grether y Plott (1979; 1982), intentaron expulsarlo del campo de la economía mediante una serie de hipótesis auxiliares. Se conjeturó la presencia de *bajos incentivos*, los cuales implican que los individuos experimentales de Slovic y Lichtenstein no tenían razones para comportarse como lo harían en circunstancias reales, lo que descalificaba la representatividad de la conducta observada en laboratorio para contextos económicos más amplios; *efectos ingreso*, de acuerdo a los cuales al cambiar la riqueza de los participantes se modificaría su actitud frente al riesgo, lo que podría explicar la reversión con un mecanismo de la caja de herramientas económica. Por otra parte, entre las preguntas formuladas experimentalmente no figuraba la opción de mostrarse *indiferente* ante las loterías, lo que supone un *pobre diseño experimental*, y abre la puerta para una explicación alternativa de la reversión según la cual las preferencias “anómalas” resultan en realidad formas desviadas de expresar

indiferencia. Y si no es así, todavía pueden explicárselas suponiendo individuos que actúan de manera *estratégica* al poner precios de venta: es una práctica habitual pedir un precio mayor al de reserva, con fines puramente estratégicos y de negociación. Algo así podría estar pasando en los experimentos de 1971.

En 1979, Grether y Plott llevaron a cabo una nueva serie de experimentos diseñados para explicar la reversión de preferencias mediante la incorporación y testeo de las hipótesis auxiliares mencionadas. Se incorporaron premios monetarios y se fueron variando los incentivos; luego se controló la variación de ingresos, diferenciando entre grupos, pero nada de esto logró hacer desaparecer el fenómeno. También se modificó el diseño experimental, el cual permitía ahora la expresión de *indiferencia* entre las loterías, y se controló la puesta estratégica de precios, y la reversión anómala de preferencias se mantuvo constante. La conclusión que extraen Grether y Plott es clara:

“Needless to say, the results we obtained were not those expected when we initiated this study. Our design controlled for all the economic-theoretic explanations of the phenomenon which we could find. The preference reversal phenomenon which is inconsistent with the traditional statement of preference theory remains...” (Grether y Plott, 1979, p.634, cit. en Hausman, 1992, p. 233)

Además de falla en la invariancia *procedural*, existe evidencia relativa a violaciones de la invariancia *descriptiva*. Cambios menores en la *descripción* de una situación dada provocan modificaciones de peso en las decisiones de los agentes. Incluso se puede dar el caso de preferencias revertidas ante descripciones *equivalentes* de una misma situación de elección. El caso más conocido es sin duda el de la *Enfermedad Asiática*, propuesto por Tversky y Kahneman en 1981.

The Asian Disease

“Imagine that the United States is preparing for the outbreak of an unusual Asian disease, which is expected to kill 600 people. Two alternative programs to combat the disease have been proposed. Assume that the exact scientific estimates of the consequences of the programs are as follows:”

La versión original fue realizada sobre estudiantes de la universidad de Stanford y de Columbia, y es replicable en nuestras universidades. Los sujetos experimentales suelen ser divididos en dos grupos, al primero de los cuales se les pide que seleccionen entre los dos programas A y B siguientes:

A: If Program A is adopted, 200 people will be saved

B: If Program B is adopted, there is a one-third probability that 600 people will be saved and a two-thirds probability that no people will be saved

El segundo grupo debe elegir entre A' y B':

A': If Program A' is adopted, 400 people will die

B': If Program B' is adopted, there is a one third probability that nobody will die and a two-thirds probability that 600 people will die

Es claro que los dos pares de opciones son *idénticos* en términos numéricos. La única diferencia reside en la expresión de los resultados en términos de “vidas salvadas” en el primer grupo y “vidas perdidas” en el segundo; por supuesto, salvar 200 vidas de un grupo de 600 personas (opción A) equivale a perder 400 vidas (opción A'). Si las decisiones de los agentes reales respetaran la *invariancia descriptiva* se observaría que el porcentaje a favor de la opción A (B), sería semejante al de la opción A' (B'). Pero esto no sucede. En el experimento original, cerca del 72 por ciento de la muestra (N: 155) se inclinó por la opción A sobre la B, mientras sólo un 22 por ciento prefirieron la opción A' sobre la B'.

10-Impacto de *Preference Reversals* sobre el *Golden Dream*

No es difícil ver por qué los economistas confiaban en que se resolvería el fenómeno de *Preference Reversals*. EUT trajo consigo una triple promesa: la de ser una teoría capaz de brindar un tratamiento *unificado* de las elecciones bajo riesgo, y de ser adecuada, tanto *descriptiva* como *normativamente*. Este esperado triple desiderátum puede considerarse como el *Golden Dream* de la economía, la posibilidad de contar con una Teoría de la Decisión *Racional*.

“Because framing effects and the associated failures of invariance are ubiquitous, no adequate descriptive theory can ignore these phenomena. On the other hand, because invariance (or extensionality) is normatively indispensable, no adequate prescriptive theory should permit its violation. Consequently, the dream of constructing a theory that is acceptable both descriptively and normatively appears unrealizable” (Tversky y Kahneman, 1986, p. S272).

La reversión de preferencias supone el final para el *Golden Dream*. Ya no es posible contar con una teoría adecuada desde el punto de vista *normativo* y correcta desde el punto de vista *descriptivo*.

EUT y la estrategia convencional de desarrollo de estas ideas parecían ser la única manera *científica* de abordar las elecciones bajo riesgo. En esta línea, y en consonancia con el operacionalismo reinante, el concepto de utilidad fue despojado de toda connotación psicológica o utilitarista, y su significado reducido al de *índice de preferencias*. (Hausman, 1992) Atar el concepto de utilidad a la escala de preferencias pudo parecer una maniobra simplificadora, pero si la escala de preferencias se revierte ¿qué sucede con el concepto de utilidad convencional?

Entre las alternativas propuestas para lo que se percibía como una limitación de los modelos convencionales, la de Simon (1955) abrió un camino que continúa hasta nuestros días. Simon señala que los modelos estándar suponen individuos con capacidades computacionales ilimitadas, y acceso a la información

instantáneo y sin costos, y postula que gran parte de las limitaciones obedecen a tal punto de partida “irrealista”. En contraposición ofrece un modelo donde agentes con racionalidad limitada (*Bounded Rationality*) deciden en contextos de riesgo o incertidumbre. Para darle sustento a estas afirmaciones, y comprender la naturaleza de las restricciones que EUT impone a la situación de elección, veremos en el Apéndice II las características *generales* de los modelos de decisión racional.

Apéndice I: Procedimiento de Medición de la Utilidad

Supongamos tres opciones alternativas, A, B y C. El procedimiento para medir utilidad puede ser reducido a tres pasos:

1-Tomar la opción más preferida y la menos preferida. Asígnele el numeral 10 a la primera y 0 a la segunda. En nuestro ejemplo, supongamos que $U(A)=10$ y $U(C)=0$

2-Para asignarle un numeral a B, necesitamos responder la siguiente pregunta: ¿Cuál es la probabilidad p^* que lo haría indiferente entre la lotería formada por A con probabilidad p^* y C con probabilidad $(1-p^*)$, por un lado, y por el otro B con certeza? Es decir el individuo debe seleccionar p^* tal que:

$$U(B) \sim p^* U(A) + (1-p^*) \cdot U(C)$$

Supongamos que p^* es 0.4. En tal caso:

$$U(B) \sim 0.4 \cdot 10 + 0.6 \cdot 0.$$

$$U(B) = 4$$

3-Llame a los números asignados “utilidad”. En nuestro ejemplo $U(A) = 10$, $U(B) = 4$ y $U(C) = 0$

Estrictamente hablando, los números obtenidos no expresan solamente preferencias sino además actitudes ante el riesgo. En Alchian (1953) figura una exposición detallada de este método. Algunos comentarios pueden ser útiles aquí. El procedimiento clarifica por qué Friedman nos recuerda repetidas veces que la utilidad es “tan medible como la temperatura y la extensión” (Friedman, 1955). Una vez elegidos los valores extremos de la escala (en nuestro ejemplo, 10 y 0), los puntos intermedios quedan determinados en una forma subjetiva, pero no arbitraria. Como en las escalas celsius y centígrada, los puntos superior e inferior son arbitrarios, pero nada más lo es.

La perspectiva de contar con un concepto de utilidad cardinal fue recibida con entusiasmo en los años '50, si bien, en su exposición original, Von Morgenstern y Von Neumann se sienten obligados a defenderse del cargo de estar mostrando un complejo dispositivo que nadie utiliza en la práctica:

“The objection could be raised that it is no necessary to go into all these intricate details concerning the measurability of utility, since evidently the common individual, whose behavior one wants to describe, does not measure his utilities exactly but rather conducts his economic activities in a sphere of considerable haziness. The same is true, of course, for much of his conduct regarding light, heat, muscular effort, etc. But in order to build a science of physics these phenomena had to be measured. And subsequently the individual has come to use the results of such measurement –directly or indirectly- even in his everyday life. The same may obtain in economics at a future date. Once a fuller understanding of economic behavior has been achieved with the aid of a theory which makes use of this instrument, the life of the individuals might be materially affected.” (1944, p. 18)

Es útil notar que este procedimiento presupone la validez del *Principio de Extensionalidad*. Para ilustrar el punto, supongamos que A significa Shakespeare, B Cervantes y C, Goethe. Siguiendo los pasos 1-3, el orden es

$U(A) > U(B) > U(C)$; es decir $10 > 4 > 0$

Ahora reemplacemos “Cervantes” por “El manco de Lepanto”. Si la reversión de preferencias tiene lugar, es posible que $A > C > B$, por lo que la utilidad asignada a B cambiaría. Esto quiere decir que *bajo dos descripciones extensionalmente equivalentes el mismo objeto B tiene dos utilidades diferentes.*

Apéndice II- Características Generales de los Modelos de Decisión Racional

Los modelos de decisión racional típicamente poseen una estructura formada por los siguientes elementos³⁴:

a- Un conjunto de *estados del mundo* posibles, que son exhaustivos y excluyentes.

El decisor no tiene influencia alguna sobre ellos. Los representaremos como S.

b- Un conjunto de *acciones* posibles, entre las que decide el individuo. Las denotaremos por A

c- Los *resultados futuros* que se obtendrían, y que dependen tanto de las acciones emprendidas como del estado del mundo que finalmente prevalezca.

d- Una función de valuación, la cual representa el valor o *utilidad* atribuida por el decisor a los diferentes resultados posibles.

e- *Información* acerca de la probabilidad de obtener un resultado si se decide seguir un curso de acción cualquiera. Así, $P_a(R)$ significa la probabilidad de obtener el resultado R si la acción *a* es emprendida.

Un ejemplo sumamente sencillo puede ser de utilidad aquí para dotar de contenido intuitivo a los elementos de la estructura recién esbozada. Supongamos un individuo que dirige su mirada hacia lo alto antes de salir de su casa, y percibe un cielo nublado. En tales circunstancias, los estados posibles del mundo (S) son dos: puede llover (S_1) o puede no llover (S_2). Ambos estados son exhaustivos (agotan todas las posibilidades) y excluyentes (la presencia de uno implica la ausencia del otro). En tal contexto de incertidumbre, nuestro individuo enfrenta dos opciones de elección:

A_1 : *Salir con paraguas*

³⁴ Una presentación alternativa puede ser encontrada en Simon (1955), pág. 102. . En particular, la reconstrucción de Simon le asigna probabilidades a *estados del mundo*, en lugar de hacerlo directamente sobre resultados, y distingue entre el conjunto de alternativas de conducta y el subconjunto de alternativas de conducta *percibidas*.

A₂: Salir sin paraguas

Cada acción conduce a una serie de resultados posibles, en función del estado del mundo S_i que finalmente prevalezca. Gráficamente podemos resumir la información en una matriz simple:

	S_1	S_2
A_1	R_{11}	R_{12}
A_2	R_{21}	R_{22}

En la columna 1 aparecen las dos acciones sobre las que tomar la decisión: A_1 es la acción de salir con paraguas, y A_2 prescindir del paraguas. En la fila 1 figuran los estados del mundo: S_1 significa en este contexto “llueve” y S_2 “no llueve”. Las celdas de la matriz contienen los resultados posibles, que son los siguientes:

R_{11} : no mojarse en un día de lluvia

R_{12} : no mojarse, cargar con paraguas inútil

R_{21} : mojarse

R_{22} : no mojarse en un día sin lluvia

El agente debe todavía definir dos elementos adicionales. En primer lugar, necesita establecer las probabilidades de cada S_i y segundo –y más importante– aplicar su *función de valuación* para asignarle un valor a cada resultado. Con respecto al primer punto, supongamos arbitrariamente que las probabilidades asignadas por el servicio meteorológico son 50% para cada escenario. En cuanto a la valuación, el agente asigna los siguientes números:

$U(R_{11})$: 8

$U(R_{12})$: 2

$U(R_{21})$: 1

$U(R_{22}): 10$

Tales números son personales, y varían de un individuo a otro. Pueden ser interpretados como la *utilidad* que cada resultado le representa a nuestro sujeto. El peor resultado posible (R_{21}) posee la utilidad más baja, y el mejor la más alta.

Con estos elementos podemos completar la matriz, y definir *procedimientos* de decisión racional:

	$S_1 (p_1: 0.5)$	$S_2 (p_2: 0.5)$
A_1	8	2
A_2	1	10

Reglas de Decisión Racional

1-Regla Max-min: Asuma que obtendrá el *peor* resultado posible y seleccione el curso de acción A que posea el *máximo* valor para el resultado *mínimo*.

En nuestros términos, la regla max-min conduce a seleccionar como racional la fila A_1 (salir con paraguas), porque es la acción donde el *peor* resultado es *mejor* respecto al resultado inferior de la acción alternativa (posee una utilidad de 2 frente a 1). Siguiendo esta regla el agente minimiza prudentemente las pérdidas posibles, independientemente del estado del mundo que se imponga. El procedimiento no le reserva papel alguno a las probabilidades de los estados del mundo alternativos.

2-Regla de Certeza: Asuma que el *mejor* resultado posible será obtenido, y seleccione aquel curso de acción que provea la *máxima* utilidad. Alternativamente, si tiene acceso solamente a un S específico, elija el curso de acción con mayor utilidad.

La Regla de Certeza tampoco involucra el uso de las probabilidades; su defensa normativa parece tener mejores prospectos en aquellos casos donde el agente no

tiene acceso a todos los estados del mundo relevantes. En el ejemplo, la alternativa racional es A_2 , cuya utilidad en el mejor de los casos es de 10.

3-Regla probabilística: Elija el curso de acción que haga máximo el producto de la utilidad de los resultados y la probabilidad de su ocurrencia; es decir seleccione el curso de acción tal que posea máx: $\sum U(R_i) \cdot P_i$

La Regla probabilística incorpora la información de las probabilidades y las pondera con la utilidad de los resultados. Recoge la vieja intuición de Pascal y la teoría de los juegos de azar en general según la cual una apuesta *racional* computa tanto las ganancias como la *probabilidad* de obtenerlas, junto con la sugerencia de Bernoulli de incorporar la esperanza moral (utilidad) de los resultados. Aplicada la regla a nuestro problema, vemos que la opción A_1 posee una utilidad esperada de 5 ($8 \cdot 0.5 + 2 \cdot 0.5$), mientras que la opción A_2 se caracteriza por una utilidad esperada de 5.5 ($1 \cdot 0.5 + 10 \cdot 0.5$). La decisión racional que debería tomar nuestro hombre aplicando la Regla Probabilística es A_2 , salir sin paraguas. Nótese que este procedimiento es sensible al contexto: si las probabilidades de lluvia aumentan levemente, es fácil ver que la opción con mayor utilidad esperada pasará a ser la A_1 , por lo que el individuo racional saldrá con paraguas.

Es interesante notar los requisitos claramente “irrealistas” impuestos sobre el decisor por este sencillo modelo. Debe conocer de antemano todos los resultados posibles (no hay lugar para mundos no ergódicos o consecuencias no anticipadas), y asignarles probabilidades definidas, además de ser capaz de manipular ambos elementos algebraicamente. En la medida en que la Regla probabilística incorpora la utilidad esperada dentro de su operatoria, constituye una parte de *Expected Utility Theory*. Esto explica las abiertas críticas respecto a su falta de realismo que formulará Simon, y que aun en nuestros días se presenta espontáneamente como un argumento que tiñe de suspicacia las posibilidades teóricas de la visión convencional.

Históricamente, los inconvenientes empíricos de EUT determinaron que el panorama se presente sin duda poco alentador. La reacción de una parte de la disciplina fue considerar que la raíz del problema residía en la *clase de supuestos* utilizados por la doctrina económica neoclásica. Su naturaleza parece demandar la existencia de un *homo economicus*, algo muy alejado de los estándares corrientes. En la vida real, se pensó, nunca existió evidencia de que los agentes se comporten de la manera establecida por EUT; lo cual no es extraño, porque los agentes “reales” no cuentan ni con la capacidad computacional ni con la información relevante para aplicar los cálculos supuestos por la teoría. En palabras de Selten (2001)

“Full rationality requires unlimited cognitive capabilities. Fully rational man is a mythical hero who knows the solutions to all mathematical problems and can immediately perform all computations, regardless of how difficult they are. Human beings are in reality very different. Their cognitive capabilities are quite limited. For this reason alone, the decision-making behavior of human beings cannot conform to the ideal of fully rationality” (p. 14)

El pionero en señalar los defectos de la teoría convencional y de mostrar el camino por el cual debía avanzar la teoría de la decisión –renunciando a las pretensiones hiperbólicas de racionalidad y contentándose con modelar individuos con racionalidad acotada- fue Herbert Simon. A él está dedicado el próximo capítulo.

Capítulo III

***Bounded Rationality.* El modelo de Simon, 1955**

3.1-La crítica del “irrealismo” de los supuestos

Uno de los supuestos de la economía neoclásica consiste en empresarios optimizadores. En términos de la teoría de la decisión, las decisiones productivas de los empresarios están mediadas por un cálculo optimizador –maximización de los beneficios y minimización de los costos-, lo cual requiere modelos con agentes perfectamente racionales, que poseen acceso a la información instantánea, y capacidad computacional ilimitada. Para poner en marcha el cálculo optimizador se necesita una suerte de demonio de Laplace que conozca de manera instantánea todos los estados posibles futuros del mundo, la probabilidad numérica relativa a la actualización de cada uno de ellos, y la capacidad de operar matemáticamente con esa enorme cantidad de información. El hecho de que la llamada Teoría de la Firma descansa en tales supuestos “irrealistas” desde el punto de vista psicológico ha suscitado escepticismo y desconfianza en quienes consideran que las teorías deben poseer puntos de partida razonables, más aproximados a la realidad, y creen necesario en consecuencia imponer límites a la naturaleza de los supuestos, descartando aquellos que resulten “heroicos” o directamente imposibles de cumplir. Una postura metodológica de tal tipo ha originado históricamente la línea de investigación que hizo eclosión alrededor de la década de 1940 con la publicación de varios trabajos en revistas de primer nivel, en los cuales se ponen en cuestión los fundamentos de la visión tradicional.

Los críticos de la visión tradicional³³ sostienen que los empresarios no maximizan beneficios ni minimizan costos al tomar decisiones de producción, sino que se manejan por reglas *had hoc* o siguen heurísticas simples, tales como adicionar un margen determinado (“mark-up”) a los costos medios de producción para tomar decisiones relativas al precio de sus productos. La motivación de estos trabajos descansa en la idea de que la Teoría de la Firma *implica*

³³ Ver “Price Theory and Business Behavior” (R.L. Hall y C. J. Hitch, 1939), y “Shortcomings of Marginal Analysis for Wage-Employments Problems” (Lester, 1946), “Marginal analysis and Empirical Research” (F. Machlup, 1946)

lógicamente la naturaleza optimizadora de los empresarios. Esta afirmación es por lo tanto una consecuencia observacional y además una consecuencia observacional *falsa*, por lo cual la Teoría de la Firma requiere como mínimo ser modificada y quizás hasta ser reemplazada por una teoría alternativa.

Es claro que no existen en la realidad agentes como los descritos por la Teoría de la Firma; nadie espera exactitud descriptiva de tales modelos, pero sí al menos que funcionen como una aproximación imperfecta a la realidad. El debate de los años 40' tiene su origen en la comprobación de que tal función es inexistente. Los empresarios no son agentes "aproximadamente" optimizadores.

En este contexto Herbert Simon publica, en 1955, su influyente trabajo, *A Behavioral Model of Rational Choice* (en adelante, BM), partiendo de la convicción de que el individuo perfectamente racional presentado por los modelos convencionales no solamente no existe, sino que no *puede* existir. Los supuestos convencionales están atravesados por una suerte de incompatibilidad psicológica, y el clima de efervescencia suscitado por las críticas no debe hacer perder de vista que la dirección de su revisión debe encaminarla hacia un mayor realismo psicológico. Lo que necesita la economía, en suma, es partir de individuos modelados con una capacidad computacional y de acceso a la información semejantes a los reales; es decir, individuos con *racionalidad acotada*.

3.2-A Behavioral Model of Rational Choice

El modelo de Simon se ofrece como una alternativa radical a la teoría estándar de la decisión. Haciéndose eco de las críticas a los modelos tradicionales, el punto de partida del autor pasa por la comprobación de que ha sonado la hora del *homo economicus*:

"Recent developments in economics, and particularly in the theory of the business firm, have raised great doubts as to whether this schematized model of economic man provides a suitable foundation on which to erect a theory (...). It

is not the purpose of this paper to discuss these doubts, or to determine whether they are justified. Rather, I shall assume that the concept of "economic man" (...) is in need of fairly drastic revision, and shall put forth some suggestions as to the direction the revision might take" (Op.cit., p. 99)

La necesidad de un cambio fundamental en los análisis de las decisiones de los seres humanos está conectada con las limitaciones percibidas en los modelos convencionales, limitaciones que Simon presenta de una manera compleja. En ocasiones, el autor parece cuestionar la validez empírica "directa" de los axiomas de EUT. Sostiene, por ejemplo, que

"The classical model calls for knowledge of all the alternatives that are open to choice. It calls for complete knowledge of, or ability to compute, the consequences that will follow on each of the alternatives. It calls for certainty in the decision maker's present and future evaluation of these consequences, no matter how diverse and heterogeneous, in terms of some consistent measure of utility" (Simon, 1979, p. 500).

Por lo regular, los individuos no pueden satisfacer los axiomas de la teoría (en particular, los de *completitud y continuidad*), con lo cual no es posible que efectúen un orden débil entre las opciones que enfrentan. Este es un sentido en el cual los axiomas de EUT son considerados "irrealistas".

La dirección hacia la cual debe avanzar la economía es hacia modelos más realistas. Esto, en palabras de Simon, implica dotar a los agentes modelados con características psicológicas:

"Broadly stated, the task is to replace the global rationality of economic man with a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms, including man, in the kinds of environments in which such organisms exist" (loc. cit)

Dado el deficiente saber psicológico de la época, las características atribuibles al agente serán consideradas *realistas* apelando al sentido común.

3.3-Naturaleza del proyecto

El objetivo de Simon es doble: se busca una teoría de la decisión que:

1-no sea pasible de incompatibilidad psicológica; es decir que no acuda a rasgos psicológicos incompatibles con los de los seres humanos “reales”.

2-y que además permita definir e identificar conductas racionales. En términos de Simon, se pretende un modelo que:

“(...) include explicitly some of the properties of the choosing organism as elements in defining what is meant by rational behavior in specific situations and in selecting a rational behavior in terms of such a definition”

Los dos objetivos están interrelacionados. La inclusión de propiedades “realistas” en los agentes de decisión es necesaria para sortear el escollo de la incompatibilidad psicológica. El modelo que se necesita es uno que guarde alguna correspondencia con la realidad, un modelo *factible*. Por otra parte, no puede ser un modelo *vacío*; con su ayuda se deben poder determinar qué acciones son racionales y cuáles no.

Resulta conveniente introducir la siguiente distinción: una decisión racional puede referir al *proceso* por el cual se arriba a una decisión, o al *resultado* de ese proceso. Un modelo puede fallar al describir el proceso, o incluso negarse por razones metodológicas a describirlo (postulando, por ejemplo, una suerte de “caja negra”), y sin embargo describir acertadamente el resultado del proceso “real”. No es necesario contar con un modelo realista del *proceso* para obtener predicciones de los resultados “realistas”, en el sentido de acertadas. Nótese que no distinguir entre el proceso y el resultado del proceso conduce a ambigüedades: si alguien afirma que EUT no es descriptivamente realista ¿se refiere a que sus axiomas no describen el proceso por el cual los individuos

deciden racionalmente? ¿O que el resultado no es uno empíricamente adecuado (algo que se desprender de las paradojas de Allais y Ellsberg, por ejemplo)? La primera es una posición más afín a las intenciones de Simon (1955), quien por otra parte no cuestiona en ningún momento el valor *normativo* de los resultados de EUT. La teoría de la decisión racional identifica exitosamente, de entre un conjunto de loterías, cuál es aquella que resulta razonable elegir. Pero el *proceso* por el cual se llega a esa decisión no es el que siguen los hombres. Y sus demandas computacionales son tan elevadas que ningún ser humano podría adecuarse a ellas, lo que deja afuera la posibilidad de que el aprendizaje de las herramientas de cálculo apropiadas por parte de los electores llegue en algún momento a auto-validar EUT. *Es el proceso el que, por lo tanto, debe revisarse.* Por otra parte, en la medida en que se pretenda contar con un modelo adecuado, ese proceso “realista” debe dar lugar a un resultado razonable, cuyo conjunto ya ha identificado la teoría convencional (EUT). En suma, se pretende un modelo descriptivo a nivel del proceso que arroje *decisiones racionales*.

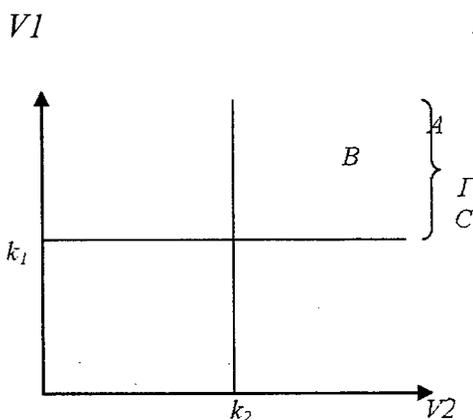
3.4-El proceso racional en Simon (1955)

Despojado de todo tecnicismo el proceso es el siguiente. Supongamos que alguien quiere vender una casa. En lugar de analizar todos los escenarios alternativos relevantes, con la innumerable serie de ofertas posibles en cada uno de ellos, descontados a la tasa apropiada, el hombre del modelo de Simon (1955) toma un camino mucho más simple. Decide cuánto tiempo está dispuesto a esperar para venderla (es decir, tiene en mente un *rango* de búsqueda de ofertas) y decide además cuál es el precio mínimo que está dispuesto a aceptar, o en términos más generales, cuál es el *nivel de aspiración* (NA) que tiene para la venta. Una vez que recibe una oferta que satisface ese nivel de aspiración, la vende; y si recibe varias ofertas satisfactorias, existe un proceso dinámico el cual permite seleccionar la oferta más alta.

El mayor realismo de la descripción descansa en el hecho de que aunque generalmente los sujetos no pueden conocer ni concebir todas las opciones a que

se enfrentan, podrían en principio identificar el subconjunto de opciones satisfactorias. Esto no significa que conocen *a priori* todos los elementos de este subconjunto (pues podrían ser infinitos o tan numerosos que, en la práctica, su conocimiento sería inalcanzable). Más bien significa que si una opción satisfactoria se presenta, el individuo la reconoce en cuanto tal. Lo conocido *a priori* no son las alternativas, sino el *estándar* para determinar que una opción cualquiera califica como satisfactoria. Por esta razón resulta necesario suponer que los individuos poseen un *nivel de aspiración* (NA). Dado NA para un sujeto S , se puede ahora atribuir a S una función de utilidad (valor) que tenga sólo dos valores: 0, si la opción no es satisfactoria y 1 si lo es. Utilizando esta función, S puede ahora identificar a todas las opciones satisfactorias que se le presenten. Son todas aquellas a las que la función valor asigna 1. Queda determinado así lo que podríamos llamar un “conjunto satisfactorio” (debido a que sus componentes satisfacen ciertas restricciones que permiten descartar a las opciones no elegibles).

Supongamos, para ilustrar la idea, que S selecciona de entre los componentes de un conjunto de opciones M según dos características que considera valiosas ($V1$ y $V2$), pero impone la restricción de que una opción es satisfactoria solo si contiene la propiedad $V1$ en una cantidad mayor o igual a $k1$ y la propiedad $V2$ en una cantidad igual o mayor a $k2$. Los pay-offs satisfactorios son, pues, los pertenecientes a la región Γ del cuadrante inferior.



Las alternativas satisfactorias pueden ser muchas, en principio infinitas. Entre ellas se encuentran A , B y C . Como la función asigna el valor 1 a cada una de ellas, no permite ordenarlas según lo que podríamos llamar su “grado de satisfactoriedad”. Permite construir un *conjunto* satisfactorio, pero *no* permite establecer un *orden débil* entre sus componentes. En realidad, no puede hacerlo, porque ello obligaría a reintroducir la función de utilidad estándar (es decir, la vigencia del conjunto de axiomas tradicional), que se desea descartar, por cuanto S debería poder conformar un ordenamiento *completo y continuo* entre las opciones satisfactorias. Cerrada esta vía, para sugerir una regla de decisión alternativa a la maximización Simon distingue dos escenarios básicos de elección:

1-*Estático*: S identifica una opción satisfactoria y la elige. Termina el proceso de elección.

Es necesario admitir que el escenario estático es algo decepcionante. Un sujeto que ha puesto en venta su casa y ubica su nivel de aspiración en $\$X$ procedería racionalmente si aceptara la primer oferta igual o mayor que X que se le presentara. Pero en el ejemplo se ve claro que todo depende ahora del nivel de X (si fuera igual a $\$10$, probablemente no diríamos que su procedimiento es racional, excepto en el sentido restringido en que racionalidad se reduce a coherencia: el individuo elige de acuerdo a la regla que él mismo se ha puesto). Es importante enfatizar que Simon necesita que el nivel de aspiración no esté meramente dado, sino que esté ubicado en un nivel *racional*. Una solución podría ser que estuviera dado por el precio de mercado de un inmueble como el ofrecido. Esta es la manera en que el hombre práctico resolvería el problema. Sin embargo, Simon no opta por esta alternativa. ¿Por qué no lo hace? Una interpretación plausible es que esto significaría que en su teoría la decisión racional ya presupone los resultados del mercado, cuando la teoría estándar de la decisión procuró brindar micro-fundamentos. Dicho de otro modo, en la visión estándar las decisiones individuales producen los resultados a nivel agregado (los

precios) y ésta es la importancia central que la teoría económica asigna a EUT: a partir de la acción separada de agentes maximizadores se deriva la curva individual de demanda de pendiente negativa, cuya suma horizontal da lugar a la curva agregada y junto con la oferta agregada se determinan los precios de mercado. En el caso que estamos examinando, en cambio, las decisiones individuales dependerían para ser racionales de la existencia de fenómenos agregados (como los precios), que quedarían sin explicar. Si Simon deseara proporcionar una teoría de la decisión alternativa a la estándar, que, que retuviera la capacidad de aquella de brindar micro-fundamentos a la teoría económica, no podría aceptar una solución semejante.

2-Dinámico: S va recibiendo opciones de manera secuencial y, a medida que las recibe, las compara con su NA. Si recibe inicialmente muchas buenas ofertas, eleva su NA; si no recibe ninguna aceptable, lo baja hasta que alguna inicialmente no satisfactoria se convierta en satisfactoria.

Como por lo regular los sujetos no conocen *a priori*, ni con certeza ni con probabilidad, *todas* las alternativas a que se enfrentan, se ven obligados a emprender una búsqueda. La *movilidad* del NA es lo que asegura la *existencia* (que haya al menos una) de opciones satisfactorias y su *unicidad* (que al final del proceso sólo quede una). Debe notarse que, a diferencia de lo que ocurre en la teoría económica estándar de la elección, no es el índice de utilidad lo que permite identificar la mejor opción en este caso. En vez de seleccionar directamente (mediante un cálculo consistente en aplicar $\sum U(R_i) \cdot p_i$ la mejor opción, el sujeto va descartando las no satisfactorias hasta que sólo queda en pie una única alternativa, o bien va ampliando el rango de búsqueda o bajando el nivel de aspiración hasta que aparece una alternativa. Por otra parte, ni siquiera se presume que el proceso de generación o eliminación de opciones (vía corrimiento del NA) culmine en la elección de la *mejor*; basta con que seleccione una que sea *satisfactoria*. El sujeto deja de buscar sin que necesariamente haya maximizado. Aunque la decisión no sea completamente racional en el sentido

estrecho de óptima, se consigue una aproximación conveniente con mínimos recursos.

3.5-Racionalidad de la Decisión

A pesar de su fuerte poder sugestivo y heurístico, el modelo dinámico de decisión propuesto por Simon es problemático. Que los individuos elijan una alternativa que satisface su nivel de aspiración, y que “ajusten” sus elecciones mediante movimientos de su NA, ilustra un mecanismo que, en el mejor de los casos, poseería relevancia meramente *descriptiva* (ya matizaremos esta afirmación algo más adelante). Pero no implica nada para la cuestión *normativa* que atañe a la *racionalidad* de esa decisión. Una cosa es describir cómo elige la gente y otra es sostener que esas elecciones son racionales. Pero Simon concibe BM como una teoría alternativa (a EUT) de la decisión *racional*. Debe, pues, proporcionar un criterio para determinar en qué circunstancias un cierto nivel de aspiración es racional, o, lo que parece ser equivalente, el modelo debe determinar cuándo es racional *cesar* de buscar. Efectivamente, el carácter racional de la elección depende de la racionalidad de la búsqueda. Más precisamente, de la existencia de “stopping rules” racionalmente defendibles. La racionalidad se desplaza por ende al *proceso*.

¿Por qué es racional este *proceso*? ¿Conduce a resultados que estaríamos dispuestos a llamar “racionales”? O puesto en términos más concretos: ¿qué le impide al sujeto de Simon vender su casa, para seguir con el ejemplo dado, a un precio irrisorio? Como hemos visto, la respuesta de Simon pasa por incorporar una *dinámica* al modelo: si el sujeto recibe ofertas por encima de su nivel de aspiración, puede moverlo hacia arriba. Del mismo modo, si las ofertas que obtiene son muy bajas en relación al nivel de aspiración, puede bajarlo (manteniendo constante el rango de búsqueda), o dejarlo constante, ampliando el horizonte temporal (es decir, se resigna a vender más barato, o a tardar más tiempo en la búsqueda de compradores).

La racionalidad de todo el proceso depende de la *racionalidad del nivel de aspiración*, que es, en última instancia, el que dispara la decisión: para ser aceptada como satisfactoria, una oferta debe superar el nivel de aspiración. Si damos un paso más y nos preguntamos por la racionalidad del nivel de aspiración, una respuesta posible en defensa de Simon es la siguiente: no importa de dónde salga, puede muy bien ser arbitrario. Y en todo caso, se puede mover con las circunstancias. Lo que interesa es que el *resultado* de la decisión es semejante al de EUT (el sujeto vende su casa al precio más alto que es capaz de obtener) y el *proceso* es ahora compatible con la capacidad computacional limitada del ser humano. Simon parece cumplir con los dos puntos de su objetivo: el modelo conduce a decisiones racionales sin caer en incompatibilidad psicológica.

Notemos, sin embargo, que si NA puede ser fijado arbitrariamente, la “racionalidad” de la decisión es de tipo *subjetivo y coherencista*. Uno tiene la sensación de que *cualquier* comportamiento puede ser presentado como racional en el marco de este esquema. De lo contrario, ¿qué constituiría una conducta irracional en esta situación? No habiendo una regla para ordenar las opciones satisfactorias ni un criterio independiente para determinar la racionalidad del nivel en que es fijado el nivel de aspiración, sólo sería irracional ponerse un nivel de aspiración y elegir una opción que no sea satisfactoria respecto del mismo. Debe darse que el sujeto declare (honestamente) que su NA está fijado al nivel X y elija una opción por debajo de X . Huele a simple incoherencia: al fijar X como NA se compromete a que sólo hay opciones satisfactorias (es decir elegibles) por encima de X ; sin embargo, elige una que se encuentra por debajo. La irracionalidad es reducida entonces a incoherencia entre sus creencias y fines.

3.6-Incompatibilidad Psicológica

Los seres humanos “reales” no tienen acceso a toda la información que sería necesaria para maximizar, ni la capacidad computacional para procesarla y llevar a buen término el cálculo optimizador. La limitación es fuerte: aunque tuvieran

acceso gratuito e ilimitado a la información, no serían capaces de *procesarla*, dadas sus limitaciones de cálculo. Y este es el inconveniente al que Simon (1955) promete encontrarle remedio, o al menos mostrar en qué dirección hay que buscarlo.

Recordemos que el problemal hombre modelado por Simon es vender una casa a un precio respetable, sin contar con los precios de mercado para establecer un nivel de aspiración razonable (porque de lo contrario, no sería un modelo *económico* de la decisión racional, es decir un modelo que explique el precio de mercado por las decisiones individuales, y no las decisiones individuales por el precio de mercado). Vimos que si de hecho están todas las ofertas frente al sujeto, se puede mover el nivel de aspiración de tal manera de asegurar un precio semejante al de mercado. De hecho, *el que todas las ofertas sean simultáneas es el sucedáneo del conocimiento perfecto*. Pero cuando las ofertas son *sucesivas*, y hay que renunciar a una para escuchar la siguiente, se origina un problema de información. Planteemos el siguiente experimento mental: alguien quiere vender su casa, digamos, a 10 centavos (en términos más técnicos, el nivel de aspiración del sujeto está en 10 centavos). Claramente esto no es razonable, y uno esperaría que el modelo de Simon lo prohíba. Si las ofertas fueran simultáneas, rápidamente la plétora de gente ofreciendo cantidades mayores impulsaría al nivel de aspiración al alza, hasta que el precio sea, digamos, razonable. ¿Pero si las ofertas son sucesivas, es decir, si hay *incertidumbre* en cuanto a la próxima oferta que voy a recibir?

Simon era consciente de este problema, y le reserva el apéndice de su trabajo para tratarlo. Veamos lo que dice:

"We suppose that an individual is selling a house. Each day (or other unit of time) he sets an acceptance price: $d(k)$, say, for the k th day. If he receives one or more offers above this price on the day in question, he accepts the highest offer; if he does not receive an offer above $d(k)$, he retains the house until the next day, and sets a new acceptance price, $d(k+1)$.

Now, if he has certain information about the probability distribution of offers on each day, he can set the acceptance price so that it will be optimal in the sense that it will maximize the expected value, $V[d(k)]$, of the sales price.” (Simon, 1955. p.115)

Como puede verse, la respuesta de Simon no puede ser más decepcionante: *se necesita que el individuo conozca la distribución de probabilidad de las ofertas futuras*. Contar con agentes del tipo del demonio de Laplace es la única manera en que se puede poner un nivel o precio de aceptación óptimo, en el sentido de que maximice el valor esperado del precio de ventas. Por supuesto, esto es algo que requiere tanta capacidad computacional y de acceso a la información como los modelos *mainstream*, y reaviva el problema de la incompatibilidad psicológica, algo que por otra parte Simon admite. En efecto, luego de presentar los cálculos maximizadores, aclara:

“It is interesting to observe what additional information the seller needs in order to determine the rational acceptance price (...) He needs, in fact, virtually complete information as to the probability distribution of offers for all relevant subsequent time periods”(*Op. cit.*, p. 117, resaltado propio)

Todo el contenido empírico de EUT proviene de las restricciones formales impuestas a preferencias y expectativas. Al abandonar estas restricciones y no poner otras en su lugar el modelo behaviorista se queda con las manos vacías. Llamativamente, en un trabajo ulterior Simon considera que EUT es casi vacía:

“the conclusions that are reached by neoclassical reasoning depend very much on the ‘auxiliary’ factual assumptions that have to be made to define the situation and very little on the assumption of substantive rationality –in particular, the utility maximization assumptions.... Almost all the action, all the ability to reach nontrivial conclusions, comes from the factual assumptions and very little from the assumptions of optimization” (Simon, 1986, p. 28)

Sin duda tiene razón: el contenido informativo de los modelos económicos proviene en mayor medida de sus hipótesis “auxiliares”, pero comparando una concepción de la racionalidad con otra no cabe duda de que EUT tiene mayor contenido empírico que el modelo behaviorista presentado por Simon en 1955.

Esto remite a una discusión tradicional en metodología: la creencia de que una concepción es tanto mejor cuanto más resultados diversos logra acomodar. Las diversas generalizaciones de EUT, han sido debilitamientos de la misma con el objetivo de acomodar anomalías. El tipo de desarrollo descrito por Starmer (2000) observado en la teoría estándar de la racionalidad es claramente anti-Popperiano: cada nueva versión de la teoría la debilita un tanto para convertir un contraejemplo en algo esperable. Desde esta perspectiva, el modelo behaviorista no altera esta forma de desarrollo, sino todo lo contrario: posee *menos* contenido empírico que los sucesivos debilitamientos de la teoría de la decisión inicial, los cuales pese a todo imponen restricciones al comportamiento de los agentes. Este punto no ha sido entendido suficientemente en metodología de la economía. Todavía hay autores que ven en la propiedad proteica de las teorías (mejor sería decir, marcos interpretativos) una virtud y no un defecto. Conlisk, por ejemplo, se congratula de que BM pueda explicar prácticamente todo (no salen resultados determinados).

El propio Simon, sin embargo, nota que algo anda mal en una teoría capaz de acomodar cualquier fenómeno y su contrario. Por ello, en su *Apéndice* al modelo behaviorista se esfuerza por obtener resultados determinados, pero para ello se ve forzado a reintroducir de manera *ad hoc* los recursos explicativos de EUT. Su artículo de 1955 es un ejercicio de psicología *folk* revestido con variables y algunos términos matemáticos.

Simon pretende realizar un modelo de la decisión económica racional que no sea pasible de incompatibilidad psicológica. La racionalidad del ser humano no es ilimitada, es *bounded rationality*. Para dotar de sentido a lo que de otra manera sería un mero slogan, construye un modelo de decisión compatible con las

restringidas capacidades computacionales humanas, y el acceso a la información limitado, las cuales sientan las bases para una imagen del “hombre económico” más realista.

Sin embargo, su modelo de 1955 enfrenta problemas. Si bien postula un proceso psicológicamente persuasivo, falla a la hora de poner límites a los resultados de ese proceso. Curiosamente la solución deja entrar en el apéndice aquello que había conjurado en el cuerpo de su *paper*, renunciando por lo tanto a obtener un verdadero modelo de *bounded rationality*.

El legado de Simon es indiscutible. Nadie puede poner en duda que abrió un sendero hasta entonces inexplorado en la economía, por el que circula ahora aquella multitud de modelos heterogéneos agrupados bajo el rótulo de *Bounded Rationality*. En tal sentido, Simon ha sido exitoso en “señalar el camino por el que deben hacerse los cambios” (Simon, 1955, p.99); pero es dudoso que él mismo haya podido transitarlo. La *herencia* dejada es sin embargo importante, y antes de emprender ulteriores críticas, esta vez dirigidas al supuesto contenido empírico del modelo de Simon, diremos algunas palabras con respecto a ella.

3.7-Derivaciones ulteriores

Una de las alternativas más fructíferas al paradigma de elección racional convencional está dada por los modelos que incorporan la racionalidad limitada (*bounded rationality*). No resulta sencillo definir estos modelos. Una caracterización significativa en la medida que alude a los problemas para circunscribir los límites de la corriente es la siguiente:

“Theories of bounded rationality are worse than pornography, at least in one sense. Pornography has been famously described as something which you can't define but you know it when you see it. I'm not sure that we know theories of bounded rationality when we see them.” (Lipman, 2002, p.926)

Sin duda, el núcleo común de estos modelos es el *abandono de la idea de optimización libre* que caracteriza a los modelos convencionales. Esta idea por

simple que parezca lleva implícita al menos una dualidad de visiones incompatibles entre sí, que han tomado caminos separados. Siguiendo con el paradigma tradicional, un grupo de economistas continúa suponiendo la optimización pero esta vez sujeta a *restricciones* (de tiempo, o computacionales, etc) Tal es el caso de Stigler (1963), o más recientemente Sargent (1993). Los agentes modelados continúan optimizando, pero su búsqueda está ahora restringida por limitaciones de tiempo, o costos de oportunidad. En el conocido ejemplo de Stigler supongamos que alguien quiere comprar un auto usado y consiguientemente comienza a buscar alternativas de unidades posibles. El individuo se detiene cuando los costos asociados a la búsqueda de nueva información sobrepasan los eventuales beneficios; de otra manera, el proceso continuaría indefinidamente. El automóvil elegido no es el óptimo absoluto, sino el óptimo dada la restricción impuesta por los costos asociados a la búsqueda de nueva información, y en tal sentido la racionalidad del modelo está *acotada*.

La estrategia de retener la optimización e incorporar restricciones permite introducir las herramientas matemáticas tradicionales, y construir modelos semejantes a los convencionales, sólo que un poco más complejos. Pero quienes reclaman mayor realismo psicológico ven como un error seguir este camino. En palabras de Gigerenzer (2001, p. 5):

"[The] attempts to model limited search leads to the paradoxical result that the models become even less psychologically plausible. The reason is that the desire for optimization is retained. Because optimization under constraints is often referred to as "bounded rationality" (e.g., Sargent 1993), this confusion has led to the dismissal of bounded rationality as a hidden form of optimization—one that just makes everything more complicated. (...) Models of bounded rationality use fast and frugal stopping rules for search that do not involve optimization"

Gigerenzer encabeza, desde el instituto Max Planck de Berlín, el segundo grupo de teóricos de *Bounded Rationality*. A diferencia del de Sargent, grupo del que aspiran a alejarse y al que ven como una continuación de la desalentadora

práctica oficial de construir modelos con agentes desesperadamente irrealistas desde el punto de vista psicológico, su recepción de las sugerencias de Simon es bastante más literal, y aspira a una ruptura mucho más profunda con la economía convencional. Esto lo lleva a romper con la idea de optimización matemática. Como la economía tradicional suele entender racionalidad como sinónimo de maximización (Cfr. Cap. VI), los herederos de Simon parecen condenados a modelar individuos irracionales. Nada de esto sucede, sin embargo:

“Bounded Rationality is not irrationality. A sharp distinction should be made here. The theory of bounded rationality does not try to explain trust in lucky numbers or abnormal behavior of mentally ill people. In such cases, one may speak of irrationality. However, behavior should not be called irrational simply because it fails to conform to norms of full rationality. A decision maker who is guided by aspiration adaptation rather than utility maximization may be perfectly rational in the sense of everyday language use” (Selten, 2001. p. 15)

Ambas ideas (construir modelos con agentes no maximizadores, separar la racionalidad de la maximización) aparecen, como hemos visto, en el modelo de Simon (1955). Allí los agentes no maximizan, sino que satisfacen, y esta satisfacción está mediada por un mecanismo donde el nivel de aspiración se adapta convenientemente a las circunstancias. Como veremos, los agentes no examinan todas las opciones posibles (algo que además de costoso, es imposible), sino que buscan hasta que encuentran una alternativa satisfactoria, que bien puede no ser la mejor.

El tercer punto heredado por el grupo de Gigerenzer consiste en la desaparición de toda mención al procedimiento cardinal para medir utilidades y al cálculo posterior consistente en adicionar productos de probabilidades y utilidades.

“Bounded rationality dispenses with the notion of optimization and, usually, with probabilities and utilities as well. It provides an alternative to current norms, not an account that accepts current norms and studies when humans deviate from these norms. Bounded rationality means rethinking the norms as well as studying

the actual behavior of minds and institutions" (Gigerenzer, 2001, p.6) Como veremos, la propuesta de Kahneman y Tversky puede interpretarse como un *tercer sentido* de *Bounded Rationality*, equidistante tanto de Sargent como de Gigerenzer, lo cual constituye un indicador claro de lo prolífico de la herencia de Simon.

Capítulo IV

Prospect Theory

4.1-Introducción

El Premio Nobel de Economía fue otorgado en el año 2002 a dos eminentes académicos. El primero es uno de los economistas pioneros de *experimental economics*, Vernon Smith; el segundo, un psicólogo que jamás tomó un curso de economía, Daniel Kahneman.

El interés académico temprano de Kahneman se orientó hacia la psicología de la percepción, un campo alejado de la Teoría de la Decisión Económica. A pesar de sus orígenes comunes, los desarrollos ulteriores determinaron el alejamiento entre psicología y economía. No extraña que la divergencia disciplinar haya ocasionado un primer contacto un tanto curioso:

“My first exposure to the psychological assumptions of economics was in a report that Bruno Frey wrote on that subject in the early 1970's. Its first or second sentence stated that the agent of economic theory is rational and selfish, and that his tastes do not change. I found this list quite startling, because I had been professionally trained as a psychologist not to believe a word of it. The gap between the assumptions of our disciplines appeared very large indeed”
(Kahneman, 2003, p. 1)

En la *Nobel lecture* brindada en 2002, Kahneman señala que fue a instancias de economistas como Richard Thaler que tanto él como Tversky comenzaron a enfocarse de manera creciente en el trabajo interdisciplinario. Resulta interesante notar que años más adelante el mismo Thaler forma parte integral de un conjunto de economistas que utilizan *Prospect Theory* como base de un conjunto de modelos económicos agrupados bajo el nombre de *Behavioral Economics*. No es aventurado conjeturar que sus tempranos esfuerzos por el trabajo interdisciplinario hayan estado motivados por la vieja idea de Herbert Simon de introducir supuestos más *realistas* en los modelos económicos (Cfr. Camerer *et al.*, 2002).

Prospect Theory se presenta como un modelo de decisión individual en condiciones de riesgo, que resulta psicológicamente realista y descriptivamente exitoso en un rango significativo de circunstancias. En este sentido pertenece a la tradición inaugurada por Simon, y sus mismos autores adscriben a la corriente de modelos *Bounded Rationality* (Kahneman y Tversky, 1979). La teoría incorpora características propias del aparato cognitivo humano, en particular de la psicología de la percepción, campo en el cual Kahneman ha realizado aportes en los comienzos de su carrera. En 1955 Simon se lamentaba del atraso relativo de la psicología en teoría de la decisión, lo que obligaba a hacer uso del conocimiento de sentido común:

“Psychologists have certainly been concerned with rational behavior, particularly in their interest in learning phenomena. But the distance is so great between our present psychological knowledge of the learning and choice processes (...) that a marking stone placed halfway between might help travelers from both directions to keep their courses. Lacking the kind of empirical knowledge of the decisional processes that will be required for a definitive theory, the hard facts of the actual world can, at the present stage, enter the theory only in a relative unsystematic and unrigorous way” (Op. cit, p. 100)

Bajo esta luz puede considerarse el trabajo de Kahneman y Tversky como la respuesta de la psicología cognitiva a la teoría de la decisión de los años '50. En los años que las separan se han producido avances y nuevas conceptualizaciones (ver Cap. V, sección 2) de acuerdo a las cuales surge una imagen novedosa del *proceso* de toma de decisión.

4.2-Intuición y razonamiento

El aparato cognitivo humano posee dos maneras de pensar y realizar juicios, las cuales se corresponden en varios sentidos con aquello que la tradición filosófica cartesiana llamó *intuición* y *razonamiento*. El ejemplo paradigmático de razonamiento es el del cálculo matemático, y sus características pueden ser

sacadas a la luz a partir de las notas que se juegan en la operación de hacer alguna operación algebraica. Al multiplicar 12 por 258 (Kahneman, 2001, p. 1450) se pone un marcha un proceso *deliberado* y que requiere *esfuerzo consciente*. En contraposición, la intuición es inmediata, espontánea y no requiere esfuerzo consciente.

En la vida cotidiana suele predominar el modo *intuitivo* de pensar, aunque es posible identificar también un cierto control que se ejerce sobre las operaciones mentales y la conducta: en general los individuos no expresan todo lo que pasa por su mente, ni obedecen cualquier sugerencia interna. Hay una suerte de barrera que imposibilita la salida de ciertos impulsos. Sin embargo, el control ejercido es muchas veces *laxo* y deja permear acciones y pensamientos erróneos.

Un ejemplo experimental fácilmente reproducible es el siguiente:

“Una raqueta y una pelota cuestan 110 \$ en total. La raqueta cuesta 100 \$ más que la pelota. ¿Cuánto cuesta la pelota?”

Investigaciones empíricas revelan que una mayoría de personas a las que se les realiza esta pregunta afirman que la pelota cuesta 10 \$.

“Frederick (2003) found that many intelligent people yield to this immediate impulse: 50 percent (47/93) of a group of Princeton students and 56 percent (164/294) of students at the University of Michigan gave the wrong answer” (Kahneman, 2001, p. 1450),

lo que testimonia el escaso control ejercido por el Sistema II sobre el Sistema I.

El pensamiento intuitivo no está asociado necesariamente con errores. Las respuestas brindadas son muchas veces exactas y rápidas; el ajedrecista que advierte instantáneamente una posición de jaque mate en tres jugadas no está calculando las innumerables combinaciones posibles, ni razona el conductor avezado capaz de conversar o dejar vagar la imaginación al mismo tiempo que

obedece las señales de tránsito, esquivando obstáculos, etc. Ambos son ejemplos de la eficiencia del modo intuitivo de pensamiento. El grupo reunido bajo el nombre de *Adaptive Behavior and Cognition* (ABC), del Instituto Max Plank, en Berlín, exploran este campo. (Gigerenzer y Selten, 2001; Hutchinson y Gigerenzer, 2005)

4.3-Sistema I y Sistema II

La distinción entre intuición y razonamiento ha ido ganando espacio en las publicaciones de psicología en las últimas décadas. Tal situación es relativamente nueva: hasta mediados del siglo pasado, el cuerpo de conocimiento científico disponible acerca de la *percepción* como proceso y de los fenómenos perceptivos en general era muy superior al de los *procesos internos* como la intuición o el razonamiento. Utilizando la nomenclatura introducida por Stanovich y Est (2000), Kahneman utiliza el nombre neutral “Sistema I” y “Sistema II” para referir a ellos. La intuición se encuentra entre el sistema perceptivo y el razonamiento, y comparte características de ambas estructuras. Un cuadro puede ayudar a presentar la información:

Cuadro 1: Tres sistemas cognitivos (adaptado de Kahneman, 2001, pág. 1451)

	PERCEPCION	SISTEMA I (INTUICIÓN)	SISTEMA II (RAZONAMIENTO)
PROCESO	Rápido, Paralelo, Automático, Sin Esfuerzo	Sin	Lento, Serial, Controlado, Demanda Esfuerzo
CONTENIDO	Percepciones / Estímulo presente / Atado al Estímulo	Representaciones Conceptuales /Pasado, presente, futuro/ Puede ser evocado mediante el lenguaje	

La intuición como *proceso* comparte sus notas con la percepción: como no demanda esfuerzo consciente puede ser paralelo, automático y en relación al

proceso que pone en marcha el Sistema II, es rápido. En cuanto al *contenido* la intuición comparte notas con el razonamiento, en la medida en que trabaja con representaciones conceptuales pasibles de ser suscitadas por el lenguaje. No está restringida, como la percepción, a procesar estímulos presentes.

El lugar ocupado por el Sistema I y sus vínculos con la percepción, han sugerido la posibilidad de utilizar el vasto conocimiento científico relativo a los fenómenos perceptivos como fuente indirecta de hipótesis aplicables al proceso intuitivo primero, y al proceso de toma de decisión más adelante.

4.4-El Error de Bernoulli

Una característica central del sistema cognitivo es que está diseñado para percibir *cambios y diferencias*. El *objeto* de percepción resulta del contraste entre las características del estímulo focal y las del contexto previo y concomitante. Las razones pueden muy bien ser evolutivas: anulando los estados uniformes o constantes, el aparato perceptivo humano es capaz de maximizar el flujo de información a procesar.

Un ejemplo de la vida cotidiana puede ayudar a aclarar el punto: la sensación al sumergir la mano en agua a 20 grados centígrados es de *calidez* si el estímulo previo es el de un contacto prolongado con agua considerablemente más fría³⁶. Es evidente que la percepción se ejerce sobre el *cambio* entre un estado inicial (agua fría) y un estado final (agua a 20 grados centígrados). Para anticipar la temperatura percibida por un sujeto al sumergir la mano en agua a 20 grados centígrados, es necesario conocer el estado previo. El impacto inicial, asociado al cambio abrupto de estado, va perdiendo progresivamente fuerza. Una vez que la mano ha permanecido algún tiempo en el agua a temperatura constante, el aparato cognitivo deja de percibirla como cálida o fría. El fenómeno de *adaptación*, del cual nos ocuparemos más adelante, ha tenido lugar. Sólo cuando

³⁶ El ejemplo figura en Kahnman (2001) pág. 1454.

cambie nuevamente el estado actual se pondrá en marcha la percepción, arrastrada por la modificación del punto de referencia, lo cual revela la naturaleza *reference-dependent* característica de la percepción.

Algo análogo podría suceder con la *intuición*. Si este fuera el caso cabría esperar que los seres humanos, enfrentados a un problema de decisión, valúen de manera intuitiva los resultados *con respecto a un punto de referencia*, y no en sí mismos. Multitud de experimentos muestran lo acertado de la conjetura. Aquello que brinda utilidad es el pasaje de no tener un bien a tenerlo, o de ganar o perder unidades monetarias. Así como el sistema perceptivo anula los estados que permanecen constantes –mediante el fenómeno de *adaptación*–, una súbita ganancia es valorada durante un cierto tiempo, pero en cuanto pierde su carácter novedoso simplemente deja de ser fuente de placer o displacer, y ya no da lugar a experimentar utilidad alguna.

El descubrimiento del carácter *reference-dependent* de la valuación es un paso fundamental, porque inicia un camino que se aleja rápidamente del seguido por la Teoría de la Decisión Racional. En ella la valoración de los resultados no depende del punto de referencia, sino de su valor absoluto o estado final:

“From the vantage point of a student of perception, it is quite surprising that in standard economic analyses the utility of decision outcomes is assumed to be determined entirely by the final state of endowment, and is therefore reference-independent. In the context of risky choice, this assumption can be traced to the brilliant essay that first defined a theory of expected utility (Daniel Bernoulli, 1738). Bernoulli assumed that states of wealth have a specified utility, and proposed that the decision rule for choice under risk is to maximize the expected utility of wealth (...). The proposition that decision makers evaluate outcomes by the utility of final assets positions has been retained in economic analyses for almost 300 years. This is rather

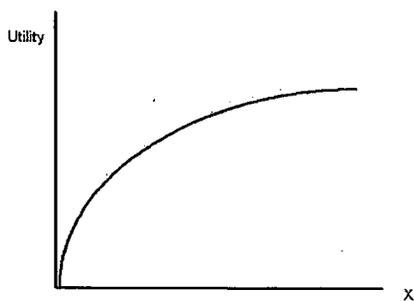
remarkable, because the idea is easily shown to be wrong: I call it Bernoulli's error" (Kahneman, *Op.cit.*, pág. 1455, subrayado propio)

La teoría de la decisión, desde Bernoulli, *predica utilidad de estados finales*: son los montos monetarios, los activos o los bienes en sí mismos los objetos de los cuales extraen utilidad los individuos. El argumento de Kahneman apunta al núcleo del saber económico tradicional. Los portadores de utilidad no son estados finales, sino *cambios* con respecto a un punto de referencia. La Teoría de la Decisión convencional ha partido de un supuesto psicológicamente falso. De acuerdo a Kahneman, el "error de Bernoulli" (el error que consiste en suponer que los portadores de utilidad son los *estados finales*, en lugar de los *cambios*) ha desencaminado a la economía durante 300 años.

4.5-Función valor y función de utilidad

Un gráfico ayudará a mostrar los alcances de la cuestión. La *función de utilidad* convencional constituye una pieza clave en la teoría económica. Su representación habitual es la siguiente:

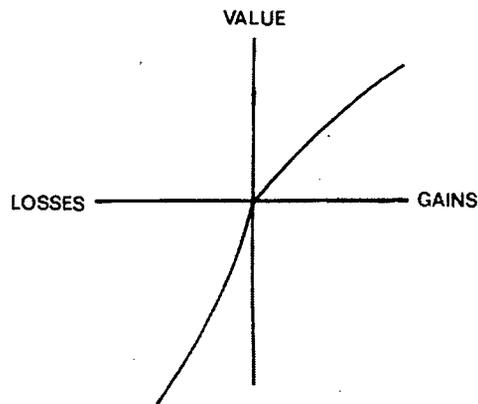
Función de Utilidad



En el eje horizontal figuran bienes o montos monetarios *finales*; ellos son los "portadores de utilidad" de acuerdo a la economía convencional. En condiciones de certeza, la *forma* de la pendiente se interpreta como utilidad marginal decreciente: a más bienes más utilidad, pero la utilidad individual que

proporcionan las unidades adicionales es cada vez menor. Veamos ahora la función alternativa propuesta por Kahneman y Tversky:

Función Valor (Fuente: Kahneman y Tversky, 1979, p. 279).



Resulta útil comparar este gráfico con el convencional, porque ambos representan de manera sucinta dos concepciones muy diferentes relativas a cómo asignan valor o utilidad los seres humanos.

1-La función valor de *Prospect Theory* divide el espectro de objetos en dos grandes campos: *pérdidas* y *ganancias*. Ambos se definen con relación a un *punto de referencia*, el cual no es evaluado (ni percibido) de manera negativa ni positiva, sino neutral, y coincide en líneas generales con el *statu quo* de los individuos. Por su parte, en el gráfico correspondiente a la función de utilidad convencional no hay ganancias ni pérdidas, sino *estados*. Tampoco existe el concepto de *punto de referencia*.

2-La pendiente de la curva de ganancias es menos pronunciada en relación a la de pérdidas en la función valor de *Prospect Theory*. La diferencia refleja el mecanismo psicológico conocido como *loss aversion*, según el cual las pérdidas generan una pérdida de satisfacción mayor a la satisfacción obtenida a partir de ganancias de igual magnitud. Perder una cantidad de dinero entraña una desutilidad mayor a la utilidad asociada a ganar o encontrar inesperadamente la misma cantidad de dinero. El valor numérico de *loss aversion* oscila entre 2 y 2.5

(Tversky y Kahneman, 1992). La posibilidad de contar con una medida cuantitativa de *loss aversion* no es una característica menor, porque permite realizar predicciones determinadas³⁷.

3-La *actitud* frente al riesgo se revierte al desplazarnos del campo de pérdidas al de ganancias. Enfrentado a una situación percibida como pérdida, el individuo modelado en *Prospect Theory* se inclina al riesgo; en el campo de ganancias, revela aversión al riesgo. La teoría convencional, por su parte, no relaciona la actitud frente al riesgo con pérdidas o ganancias.

4.6-Función Ponderación

Veamos de qué manera reacciona el aparato cognitivo humano ante las probabilidades dadas en un problema de decisión. La concepción estándar había dado por sentado que las probabilidades de un evento aleatorio eran percibidas de manera transparente por los individuos. Una probabilidad de 0.1 es percibida aquí como poseyendo un peso de 0.1, y los resultados se ponderan por tal número.

³⁷ *Loss aversion* permite dar cuenta de un fenómeno curioso observado desde antiguo en el campo de la teoría de los juegos de azar. La mayoría de los individuos se muestran renuentes a participar en juegos “justos”, del tipo:

“Suponga la siguiente situación: se arroja una moneda normal, y si sale cara usted gana 100 pesos; si sale ceca, usted paga 100 pesos. ¿Aceptaría este juego?”

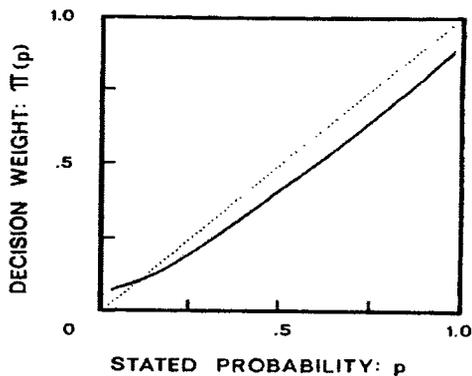
La respuesta suele ser negativa. De acuerdo con *Prospect Theory*, la pérdida de 100 pesos posee un peso negativo mayor al efecto positivo asociado a la ganancia de igual monto. Con el objeto de averiguar en qué medida el monto a ganar debe superar la eventual pérdida para hacer atractivo el juego, se puede ir aumentando el premio a obtener. De manera compatible con *loss aversion*, estudiantes universitarios de los cursos de Epistemología de la Economía, FCE, UBA, periodos 2007-2010, comenzaron a dar respuestas positivas cuando el juego es formulado de la siguiente manera:

“Suponga la siguiente situación: se arroja una moneda normal, y si sale cara usted gana 200 pesos; si sale ceca, usted paga 100 pesos. ¿Aceptaría este juego?”

Lo cual sugiere una medida de *loss aversion* de aproximadamente 2 puntos

La visión estándar puede tener el mérito de la simplicidad. Pero el problema de cuál es efectivamente la forma en que los individuos reales tratan a las probabilidades es una cuestión empírica, y los resultados de Kahneman y Tversky revelan que, lejos de ser percibidas directamente, son transformadas de manera típica al ser procesadas en la decisión. La Función Ponderación de *Prospect Theory* recoge esa información.

Función Ponderación (Fuente: Kahneman y Tversky, 1979, p. 283)



La línea diagonal representa los puntos donde las probabilidades dadas, que aparecen en la base del cuadrado, coinciden con los números que el aparato cognoscitivo les asigna, representados en el lado izquierdo de la figura. Como puede apreciarse las probabilidades pequeñas reciben una ponderación mayor (*overweighting*) y están sujetas a sub aditividad, y las probabilidades intermedias exhiben subcerteza (algebraicamente, $\pi(p) + \pi(1-p) < 1$) En los extremos la función es irregular. Las probabilidades altas se perciben unas veces como idénticas con la certeza, mientras que en otras se exagera su diferencia con la certeza. La reacción ante las bajas probabilidades también se caracterizan por un comportamiento errático: en algunas ocasiones colapsan con la probabilidad nula, y en otras se las sobreestima.

El origen de la Función Ponderación es inductivo, en el sentido que surge de observaciones empíricas. Sin embargo, su rango de validez excede los casos

observados, y es capaz de explicar (y predecir) conductas anómalas para la Teoría de la Decisión Racional.³⁸

4.7-Proceso de toma de decisión en *Prospect Theory*

El proceso de toma de decisión en *Prospect Theory* se divide en dos pasos. En un primer momento, el problema es *editado* de tal manera que el punto de referencia relativo a la decisión queda determinado. Los resultados que lo componen son *codificados* como *ganancias* cuando exceden este punto y como *pérdidas* cuando no lo alcanzan. El proceso de “edición” es fundamental porque determina los *prospects* o loterías a que se enfrenta el sujeto.

Identificados los *prospects*, *Prospect Theory* indica cómo deben ser representados en términos de la teoría. La idea básica es muy simple. Se clasifica a los *prospects* en dos grandes categorías: estrictos (positivos o negativos) y regulares, y luego se especifican dos ecuaciones básicas: una para *prospects* estrictos y otra para regulares. Estos dos procedimientos no son arbitrarios, sino que intentan deliberadamente “simular” las operaciones de *edición* que practican los individuos mismos al enfrentar opciones. La fase de edición es seguida por una de *evaluación*, en la cual se elige aquella alternativa que haga máxima la siguiente función³⁹:

$$V^{40}(x, p; y, q) = \pi(p) \cdot v(q) + \pi(q) \cdot v(y)$$

Donde:

³⁸ La *paradoja de Allais* queda explicada, por ejemplo, a partir de la sobreestimación de las probabilidades pequeñas y la subestimación de las relativamente grandes que caracterizan la Función Ponderación.

³⁹ Kahneman y Tversky, 1979, pág. 276. Una extensión de *Prospect Theory*, conocida como *Cumulative Prospect Theory* (Tversky y Kahneman, 1992) permite trabajar con loterías con gran número de resultados posibles, y soluciona algunos puntos débiles de la formulación original.

⁴⁰ En el caso de las loterías estrictamente positivos o negativos, la ecuación varía ligeramente. Ver Kahneman y Tversky, 1979, p.276.

x e y son los resultados (no necesariamente monetarios)

p y q son las probabilidades asociadas a cada resultado

$\pi (*)$ es el peso asignado a las probabilidades por la Función Ponderación

v (*) es el valor dado a los resultados, de acuerdo a la Función Valor

Describiendo la situación de elección tal como la ven los individuos (es decir, tal como es *editada* por los propios individuos), *Prospect Theory* permite identificar las opciones a que los propios individuos consideran que se enfrentan y proporciona la manera adecuada de formularlas mediante ecuaciones que, se espera, arrojen resultados aceptables desde el punto de vista descriptivo. En la sección siguiente evaluaremos tal aspiración.

4.8-La predicción de *Preference Reversals*

Uno de los contraejemplos más revulsivo para EUT lo constituye la reversión de preferencias; de hecho, es una amenaza de fuste no sólo para el *Golden Dream*, sino también para los mismos fundamentos ontológicos sobre los que está erigida la Teoría de la Decisión Racional, entre otras razones porque termina con la posibilidad de *medir* la utilidad (Cap. V). Sin embargo, el fenómeno se explica fácilmente con las herramientas dispuestas por *Prospect Theory*. Consideremos el caso de *Asian Disease*, presentado más arriba. Allí habíamos visto que las preferencias por el Programa A sobre el B se revierten de acuerdo a la descripción en la que se enmarca la situación. Cuando se formula en términos de *ganancias*, los agentes le atribuyen mayor valor a la cantidad cierta (salvar 200 vidas) que a la lotería correspondiente cuyo valor esperado es de 200 vidas. Lo contrario sucede con la descripción en términos de *pérdidas*. Aquí la gente se comporta huyendo de la certeza.

Lo mismo sucede en el siguiente caso⁴¹:

Problema I

Asuma que acaba de ganar 300 \$. Ahora elija entre:

A: Una ganancia segura de 100 \$

B: 50% de probabilidades de ganar 200 \$ y 50% de probabilidades de ganar 0 \$.

Problema II

Asuma que acaba de ganar 500 \$. Ahora elija entre:

C: Una pérdida segura de 100 \$

D: 50% de probabilidades de perder 0 \$ y 50% de probabilidades de perder 200 \$

Como los estados finales son idénticos, la teoría convencional implica que la decisión de A sobre B (B sobre A) es la misma que C sobre D (D sobre C); después de todo, el Problema II surge de agregar 200 \$ al activo inicial del sujeto, y restar esa misma cantidad de las opciones. Es simplemente un cambio en la presentación que no debería engañar a nadie. Sin embargo, los resultados empíricos muestran una vez más el conocido patrón de *preference reversals*. Enfrentados a la descripción del Problema I, los individuos eligen A sobre B; pero cuando se les ofrece la descripción equivalente del Problema II prefieren la alternativa D sobre la C. El conjunto de decisiones observado es inconsistente con EUT, pero plenamente compatible con *Prospect Theory*. De acuerdo a la Función Valor, en el terreno de las ganancias los individuos huyen del riesgo y se

⁴¹ Kahneman y Tversky, 1981, p. S258. El experimento se realizó con una muestra de 126 individuos, de los cuales un 72% eligió la opción A, y el 28 % la opción B en el primer problema. El segundo contó con 128 participantes, que eligieron la opción C y D en un 36% y 64% respectivamente.

refugian en la opción segura (A), mientras que en el dominio de las pérdidas reaccionan alejándose de la opción segura y tomando en su lugar la incierta (D).

Los distintos casos de reversión de preferencias ante variaciones supuestamente irrelevantes en la descripción constituyen un resultado *anómalo* para EUT, y al mismo tiempo una *confirmación* de *Prospect Theory*, sacando a la luz el conocido patrón de un *experimento crucial*, en el cual coexisten dos teorías rivales, y la evidencia a favor de una es evidencia disconfirmatoria de la rival. En lo que sigue analizaremos las consecuencias de este hecho.

4.9-Decisores Extensionales

La economía convencional daba por descontado que cualquier teoría empírica científica de la elección debía ser edificada sobre alguna versión del *homo economicus* entendido como agente *racional*. Un aspecto esencial del comportamiento racional es el respeto del Principio de Extensionalidad. Llamaremos individuo *extensional* al sujeto (ideal o real) que respeta tal principio.

Informalmente, el Principio de Extensionalidad nos permite determinar cuando dos entidades son iguales. Aplicado sobre conjuntos, establece que A y B son iguales si y sólo si están formados exactamente por los mismos miembros. En el caso de las preferencias, suele denominarse *extensionalidad* (Arrow, 1982) o también *invariancia* (Tversky y Kahenman, 1986), y establece que las preferencias no son afectadas por variaciones insustanciales en la descripción de los resultados de una lotería. En el Capítulo II vimos que este principio, bajo el nombre de *invariancia descriptiva*, constituye uno de los *supuestos* de EUT. El individuo extensional posee preferencias definidas sobre resultados, y consecuentemente, al enfrentar un problema de elección, descarta las descripciones en la que vienen envueltas las opciones. Las principales virtudes de EUT (adecuación normativa, poder unificador, relevancia descriptiva, medibilidad de la utilidad) descansan sobre esta base.

La extensionalidad es violada en los *framing effects*, en los cuales descripciones extensionalmente equivalentes conducen a elecciones diferentes. *Asian Disease* es un ejemplo. Como el Principio de Extensionalidad es fundamental para cualquier comportamiento racional, *framing effects* parece comprometer seriamente la posibilidad de construir una teoría de la decisión *racional* que tenga al mismo tiempo contenido empírico.

Prospect Theory muestra que abandonando el *homo economicus* y el decisor extensional, se puede describir y predecir las elecciones individuales. Al hacer esto parece renunciar a ser una teoría *normativa* de la decisión. (Khaneman 2003), y reemplazar el *homo extensionalis* por otra clase de sujeto, de características muy diferentes al tradicional.

El abandono del agente extensional suscita una serie de interrogantes ¿Qué clase de teoría puede construirse con decisores no extensionales? ¿Cuáles son las implicancias para la teoría de la decisión racional? ¿Qué queda de las aspiraciones de contar con una teoría unificada, descriptivamente adecuada y normativamente correcta de la decisión? En las secciones siguientes serán examinados los retos que *framing effect* genera para la teoría de la decisión, y se adelantarán algunas observaciones relativas a los inconvenientes que surgen si la teoría económica debe ser construida suponiendo agentes intensionales.

4.10-Principio de Extensionalidad

De acuerdo a la metodología convencional, las teorías científicas poseen dos tipos generales de términos: términos lógicos (como las tradicionales conectivas veritativo-funcionales de la lógica estándar, del tipo "y", "o", etc) y términos descriptivos (por ejemplo, enunciados y conceptos). En lenguajes de cierta complejidad la *identidad* de las expresiones tiene que ser definida. Esto se hace estableciendo que dos (o más) expresiones sintácticamente diferentes de una lengua son equivalentes cuando su referente es el mismo (es decir, poseen la misma *extensión*).

El *Principio de Extensionalidad* (PE), afirma que dos descripciones diferentes, pero coextensivas D1 y D2 pertenecientes a un lenguaje L, se pueden intercambiar dentro de L *salva veritate*. Es posible esbozar una teoría semántica muy simplificada, según la cual el hecho de que está lloviendo se puede describir de muchas maneras diferentes, como "llueve", "está lloviendo", "está lloviendo ahora ",etc. Análogamente, los individuos (o clases de individuos) y los predicados pueden ser descriptos por diferentes expresiones. "El autor de *El Quijote* " y "El manco de Lepanto" son descripciones equivalentes para el mismo individuo, Cervantes. "Cervantes" puede ser sustituido por la descripción definida "El manco de Lepanto" dentro de la frase "El manco de Lepanto escribió *El Quijote*", sin ninguna alteración en el valor de verdad del enunciado original⁴².

Pasemos ahora a los casos en que pueden ser formuladas dos descripciones *diferentes* de una *misma* situación de elección. Los llamados agentes *extensionales* son los individuos que al elegir entre opciones no prestan atención a la forma particular en que se describen, y se centran sólo en el "verdadero" contenido de la opción aludido por estas descripciones. Por el contrario, los agentes *intensionales* son aquellos cuyas decisiones dependen de las descripciones particulares de las opciones que enfrentan. Para ilustrar las consecuencias poco deseables que surgen a partir de la violación de extensionalidad podemos comenzar con un caso muy simple. Supongamos que un individuo tiene que elegir entre dos pares de opciones:

Opción I: A o B

Opción II: B o C

⁴² Ver, sin embargo, las *Paradojas de la Denotación* (Simpson, 1975)

Y supongamos que elige A sobre B y B sobre C. En la medida en que sólo presta atención al *contenido* de las opciones que enfrenta (es decir, en tanto sea un decisor extensional), está obligado a elegir A sobre C. Para continuar con el ejemplo, supongamos ahora que A refiere a Cervantes, B a Shakespeare, y C a Goethe. Nuestro decisor prefiere Cervantes a Goethe. Pero supongamos ahora que lo enfrentamos con un nombre alternativo para Cervantes, como "El manco de Lepanto". ¿Cómo reaccionará ahora ante las opciones? ¿Se comportará como un decisor extensional o no? Qué clase de agentes sean los hombres reales es algo que tiene que ser establecido por la investigación empírica. Y una enorme cantidad de evidencia muestra que muchas personas se comportan *intensionalmente* a la hora de tomar decisiones. Es posible que al tener que elegir entre las obras de "El manco de Lepanto" o las de Goethe, un individuo que arrastra consigo emociones fuertemente negativas hacia las limitaciones físicas elija la segunda opción (a pesar de su preferencia anterior de Cervantes por sobre Goethe) Difícilmente pueda llamarse "racional" a un individuo así, y una teoría que intente anticipar sus acciones es incapaz de tener valor normativo, al menos a primera vista.

Quizás pueda verse ahora en más profundidad de qué forma particular se viola el Principio de Extensionalidad en el caso de *Asian Disease*. El problema posee una característica importante: establece un límite superior en el universo de las personas afectadas. En el ejemplo, se describe que 600 personas están en peligro de muerte. Las dos terapias alternativas se introducen en términos de vidas salvadas y perdidas. Al cerrar el universo se da lugar a un paso crucial, porque es posible, ahora, establecer las siguientes equivalencias:

A: 200 vidas se salvarán = A': 400 vidas se perderán

B: 600 vidas se salvarán = B': Nadie va a morir.

Salta a la vista que las opciones son las *mismas* en términos de probabilidades y resultados, las dos variables que desde Pascal utiliza la teoría de la decisión como relevantes en condiciones de incertidumbre. La única diferencia es la *descripción*, la cual se había asumido (correctamente) como irrelevante para decisores *racionales*.

Es conveniente ahora retomar el Principio de Extensionalidad y definir qué papel juega en Teoría de la Decisión. En contextos declarativos el Principio de Extensionalidad es necesario para la sustitución de expresiones materialmente diferentes y extensionalmente idénticas *salva veritate*. Exportado a situaciones de elección, el Principio de Extensionalidad establece que dos descripciones alternativas D1 y D2 de la misma situación de elección puedan ser sustituidas *salva preferencias* (es decir, preservando el orden de preferencias)

Remontándonos a los orígenes de la teoría convencional, puede apreciarse ahora donde reside la raíz del problema. Al modelar una situación de elección sobre la base de la Teoría de los Juegos de Azar del siglo XVII, la disciplina aceptó implícitamente una *ontología* compuesta por loterías. Una lotería está compuesta por dos elementos, resultados R_i y probabilidades p_i asociadas a dichos resultados. Ahora bien, tanto probabilidades como resultados pueden aparecer presentados de múltiples maneras. Por ejemplo, el resultado de ganar 100 \$ se puede expresar como ganar dos veces 50 \$, etc. Esto hace necesario definir *reglas de equivalencia* entre loterías. Dos loterías se consideran equivalentes cuando poseen los mismos resultados R_i y las mismas probabilidades asociadas p_i , aunque su modo de denotación pueda variar⁴³. La relación de equivalencia se establece mediante una versión del Principio de Extensionalidad. Definido sobre *conjuntos*, el principio establece que:

⁴³Hausman(1992) menciona un “Postulado de Reducción” que permite reducir loterías complejas a loterías simples.

Dado un par de conjuntos cualquiera, A y B, si para cada elemento C, C es un elemento de A si y sólo si C es un elemento de B, entonces A es igual a B.

En nuestros términos, dos loterías son iguales si y sólo si sus componentes son iguales, o bien se pueden reducir a una forma común⁴⁴.

Los decisores extensionales habitan un mundo conformado ontológicamente por loterías. La capacidad de cálculo instantánea les permite además operar para reducir las loterías complejas a sus expresiones simples y así elegir de manera consistente. Las cuestiones éticas o emocionales no juegan ningún rol en las decisiones. Si bien el mundo real es diferente a la imagen ideal retratada por la estrategia convencional, se esperaba que los agentes *extensionales* sacaran ventajas suficientes como para eliminar a los *intensionales* –una apuesta darwinista- o que el aprendizaje fuera cerrando la brecha entre el mundo ideal y el real, de tal manera que al final del día la teoría terminaría por autovalidarse. La estrategia de postular una teoría normativa y esperar a que sea exitosa desde el punto de vista predictivo parecía ser una estrategia inteligente y prometedora, pero está basada en seres humanos extensionales.

En las siguientes secciones veremos dos reacciones disciplinarias frente al hecho de que los seres humanos no son *homo extensionalis*. La estrategia convencional, que reconoce (y desestima) el fenómeno y la heterodoxa, que acepta el hecho y trata de vivir con sus consecuencias.

4.11-La estrategia convencional

Para decirlo en una fórmula, la estrategia convencional pasa por considerar la reversión de preferencias como un fenómeno *interesante*, pero *irrelevante*. Hacia 1971, cuando las primeras pruebas de reversión de preferencias comenzaban a

⁴⁴Esto permite determinar que la lotería A con : (100, 50; 0.25) es la *misma* lotería que A': (50, 100; 0.75)

surgir (Lichtenstein y Slovic, 1971) la reacción disciplinar consistió en dudar de la probidad de los datos. Como se ha visto en el Capítulo III, Grether y Plott admitieron la existencia del fenómeno, pero adjudicaron la reversión a un mal diseño experimental. En un intento por disolver la anomalía y mostrar que podía derivarse de la teoría económica tradicional, plantearon y llevaron a cabo nuevos experimentos. Lamentablemente ni la presencia de incentivos, ni el control de los efectos ingreso, ni la cuidadosa eliminación de las posibles conductas estratégicas de los sujetos de experimentación, hicieron desaparecer la reversión de preferencias. (Grether y Plott, 1982 citado por Hausman, 1992, p. 232 y ss). La conclusión de los economistas es curiosa:

“The fact that preference theory and related theories of optimization are subject to exception does not mean that they should be discarded. No alternative theory currently available appears to be capable of covering the same extremely broad range of phenomena. In a sense the exception is an important discovery, as it stands as an answer to those who could charge that preference theory is circular and/or without empirical content.” (Grether y Plott, 1979, p.634. Citado por Hausman, 1992, p. 234)

Parafraseando a Hausman, esto suena a “dado que estos contraejemplos no pueden ser eliminados, los economistas deberían ignorarlos, aunque no sin antes felicitarlos por tener en sus manos una magnífica teoría plena de contenido empírico”. Pero esta apreciación puede no ser del todo justa. Hay una considerable confusión, que excede a Grether y Plott, con respecto al carácter tautológico o empírico de la Teoría de la Decisión convencional, la cual sigue vigente en nuestros días (Cap. VI). Por otra parte, recordemos que existen al menos cinco fines o usos de EUT, y que son conceptualmente diferentes. En rigor, si pensamos que una teoría es un intento de solución a un problema particular, quizás deberíamos hablar de cinco pares teoría-problema diferentes. Sea como sea, la reversión de preferencias compromete seriamente el uso *descriptivo* de la teoría, pero no afecta su capacidad *interpretativa*, la cual

permite iluminar una amplia variedad de fenómenos, ni su función *idealizante*. Desde este punto de vista no hay irracionalidad alguna en quienes al parecer se apegan obstinadamente a EUT a pesar de los contraejemplos. Llamativamente, este grupo incluye tanto a Kahneman, quien no parece ser un adicto a la estrategia convencional, como a los economistas involucrados en *Behavioral Economics*. En el Capítulo VII veremos qué razones hay detrás de su actitud.

Por último, siempre está abierta la puerta para refugiarse en el *Golden Dream* a la espera de que la realidad termine vindicando la teoría en su aspecto descriptivo, al menos en apariencia. Examinemos más de cerca la viabilidad este punto.

La función *descriptiva* de la Teoría de la Decisión exige que los patrones de elección observados se deduzcan del conjunto de premisas. La construcción de una teoría exitosa descriptivamente, que se derive además de axiomas racionales (es decir, normativamente aceptables) está en el núcleo del *Golden Dream* de la visión tradicional. Sin embargo, esta estrategia está en peligro y puede ser criticada por razones epistemológicas: el hecho de que los seres humanos reviertan sus preferencias ante cambios insustanciales en la descripción y en los procedimientos utilizados para sacarlas a la luz, afecta al proyecto tradicional en su forma más profunda: cuestionando la capacidad para *medir* la utilidad en el ámbito de las decisiones humanas.

Como hemos visto en el Capítulo II, EUT incorpora un procedimiento para medir la utilidad. En su momento se generó un amplio debate alrededor de qué tipo de propiedades cardinales implicaba la medibilidad de la utilidad (Mongin, 2001, Fishburn, 1976; Fama, 1972) a lo largo del cual la efervescencia inicial fue perdiendo impulso, pero en su momento fue pensada como un logro extraordinario que investía a la economía de rigor científico y permitía distinguirla del resto de las ciencias sociales. Friedman (1955) se congratula de poder contar con un concepto de utilidad que es tan medible “como la temperatura o la longitud”. La reversión de preferencias descubierta por

Lichtenstein y Slovic en 1971 pone en juicio la validez del método para medir la utilidad en el dominio de los seres humanos.

De acuerdo a la teoría de la medición, dos procedimientos equivalentes para ordenar loterías deben dar lugar al mismo orden. Para ilustrar por qué esto es así, imaginemos el orden surgido de pesar cuerpos y disponerlos de manera creciente en una superficie; intuitivamente este orden debe ser idéntico sea que pesemos los objetos con una balanza de resortes o una balanza de platillos. La forma de sacar a la luz una propiedad cualquiera no debe interferir con ella. Sin embargo, en el caso descubierto por Lichtenstein y Slovic, dos maneras irreprochables de sacara a la luz preferencias conducen a resultados opuestos⁴⁵. El experimento conduce a considerar que el concepto de utilidad implícito en EUT *no puede ser medido* en el ámbito de los humanos⁴⁶. Este resultado tiene una relevancia directa para la economía: los individuos no poseen preferencias bien formadas con respecto a loterías. Y como la maximización de la utilidad esperada depende de la posibilidad de contar con una función determinada, *no puede haber maximización en el sentido convencional del término* (veremos más adelante qué

⁴⁵ "Procedure invariance plays an essential role in measurement theories. For example, the ordering of objects with respect to mass can be established either by placing each object separately on a scale, or by placing both objects on the two sides of a pan balance; the two procedures yield the same ordering, within the limit of measurement errors. Analogously, the classical theory of preference assumes that each individual has a well defined preference order (or a utility function) that can be elicited either by offering a choice between options, or by observing their reservation price.

Procedure invariance provides a test for the existence of a measurable attribute. It would have been difficult to attribute mass to objects if the orderings of these objects with respect to mass were dependent on the measuring device.

Similarly, it is difficult to defend the proposition that a person has a well-defined preference order (or equivalently a utility function) if different methods of elicitation give rise to different choices" (Tversky, 1999, p. 189).

⁴⁶ En el Apéndice del Cap. II figura el procedimiento de medición; si las preferencias se revierten, puede demostrarse que habría al menos un elemento en la escala de preferencias que tendría asignadas *dos* utilidades diferentes al mismo tiempo, lo que implica entre otras cosas la inexistencia de un *función* de utilidad (un elemento del dominio tendría más de una imagen, algo prohibido por la definición de *función*)

modificaciones son necesarias para seguir utilizando el instrumental de maximización en *Prospect Theory*) De esta manera, lo que parecía ser un problema local y meramente técnico se extiende por toda la teoría económica.

Por otra parte, y más grave, la existencia misma del *concepto* de utilidad se pone en entredicho. Recordemos que, siguiendo el imperativo operacionalista, Samuelson había reducido a la utilidad a ser un mero *índice de preferencias*, propuesta que adquirió rápidamente aceptación dentro de la disciplina. En otros términos, todo el contenido referencial del término se agota en señalar un lugar preciso en el ordenamiento de bienes que precede a la construcción de una función de utilidad. Es decir:

Utilidad= Número que indica el orden en el cual se prefieren los bienes.

Esta jugada permitió deshacerse de las interpretaciones psicologistas vinculadas a la tradición utilitarista, y convertir a la utilidad en un dispositivo de sencilla elegancia. En su momento parecía una estrategia inteligente y poco costosa. Sin embargo, esto no fue así. Al reducir el término utilidad a su versión operacionalista, se la ató a la existencia de un orden bien formado. El Principio de Extensionalidad estaba suficientemente protegido dentro de la red conceptual como para dar lo que más adelante se revelaría como una falsa sensación de invulnerabilidad. Pero la caída del Principio de Extensionalidad implica que no existe un orden bien formado en el ámbito de las preferencias. Y al derrumbarse la posibilidad de construir tal orden, ¿qué *contenido* queda para el término “utilidad”? Esta es una segunda razón, conectada con la anterior, que muestra el impacto sobre la teoría de la decisión convencional de la caída de la extensionalidad: *el concepto de utilidad desaparece*. Más adelante veremos las implicancias de este hecho para la evaluación epistemológica del pasaje de EUT al *Prospect Theory*.

Veamos lo que sucede con respecto a la capacidad predictiva. Recordemos que un punto fuerte de la estrategia convencional descansaba en la posibilidad de que

una teoría normativa pudiera ir ganando progresivamente eficacia predictiva. A medida que los agentes fueran educados en los preceptos de la decisión racional, la teoría se autovalidaría. La estrategia convencional descansa en dos supuestos: la adecuación normativa de la teoría, y la posibilidad de que los agentes aprendan a comportarse racionalmente. Pero si los *framing effects* son ubicuos tanto en contextos artificiales como en la vida real, la estrategia convencional no solamente no desencadena el éxito predictivo, sino que lo *imposibilita*. Esto constituye un resultado devastador para las teorías convencionales. Tomar como punto de partida una teoría normativa (*i.e.*, racional) es la fórmula del fracaso para predecir las decisiones de agentes *intensionales*, porque la extensionalidad es un requisito de racionalidad central que se encuentra ausente en un amplio conjunto de circunstancias relevantes.

En este momento una objeción clásica puede ser planteada. Alguien podría verse tentado a pensar que los efectos deletéreos de los *framing effects* sobre la estrategia convencional son sólo aparentes, y en última instancia irrelevantes. Como las descripciones que intervienen en los *framing effects* son descripciones *equivalentes* de la misma situación, los agentes tarde o temprano se harían conscientes del error cometido. A medida que los individuos aprendan de sus errores, los efectos anómalos tenderían a desaparecer. Esta afirmación parece plausible a primera vista, pero carece de sustento. El hecho de que los agentes aprendan o no es una cuestión empírica; y diferentes estudios empíricos muestran que no se puede dar por sentado que lo hagan. El contexto necesario para el aprendizaje requiere la ausencia de consecuencias no anticipadas, además de una estructura ergódica, condiciones ampliamente ausentes del mundo real. Por otra parte, al parecer la conducta generada por la aceptación de un punto de referencia ligado a la descripción es *inevitable* dadas las características del aparato cognitivo humano. El argumento tradicional, según el cual los agentes se comportarán racionalmente si reciben incentivos adecuados y se los provee de suficiente tiempo para recoger información y procesarla, es *falso* desde el punto de vista empírico. Este proceso virtuoso es una imposibilidad psicológica, porque

la percepción es un proceso necesario y previo a la toma de decisiones, y uno de sus determinantes clave es la *descripción* de la situación dada.

4.12-La estrategia heterodoxa: *Prospect Theory* como teoría puramente *descriptiva* de la decisión

Prospect Theory renuncia a modelar individuos racionales y con ello gana poder predictivo. No es de extrañar entonces que sus autores adscriban a la corriente de *Bounded Rationality*, iniciada por Simon (*cf.* Cap. III). Sin embargo, a diferencia del proyecto original de Simon y de muchos de sus seguidores (*v.gr.* Gigerenzer, o Selten), la racionalidad a la que se renuncia en *Prospect Theory* no es la de la maximización. Los agentes modelados son maximizadores, como veremos. El requisito que se deja de lado es el más profundo de extensionalidad. Sin duda, una parte importante del mérito del trabajo de Kahneman y Tversky consiste en haber encontrado la manera de modelar individuos no racionales (es decir, individuos que violan extensionalidad) y al mismo tiempo, *predecibles*. Al posibilitar predicciones bien determinadas y preservar la maximización, *Prospect Theory* ha sido capaz de ir incorporándose progresivamente en un conjunto de modelos económicos más amplios, y están en la base de lo que se denomina *Behavioral Economics*. La precisión formal y el poder predictivo son características bienvenidas para la economía, y han constituido la puerta de entrada para modelos con base en sujetos psicológicamente realistas.

Las predicciones científicas están condicionadas por la naturaleza. Esta característica da lugar a la distinción tradicional (popperiana) entre *predicción* y *profecía*. Sin embargo, la manera en la que intentan predecir los modelos convencionales y *Prospect Theory* es marcadamente diferente. Mientras que en la visión tradicional la predicción se deriva del conjunto de axiomas y supuestos de racionalidad (más el principio de maximización), la predicción prospectiva descansa en la posibilidad de conocer de antemano cuál es el marco o descripción que va a ser adoptado por el sujeto. El hecho de que sea necesario conocer de

antemano de qué manera se van a enmarcar las opciones parece ser un talón de Aquiles para *Prospect Theory*. Lejos de que tal cosa suceda, las predicciones prospectivas sacan a la luz características potencialmente innovadoras vinculadas a la posibilidad de *intervenir* en las decisiones de los agentes.

4.13-Predicción convencional y predicción en *Prospect Theory*

Modelando a los individuos como decisores intensionales, *Prospect Theory* es capaz de predecir la mayoría de las anomalías conocidas de la teoría de la decisión convencional, incluidos los *framing effects*. El singular mérito de la teoría puede verse siguiendo el camino emprendido para alcanzar los resultados deseados. En agudo contraste con muchos de las propuestas “realistas” del pasado (Cfr. Cap. III), *Prospect Theory* logra introducir una variable causal –la descripción o punto de referencia- dentro de las ecuaciones de maximización.

Recordemos que en la tradición del siglo XVII la apuesta razonable –o lo que es idéntico-, la cantidad máxima que una persona racional apostaría por participar en una situación riesgosa como un juego de azar, contempla ponderar dos variables: el premio o *resultado* en juego y las *probabilidades* de obtenerlo. Esta situación queda reflejada en la ecuación básica del valor esperado:

$$VE (A): \sum R_i \cdot p_i$$

El valor esperado (VE) de la lotería A surge de operar matemáticamente sobre los resultados posibles (R_i) por las probabilidades asociadas a ellos (p_i). Estos dos son las únicas variables causales que intervienen para un apostador racional.

La modificación de Bernoulli, luego retomada por Morgenstern y Von Neumann, consiste en reemplazar los resultados por la *utilidad* que le brindaría a un sujeto:

$$UE (A): \sum U(R_i) p_i$$

La utilidad esperada (UE) de la lotería A tiene su origen en la utilidad de los resultados multiplicada por su correspondiente probabilidad. Ahora el teórico necesita conocer, además de las variables “objetivas” de la situación, la *utilidad* que el sujeto asigna a los resultados. Es por ello que las predicciones de EUT requieren observar al menos un par de elecciones previas del sujeto, con el fin de inferir la forma específica de la función de utilidad del elector (es decir, su aversión o propensión al riesgo). Pero una vez hecho, el sujeto “está atado” en sus decisiones subsecuentes (Baumol, 1952). Comparemos ahora la situación con la de *Prospect Theory*:

$$UP: \sum V(R_j) \pi (p_i)$$

La utilidad prospectiva (UP) de una lotería surge de operar sobre la Función Valor y la Función Ponderación. La primera expresa las valuaciones que los seres humanos realizan sobre los resultados y la segunda la manera en la cual las probabilidades son deformadas al percibirse. Lo que resulta central es que las valuaciones ya no se realizan sobre los “premios” o estados finales, como suponen las restantes ecuaciones. El material necesario para obtener la utilidad prospectiva es radicalmente diferente a los anteriores, porque su referente son ganancias y pérdidas con respecto a un punto de referencia. Nada de esto sucede en la doctrina tradicional, la cual siguió el derrotero impuesto por Bernoulli de asignar valores a estados finales. Hay, entonces, una tercera variable causal involucrada, que se suma a las anteriores, para determinar la decisión de los seres humanos. Podemos llamarla sucintamente la *descripción* de la situación de elección.

Resulta interesante notar que a diferencia de las anteriores, la variable introducida por *Prospect Theory* es de *naturaleza* diferente, porque no depende de la situación “objetiva” –el juego o la lotería- sino de una consideración adicional: *la descripción de la situación objetiva*.

Vemos ahora que *Prospect Theory* puede predecir la conducta individual porque ha identificado una variable causal adicional a las conocidas que es relevante en el dominio de las decisiones humanas. Tal variable determina las decisiones de los agentes junto con probabilidades y resultados. La naturaleza particular de la nueva variable explica en parte las características únicas de la nueva teoría, y los desafíos que presenta a la conceptualización epistemológica.

Algunas de las implicaciones pueden resultar difíciles de asimilar. Una primera consecuencia es la siguiente:

A diferencia de los modelos predictivos convencionales, en la cual el teórico se comporta como un verdadero espectador frente al objeto de estudio, en las predicciones de la decisión humana es necesario intervenir para asegurar el éxito predictivo.

Tanto las predicciones de la teoría convencional como las prospectivas son enunciados condicionales, en sentido lógico; pero el *antecedente* es muy distinto en uno y otro caso. Una predicción convencional parte de los axiomas de racionalidad, mientras que las predicciones prospectivas dependen de conocer el punto de referencia o descripción adoptada por el sujeto. Además, mientras que en EUT las predicciones son el resultado de una tarea puramente intelectual, que exige solamente el cálculo externo por parte del teórico –algo semejante al modelo astronómico, donde el observador se limita a realizar cálculos a partir de leyes y datos iniciales “típicos”-, las predicciones en *Prospect Theory* son posibles mediante la intrusión del teórico en el contexto de elección.

La intervención del teórico es trivialmente cierta en contextos cerrados. En el caso de la decisión humana, la situación típica de laboratorio contempla enfrentar al sujeto con opciones (loterías) que se pueden ver como una tríada de resultados, probabilidades asociadas y una cierta descripción. Retomemos una vez más el caso de *Asian Disease*. Allí la reversión de preferencias que presentaba el

carácter de una dificultad insalvable para la teoría convencional se transforma en un patrón predecible. ¿Cómo es esto posible?

Prospect Theory puede predecir exitosamente porque el teórico tiene en sus manos las tres variables causales de la elección. Conoce los resultados, las probabilidades asociadas y la *descripción* de la situación que enfrenta el sujeto (en el primer caso, el problema se plantea en términos de ganancias, en el segundo de pérdidas). En rigor, no sólo conoce la descripción, sino que él mismo es quien la ha construido. ¿Qué sucede en contextos *abiertos*? ¿Es posible predecir la conducta de los individuos cuando se sale de las condiciones restrictivas del laboratorio? La respuesta es negativa:

Los resultados experimentales conseguidos en laboratorio no son directamente aplicables para predecir la conducta en contextos abiertos, porque una de las variables causales determinante de los resultados (i.e., la descripción) no puede conocerse de antemano.

Para ilustrar el caso, veamos el siguiente ejemplo:

Problema del apostador:

“Consider a person who has spent an afternoon at the race track, has already lost \$140, and is considering a \$10 bet on a 15:1 long shot in the last race. This decision can be framed in two ways, which correspond to two natural reference points. If the status quo is the reference point, the outcomes of the bet are framed as a gain of \$140 and a loss of \$10. On the other hand, it may be more natural to view the present state as a loss of \$140, for the betting day, and accordingly frame the last bet as a chance to return to the reference point or to increase the loss to \$150.” (Tversky y Kahneman, 1981, p. 456)

Claramente sus opciones son apostar o no apostar, pero estas dos alternativas pueden ser descritas de varias maneras distintas:

- *Caso 1. El apostador no toma en cuenta las pérdidas del día. Su punto de referencia en este caso es 0\$, y sus opciones son las siguientes:*

A: No apostar $(A) = 0$

B: Apostar $(B) = (-\$10, 14/15; +\$140, 1/15)$

- *Caso 2. El apostador toma en cuenta las pérdidas del día. Su punto de referencia es ahora -\$140, y sus opciones cambian conjuntamente:*

C: No apostar $C: (-\$140)$

D: Apostar $D: (-\$150, 14/15; \$0, 1/15)$

¿Es posible, para un observador externo, predecir si nuestro hombre va a apostar o no? La respuesta es negativa.⁴⁷ No puede hacérselo, una vez más, porque su decisión está causada en parte por una variable que no depende de las circunstancias “objetivas” de la situación –resultados y probabilidades-, sino que depende de la *descripción* que el sujeto asuma como significativa. ¿Considerará el apostador sus pérdidas pasadas? En tal caso, *Prospect Theory* predice que va a apostar. ¿Verá cada apuesta como un evento independiente? En ese caso, la teoría arroja como predicción que nuestro sujeto no va a apostar. Pero en la medida que no puede saberse de qué forma efectivamente va a enmarcar la situación, no hay manera, sin intervención directa del teórico, en que pueda conocerse de antemano la elección individual. La predicción solamente puede ser exitosa en laboratorio, y si puede extenderse esa capacidad predictiva a la realidad es en la medida en que se tenga algún poder sobre las descripciones que enfrenta el sujeto.

Esto nos lleva a considerar una diferencia importante entre la predicción convencional y las predicciones prospectivas: el componente *tecnológico* de la

⁴⁷ Las consecuencias de los *framing effects* han sido desarrolladas en Marqués-Weisman (2009)

segunda. El hecho de que las decisiones de los agentes se puedan predecir a través de las manipulaciones es un notable resultado. Dejando de lado por el momento las consideraciones relativas al poder social o la libertad individual, es la capacidad para manipular las decisiones fuera del laboratorio lo que hace al Marketing posible como disciplina. La innovación de *Prospect Theory* no reside en este caso en producir nuevos fenómenos, sino en iluminar los existentes bajo una nueva luz, unificando la comprensión de fenómenos que hasta entonces estaban apartados. El fenómeno de la publicidad, por ejemplo, que en la concepción tradicional no tiene otra función que hacer “público” el conocimiento de bienes y servicios, puede ahora ser conceptualizada e incorporada de una manera menos desesperadamente ingenua dentro del marco general de toma de decisión. Sin embargo, los efectos sobre otros fenómenos son algo más complejos e inquietantes: el *homo intensionalis* parece ser demasiado débil para cargar sobre sí el peso de la competencia ciudadana, base del funcionamiento teórico de las democracias. En el Capítulo VII veremos algunas de tales implicaciones, a las que podemos llamar *indirectas*, sobre las encuestas y la democracia.

¿Cuáles son, hasta ahora, las diferencias entre *Expected Utility Theory* y *Prospect Theory*? Un gráfico puede ayudar a sistematizar la información.

	EUT	PT
Concepto de Utilidad	Formal	Sustantivo
Portadores de Utilidad	Estados Finales	Cambios
Individuos	Extensionales	Intensionales

SECCIÓN II

Análisis Filosófico y epistemológico de Prospect Theory

Capítulo V

Características del cambio teórico involucrado en el pasaje de EUT a *Prospect Theory*

5.1-Introducción

¿Cuál es la relación de *Prospect Theory* con EUT? ¿Qué clase de cambio teórico es el que describe mejor el paso de una a otra teoría? En este capítulo se sostendrá que el cambio es de naturaleza radical, en contra de la visión habitual según la cual habría una relación de *continuidad* en el pasaje de una a otra teoría.

Suele afirmarse que *Prospect Theory* guarda una relación particular con EUT, en el sentido de que ésta sirve como *benchmark* para contrastar los modelos psicológicos de decisión económica (Camerer *et al*, 2003) Lo anterior sugiere una relación de *continuidad* entre ambas teorías. Coincidiendo con esta interpretación, Starmer (2003) establece que *Expected Utility Theory* es un caso límite de *Prospect Theory*, a la cual podría ser reducida bajo ciertas circunstancias. Schoemaker (2000) señala una divergencia más profunda, al apuntar a la pérdida de transitividad como característica diferencial de *Prospect Theory*; pero su distinción juega un rol meramente taxonómico y no se involucra con la naturaleza epistemológica de la ruptura entre ambas teorías. En contra de estas tesis, y siguiendo a Marqués y Weisman (2008) se defenderá que la relación conflictiva entre ambas teorías es mejor comprendida como un *cambio revolucionario*, debido a las mutaciones en la referencia experimentadas por ciertas variables fundamentales, como el término “utilidad”, el cual posee propiedades muy distintas en ambas teorías. Si esto fuera cierto, cabría esperar una relación de inconmensurabilidad del tipo Kuhniano en el pasaje de *Expected Utility Theory* a *Prospect Theory*.

En EUT y en muchas de sus versiones convencionales, la *utilidad* esperada es una propiedad de las loterías, y se obtiene a partir de la *utilidad* asignada a los resultados que conforman la lotería. Dados dos resultados o loterías X e Y el agente convencional podría ordenarlos de una manera bien definida, y los números o símbolos que permiten realizar el ranking de las opciones son la *utilidad* que él extrae de ellas. Como hemos visto, esta característica de la teoría

convencional depende crucialmente de la *extensionalidad*. Es este principio lógico el que permite postular equivalencias entre loterías, y alternativamente el que habilita a hacer lo mismo con los resultados. Los agentes que respetan tal principio, y que por lo tanto pueden identificar dos loterías como *la misma* a pesar de que aparezca bajo diferentes descripciones, son decisores *extensionales* (Marqués y Weisman, 2009)

En la medida que los seres humanos son decisores *intensionales* sus elecciones no revelan un ranking de preferencias bien definido: el mismo conjunto de elección dispara dos órdenes distintos cuando las opciones se describen de manera distinta en términos de ganancias o pérdidas. En tal situación, *la utilidad en el sentido convencional ya no es válida en el dominio de los seres humanos* (y por ende, ya no es posible maximizarla) En un dominio en el cual la extensionalidad ha perdido validez, la noción de utilidad convencional debe ser resignada.

Prospect Theory introduce un cambio mayúsculo con esta posición al postular individuos con preferencias (y por lo tanto, funciones de utilidad) definidas sobre *descripciones de loterías* en lugar de loterías. Y tal idea ha sido incorporada en funciones y ecuaciones capaces de representar un orden único de preferencias el cual puede ser maximizado, incluso bajo la presencia de *framing effects*. Siguiendo la tradición sería perfectamente posible llamar a estos nuevos números “utilidad”, pero es importante tener en mente que esta clase de utilidad es una noción muy diferente respecto a aquella presente en la teoría de la decisión convencional: refiere a descripciones de cambios de resultados (ponderados por su probabilidad) relativos a cierto punto de preferencia. A pesar de que en algunas ocasiones se le otorga un nombre especial (por ejemplo, en Barberis y Thaler, 2003, se la denomina “utilidad prospectiva”) prevalece la utilización indistinta del término, sin tomar en cuenta la transformación experimentada.

Un cambio conceptual como el descrito, que afecta al concepto de utilidad, invita a una elaboración metodológica. Trae a la mente la descripción kuhneana⁴⁸ de la transformación del concepto “masa” cuando se exporta del marco teórico de Newton al marco teórico de Einstein. Podemos sacar provecho del análisis provisto por Kuhn (1970) respecto al cambio teórico para echar luz sobre la aparición de *Prospect Theory* y la relación que podría guardar con *Expected Utility Theory*.

Thomas Kuhn reduce a los fenómenos que desencadena el surgimiento de nuevas teorías a tres clases (*Op.cit*, p. 98). En primer lugar, hay *fenómenos ya explicados* por las teorías aceptadas, y pueden aparecer nuevas teorías que compitan por explicarlos. Pero esta situación por lo general no origina un cambio teórico de magnitud, porque no hay manera de discriminar sobre bases empíricas entre las alternativas planteadas. La historia de la ciencia muestra el frío desinterés con el cual son recibidos tales aportes. La segunda clase la constituyen *fenómenos esperados*, es decir fenómeno cuya naturaleza es señalada por la teoría aceptada, pero del cual se desconocen los detalles, situación que dispara en la comunidad científica los incentivos adecuados para poner en marcha la articulación teórica. La situación culmina cuando se logra *articular* el paradigma, generando una nueva teoría compatible con el hecho. En este proceso puede surgir la tercera clase de fenómenos, las *anomalías*, problemas que se niegan obstinadamente a entrar en los límites paradigmáticos. En la visión de Kuhn, solo este tercer tipo de fenómenos es capaz de desencadenar teorías *radicalmente* nuevas.

La aplicación de esta tipología al surgimiento de *Prospect Theory* no está exenta de dificultades. En primer lugar, el fenómeno de *preference reversals* claramente puede ubicarse como una *anomalía* en sentido kuhniano. No es un fenómeno

⁴⁸ Los problemas del término “paradigma”, ya reconocidos por el mismo Kuhn, limitan su utilidad como instrumento de análisis cuando se trata de aplicarlo en casos puntuales. En esta sección nos restringiremos en la medida de lo posible al término “teoría”, más aceptado y de límites menos imprecisos.

explicado ni esperado por la teoría convencional, y la historia relativamente reciente esbozada en el Capítulo 3 muestra a Slovic y Liechtenstein poseyendo un grado de conciencia sorprendente en relación a este punto. Ambos sabían que estaban ante un fenómeno anómalo y potencialmente disruptivo para la economía *mainstream*. Por otra parte, los esfuerzos de Grether y Plott por demostrar primero que el fenómeno era inexistente, luego que aunque existía podía ser explicado y por lo tanto reducido a la teoría convencional, y finalmente aceptando su condición de inexplicable (y dictaminando que en última instancia era irrelevante para la teoría de la decisión establecida) dan cuenta de una defensa de la teoría reinante de libro de texto. Sin embargo, debe notarse que la anomalía de *preference reversals* no surgió *dentro* de la Teoría de la Decisión de la Economía. A pesar de lo difuso del término “paradigma” (Kuhn, 1970), y sea cual sea su valor como herramienta descriptiva, cabría esperar que la Teoría de la Decisión en Economía y en Psicología Cognitiva sean, o pertenezcan a paradigmas separados. ¿Cómo se explica que una disciplina sea quien postule una anomalía para *otra* disciplina?

5.2-El acercamiento entre Psicología Cognitiva y Teoría de la Decisión Racional

Quizás una mirada a la historia reciente de las dos disciplinas arroje luz sobre la cuestión. Hacia los años 60' comenzó a popularizarse la llamada “metáfora de la computadora” en el cognitivismo y la psicología cognitiva (Camerer, 2002). Hasta entonces prevalecía el enfoque del comportamiento tipo estímulo-respuesta, donde el cerebro era modelado como una “caja negra”, a la manera instrumentalista⁴⁹. Algo similar ocurría en la Teoría de la Decisión Racional en

⁴⁹ Resulta interesante notar que Von Neumann, el mismo que ya había realizado aportes a la Teoría de la Decisión Racional brindando una axiomatización de la Hipótesis de la Utilidad Esperada, es también uno de los pioneros en señalar sugestivas analogías (y diferencias) entre el funcionamiento de la “computer machine” y el cerebro humano. Sus aportes fueron publicados de manera póstuma en 1958 (The computer and the brain. New Haven, CT: Yale University Press)

economía. El *proceso* de toma de decisión no era considerado relevante⁵⁰: la racionalidad se reducía a consistencia, y lo que es más importante se contaba con un dispositivo disciplinar que por un lado, permitía dejar afuera los procesos *individuales* de toma de decisión, y por el otro garantizaba una clase de éxito predictivo particular. Es el que hemos denominado uso *idealizante* de EUT (Cfr. Cap II). En la tradición neoclásica, el *mercado* asegura condiciones de competencia que penaliza las decisiones erróneas o irracionales, de tal manera que al final del día sólo queden en pie aquellos individuos que se comportaron *como si* hubieran tomado decisiones racionales. Este es un dispositivo característico de la economía, de carácter singular y poco tematizado dentro de la epistemología tradicional (Weisman, 2009). Más adelante diremos unas palabras al respecto. Provistos de tal mecanismo externo, los economistas interesados en Teoría de la Decisión no necesitan averiguar los mecanismos internos “reales” por los cuales los individuos efectivamente toman decisiones. Basta con conocer de antemano los resultados obtenidos por aquellos individuos *ideales*, plenamente racionales, los únicos capaces de sobrevivir a largo plazo en condiciones de libre competencia. Para conocer de antemano tales resultados, sólo basta con operar dentro del modelo. Esto explica por qué *Expected Utility Theory* puede ser interpretada como una teoría *normativa* (identifica la decisión racional) y además *descriptiva* (describe el resultado final de mercado)

Lo anterior no descarta que EUT también pudiera ser interpretada como una teoría descriptiva del *proceso* de toma de decisión racional. Savage y Raiffa parecen haber estado pensando en esta posibilidad, al igual que Simon. De hecho el interés de Simon pasa por encontrar un *proceso* racional de toma de decisión, algo que es conocido en la literatura relevante como *Procedural Rationality* (Marqués, 2010). Infortunadamente sus intentos no encontraron eco en la Economía *mainstream*, quizás por las razones señaladas en el Capítulo III: el

⁵⁰ En rigor, había quienes defendían la interpretación *procedural* de EUT, pero su posición parece haber sido minoritaria.

modelo construido resultó ser vacío de contenido empírico, y en rigor ninguna predicción puede obtenerse a partir de él (Marqués y Weisman, 2008) Podemos llamar al subgrupo de economistas interesados en los procesos reales de toma de decisión individual, teóricos *procedurales* de la decisión.

La existencia del mecanismo *externo* disciplinar, y las limitaciones de los modelos interesados en los mecanismos *internos*, unieron fuerzas para cerrarle la puerta al interés por los procesos de toma de decisión “reales” por parte de una parte importante de la profesión económica. Incluso en la actualidad aquellos economistas que sólo conciben el mecanismo externo de la competencia como el único relevante (y definitorio de la disciplina), no tienen dificultades para desdeñar los descubrimientos psicológicos como interesantes, pero ampliamente irrelevantes. Llamaremos a este grupo teóricos *substantivos* de la decisión. La experimentación sobre individuos “reales” resulta superflua para los teóricos substantivos, porque lo único que importa es modelar la conducta de un único individuo: el individuo *ideal*. El proceso de mercado terminará por hacerlo *real*, y en este peculiar sentido la modelización irrealista posee éxito predictivo garantizado. No existe en tal esquema lugar para investigar los procesos de toma de decisión reales.

Mientras tanto, en el ámbito de la psicología la metáfora de la computadora permitió pensar el cerebro como un conjunto de *procesos*, y despertó un creciente interés por fenómenos como la memoria, la evaluación, la percepción y finalmente la toma de decisiones. Esto explica que hacia la década del '70 se perciba un interés común por parte de psicólogos cognitivos por un lado, y economistas *mainstream* por el otro en el ámbito de la Teoría de la Decisión, si bien con enfoques diferentes. Como era de esperarse, Paul Slovic y Sarah Liechtenstein descubrieron la reversión de preferencias a partir de experimentos sobre individuos reales:

“The impetus for this study was our observation in our earlier 1968 article that choices among pairs of gambles appeared to be influenced primarily by probabilities of winning and losing, whereas buying and selling prices were primarily determined by the dollar amounts that could be won and lost (...)” (1983, p. 597)

El hecho de que el proceso para tomar decisiones sea distinto al proceso para poner precios en los individuos *reales* no es una anomalía para la Psicología Cognitiva, simplemente un descubrimiento interesante. Dentro del campo de la Economía *mainstream* de los años 70', sin embargo, el mismo hecho constituía claramente una anomalía para quienes consideraban que EUT describía de manera adecuada los *procesos* de toma de decisión de los individuos reales, es decir para los teóricos *procedurales*. Pero para el grupo de teóricos *substantivos*, cuyos miembros consideraban el *mecanismo externo* como único campo de interés legítimo para la disciplina, el descubrimiento simplemente carecía de interés. Esto explica la recepción relativamente tibia que tuvieron las anomalías, aun para quienes reconocían su importancia (Grether y Plott, 1979), y el profundo interés que despertó en los teóricos *procedurales*, los cuales coincidían con Simon en que la forma de mejorar la economía pasaba por construir modelos más “realistas” a nivel de los sujetos y procesos de toma de decisión.

Kuhn ha señalado que las categorías epistemológicas se vuelven muchas veces porosas o abstractas en demasía cuando se las quiere aplicar a la realidad (1970, Prefacio). Las categorías epistemológicas de Kuhn no son la excepción. La anomalía de *preference reversals* no parece haber surgido de un intento fallido de articulación intraparadigmático, sino más bien de un complejo acercamiento progresivo entre tres actores (teóricos *substantivos*, *procedurales* y psicólogos cognitivos) pertenecientes a dos disciplinas distintas que terminó por redistribuir el campo teórico y dejarlo en manos de una nueva propuesta, *Prospect Theory*. Sin duda, el surgimiento de *Prospect Theory* puede explicarse a partir de los

intentos fallidos por resolver una anomalía, como Kuhn establece, pero incluso en esbozo de reconstrucción histórica como el presente revela un proceso bastante menos lineal al sugerido.

5.3-EUT como caso especial de *Prospect Theory*

En la actualidad está ampliamente difundida la idea de que *Prospect Theory* es compatible con *Expected Utility Theory*. El mismo Kahneman le reserva a EUT la función de una “hipótesis nula” (2002, p. 1449), y más explícitamente Starmer (2003) afirma que EUT es un caso especial de *Prospect Theory*, a la cual se reduce en ciertas condiciones.

El argumento es paralelo al utilizado en ocasión de la dinámica Newtoniana como caso especial de la dinámica relativista (Kuhn, 1970) Podemos llamar a esta estrategia “reducción del campo de aplicación”, por razones que serán claras en un momento. De acuerdo a esta postura, cuando una teoría T1 enfrenta una anomalía y da ocasión al surgimiento de una nueva teoría T2, lo que se revela no es la *falsedad* de la primera, sino su *inaplicabilidad* a un campo determinado. Luego de la resolución de la anomalía, ambas teorías se conservan, cada una en un dominio de aplicación específico. Al final del proceso T1 ha experimentado una reducción en su campo de aplicación, y ha cedido terreno a T2. No hay aquí cambio revolucionario, porque ambas teorías se conservan. Más aun, T1 es un *caso especial* de T2, lo que quiere decir que bajo ciertas condiciones iniciales C1, T2 se reduce a T1.

La estrategia ha sido denunciada como un resabio del positivismo por parte de Thomas Kuhn, entre otras razones porque niega el cambio revolucionario, un fenómeno constitutivo del progreso científico. Podemos reconstruir el argumento de manera breve. La estrategia positivista encuentra apoyo en la distinción entre *testeo* y *aplicación* (Bunge, 1970), o contexto de validez y contexto de aplicación. En el primero se decide si una conjetura es *válida*, es decir si las

variables se relacionan de la manera en que se ha postulado. Dado que las teorías poseen supuestos iniciales de validez muy restrictivos (suelen realizar afirmaciones en condiciones extremas como existencia de vacío, rozamiento nulo, competencia perfecta, etc) la forma adecuada de testearlas es dentro de un laboratorio, o bien con herramientas matemáticas, construyendo modelos. Una vez que ha finalizado la función del contexto de validez, suele discutirse la *aplicabilidad* de la teoría a casos fuera de las restrictivas condiciones de partida. Conviene notar que son problemas muy diferentes. Una teoría puede ser falsa y ampliamente aplicable (v.gr., la astronomía ptolemaica), o bien válida sólo en laboratorio, o en el modelo. O también puede experimentar en ciertos casos una especie de crecimiento virtuoso, comenzando con éxito en el laboratorio y continuando luego una suerte de colonización progresiva de nuevos campos de aplicación.

El proceso de ampliación del campo de aplicación es ampliamente conjetural, en el sentido que se ponen a prueba nuevas hipótesis, del tipo:

H1: “La Teoría T1 es aplicable al dominio D1”

H2: “La Teoría T1 es aplicable al dominio D2”

H3: “La Teoría T1...

A medida que estas *hipótesis de aplicación* son puestas a prueba exitosamente se va ampliando el dominio $D = (D1, D2, \dots)$ en el cual vale la teoría. Una falla en una de la hipótesis de aplicación no compromete desde el punto de vista lógico a las otras, porque son hipótesis diferentes. Tampoco compromete el test realizado en laboratorio, el cual es independiente además desde el punto de vista lógico (Weisman, 2010)

Sin embargo, en ciertas ocasiones la refutación de una hipótesis de aplicación puede dar origen a una teoría alternativa T2, *exitosa* en el campo específico donde T1 ha fallado. En tales circunstancias, T1 no pierde el terreno ya

conquistado. El resultado de la refutación es la coexistencia de la teoría original, válida todavía en su campo de aplicación, y una nueva teoría T2, la cual vale en un contexto de aplicación diferente.

Aplicado al caso de la dinámica de Newton, la estrategia permite afirmar que es *compatible* con la de Einstein. Los contraejemplos encontrados no muestran la naturaleza errónea de la dinámica de Newton *in toto*, sino que señalan más bien la falsedad de la hipótesis de aplicación particular cuyo dominio está constituido por casos de alta velocidad relativa. Es la *aplicación* al dominio de altas velocidades relativas la que se demostró inválida, no la teoría newtoniana en sí ni su aplicabilidad a dominios previos. Esto explicaría que siga siendo utilizada por físicos e ingenieros en una multitud de circunstancias, justamente aquellas que la teoría de Einstein señalan como de bajas velocidades relativas. En consecuencia la teoría relativista no anula ni invalida la teoría de Newton. Antes bien, la teoría de Newton es un *caso especial* de la de Einstein, la cual se reduce a aquélla en ciertas condiciones (*i.e.*, condiciones de baja velocidad relativa).

De acuerdo a la posición convencional, algo semejante sucede en el terreno de la Teoría de la Decisión con el surgimiento de *Prospect Theory*. La nueva teoría no anula la anterior, ni la invalida. Simplemente muestra que existe un dominio de aplicación particular (*i.e.*, aquellos donde predominan los *framing effects*) donde la teoría falla.

5.4-Inconmensurabilidad y Cambio Teórico

Desde el punto de vista de Kuhn, la estrategia de reducir el campo de aplicación para salvar teorías presenta inconvenientes. La teoría de Newton es *incompatible* con la de Einstein, en el mismo sentido que la astronomía Ptolemaica es incompatible con la Copernicana: sólo puede aceptarse la dinámica de Einstein reconociendo que la de Newton estaba equivocada (Kuhn, *op.cit.*, p 101) Para

defender su postura acude a varios argumentos pragmáticos, pero el apoyo más eficaz proviene de su conocido argumento lógico.

El argumento está construido para señalar que la teoría de Einstein no es reducible a la de Newton, porque al derivar un conjunto de ecuaciones “newtonianas” N_i a partir de las de Einstein (E_i), se están utilizando términos homófonos con cargas referenciales diferentes:

“(...) la derivación es ilegítima, al menos hasta este punto. Aunque el conjunto N_i es un caso especial de las leyes de la mecánica relativista, no son las leyes de Newton. (...) Las variables y parámetros que en la serie einsteiniana E_i representaban la posición espacial, el tiempo, la masa, etc., se presentan todavía en N_i ; y continúan representando allí espacio, tiempo y masa einsteinianos. Pero las referencias físicas de esos conceptos einsteinianos no son de ninguna manera idénticos a las de los conceptos newtonianos que llevan el mismo nombre. (La masa newtoniana se conserva; la einsteiniana es transformable por medio de la energía. Sólo a bajas velocidades relativas pueden medirse ambas del mismo modo e, incluso en ese caso, no deben ser consideradas idénticas). A menos que cambiemos las definiciones de las variables en N_i los enunciados derivados no serán newtonianos. Si las cambiamos, no podremos de manera apropiada decir que hemos derivado las leyes de Newton, al menos no en cualquiera de los sentidos que se le reconocen actualmente al verbo "derivar".” (1970, p. 163)

Las definiciones de los términos han cambiado en el pasaje de una teoría a otra. Esto implica que las teorías no son solo *incompatibles*, sino que además son *inconmensurables*. Más importante, se puede extender el argumento contra las pretensiones reduccionistas en Teoría de la Decisión. Esto es así porque también las propiedades características del término *utilidad* cambian en el pasaje de *Expected Utility Theory* a *Prospect Theory*. El concepto de utilidad en EUT es un índice de preferencias, no se compromete con el contenido hedónico, se predica

de estados finales, y es medible; el término homófono en *Prospect Theory* posee contenido sustantivo, refiere a experiencias hedónicas, está asociado a cambios de estado y no hay hasta el momento un procedimiento para medirlo. Hay un cambio *radical* en el contenido semántico del término, lo que arroja una sombra de duda sobre la posibilidad de derivar una de otra.

Una comparación con la manera en la cual *Prospect Theory* “generaliza” las ecuaciones de *Expected Utility Theory* puede ayudar a entender la cuestión. Consideremos la Hipótesis de Utilidad Esperada:

$$\text{HUE: } UE(A): \sum U(R_i) p_i$$

La ecuación (1) de *Prospect Theory* “generaliza” HUE:

$$(1) \sum V(R_i) \pi(p_i)$$

La ecuación (1) cambia de una manera irreducible la noción de utilidad que está presente en HUE: $V(x)$ no es $U(x)$. Si la ecuación 1 pudiera ser reducida a HUE, la utilidad prospectiva $V(x)$ debería poder ser reducida a la utilidad convencional $U(x)$. ¿Es posible esto? En la medida que el cambio conceptual de una a otra se mantenga, la respuesta parece ser negativa.

Consideremos ahora la relación entre las dos ecuaciones de *Prospect Theory*. En este caso, Kahneman y Tversky también afirman que existe una generalización (y por lo tanto la reducción es posible). De hecho, la ecuación (2) puede ser reducida a (1) dadas algunas restricciones. Pero el *concepto de utilidad* incorporado en ambas ecuaciones es el mismo. Esta reducción, en contraste con la mencionada más arriba, no supone problemas relacionados con la inconmensurabilidad. La primera clase de generalización (o alternativamente, reducción) es filosóficamente problemática, a pesar de que hasta el momento no ha recibido suficiente atención por parte de la comunidad de economistas.

Las dificultades experimentadas por los economistas *mainstream* y heterodoxos en distintos debates respecto al concepto de utilidad, u otros puntos críticos implicados por esta noción, podrían tener su origen en que un fenómeno de inconmensurabilidad entorpece el diálogo disciplinar. Si este diagnóstico es correcto, una gran cantidad de trabajo filosófico debería llevarse a cabo para mejorar la conversación en torno a este punto, y quizás arribar a un consenso en relación a los cambios semánticos del término utilidad.

5.5-La reacción disciplinar

En la sección precedente se ha argumentado que el concepto de utilidad cambia de manera radical en el pasaje de una teoría a otra, lo cual podría estar relacionado con inconvenientes propios del fenómeno de inconmensurabilidad. Cabría distinguir tres tipos de efectos desplegados por el cambio conceptual: un impacto directo sobre el campo de la Teoría de la Decisión, uno indirecto ejercido sobre la Teoría Económica que lo incluye, y un efecto derrame sobre disciplinas colindantes y quizás alejadas, pero con las que existen conexiones conceptuales.

En relación al impacto indirecto ¿pueden ser extendidos los resultados anteriores desde el ámbito de la Teoría de la Decisión a la Economía? En la medida que en la base de una teoría económica se encuentra una teoría de la decisión, podría pensarse que una teoría económica que tenga pretensiones de ser exitosa desde el punto de vista descriptivo debería modelar a los agentes siguiendo a *Prospect Theory* (es decir, como agentes *intensionales*). Este es el camino seguido por *Behavioral Economics*. Sin embargo, la situación no es tan simple. La economía convencional no puede ser descartada sin más. En primer lugar, no puede ser excluida de antemano la posibilidad de que un conjunto relevante de fenómenos pudiera resolverse a partir de la economía basada en decisores racionales. Además, incluso en aquellos dominios en los cuales *Behavioral Economics*

reclama pertinencia, el acercamiento convencional podría tener algún papel después de todo. Como Barberis y Thaler (2002, p. 1076) reconoce, “*both the rational approach and behavioral approaches to finance have made progress in understanding the (...) puzzles [related with the equity premium puzzle]*”

En definitiva, una mirada prudente en relación a este punto es reconocer que el problema de cuáles fenómenos pueden resolverse mediante la visión tradicional y cuáles con la visión behaviorista, no puede zanjarse de antemano: tiene que ser decidido caso por caso. Dado que los términos teóricos básicos usados en cada una tipo de explicación son irreducibles unos a otros, y en la medida en que la coexistencia de ambas visiones persista, el futuro cercano de la teoría económica parece ser la fragmentación antes que la unidad. En el futuro podría esperarse que los modelos económicos de agentes racionales y los modelos behavioristas tendrán sus propios conjuntos de fenómenos explicados, utilizando nociones de “utilidad” diferentes y agentes extensionales e intensionales respectivamente.

Sin embargo, es difícil reconciliar este resultado con la visión que algunos economistas de *Behavioral Economics* tienen acerca de la naturaleza de su empresa y de los vínculos que los atan con la economía tradicional, cuando avanzan la tesis según la cual *Behavioral Economics* es una *generalización* de la economía convencional. Se afirma que una teoría T' es una generalización de otra teoría T cuando T' preserva los parámetros presentes en T, y además incorpora parámetros adicionales⁵¹. La idea según la cual las ecuaciones de *Behavioral Economics* se reducen a las de la teoría convencional cuando los parámetros adicionales se hacen cero, supone nuevamente una visión del cambio científico acumulativa. Precisamente da por sentado que se preserva la referencia cuando la utilidad en el sentido convencional es reemplazada por la utilidad en el

⁵¹ “Theories in behavioral economics also strive for generality –e.g., by adding one or two parameters to standard models. Particular parameter values then often reduce the behavioral model to the standard one...” (Camerer and Loewenstein, 2002, p. 2).

nuevo sentido, y viceversa. Si, por el contrario, algún tipo de inconmensurabilidad está involucrada en la relación entre ambos conceptos, hemos visto que no puede afirmarse sin más que el nuevo concepto de utilidad colapse con el primero cuando los parámetros adicionados asumen un valor nulo.

5.6-Cambio de lógica subyacente

En esta sección mostraremos un argumento adicional a favor de la naturaleza disruptiva del cambio teórico en consideración, vinculado con el posible *cambio en la lógica subyacente* experimentado al pasar de una a otra teoría. Esto nos permitirá analizar más adelante los posibles efectos de derrame que han impactado sobre disciplinas colindantes. Hemos visto que la anomalía de *preference reversals* conduce a un cambio teórico signado por modificaciones de peso en conceptos centrales de la Teoría de la Decisión, los cuales se extienden a través de la economía convencional y su relación con *Behavioral Economics*. Tales efectos se verifican sobre los puntos de contacto inmediatos, pero es posible que se encuentren además reverberaciones perceptibles sobre sectores más alejados, aunque conectados conceptualmente con el epicentro. En los Capítulos siguientes veremos algunos cambios en la naturaleza de la intervención en Economía e incluso en conceptos centrales de Teoría Política surgidos a raíz de la aparición de *Prospect Theory*. Estos efectos de derrame sugieren la existencia de conexiones *entre* las disciplinas más amplias a las sugeridas por el holismo de Kuhn. Para sentar las bases de su análisis resulta conveniente trascender la matriz disciplinar y asumir un holismo más amplio⁵², que considere las distintas disciplinas como un todo relacionado.

De acuerdo con Quine (1953) la totalidad del conocimiento “desde las más casuales cuestiones de la geografía y la historia hasta las más profundas leyes de

⁵² Ulises Moulinés (1985)

la física atómica o incluso de las matemáticas o de la lógica pura” puede concebirse como una red conceptual que tiene contacto con la experiencia sólo en sus límites más externos. Cuando una teoría enfrenta contraejemplos, la consistencia con la realidad puede ser recuperada de distintas maneras. Debido a cuestiones pragmáticas, los científicos inicialmente intentan modificar los enunciados periféricos de la red conceptual; pero ningún enunciado está a salvo de la revisión, ni siquiera aquellos que pertenecen al núcleo de las teorías centrales o incluso los supuestos lógicos del entramado teórico.

Esto quiere decir que, en última instancia, el proceso de cambio conceptual en el que se embarca una disciplina podría ser tan radical como para afectar a las leyes más profundamente encarnadas del entramado científico, lo que incluye a las llamadas leyes lógicas. Ellas no pueden reclamar para sí mismas un status diferencial o un tipo de validez distinto al de las leyes de la periferia: ambos tipos de enunciados son *sintéticos* en la medida en que son modificables a la luz de la experiencia.

Quine no habla sólo de una situación hipotética; cita además ejemplos históricos concretos. Uno de ellos es el de la mecánica cuántica, cuya simplificación requirió la modificación del principio del tercero excluido o, según otras reconstrucciones, del principio de distribución. (*Loc, cit*, p. 78). En lo que sigue se defenderá que el cambio de EUT a *Prospect Theory* constituye un ejemplo adicional, esta vez en el campo de las ciencias sociales, de la posibilidad sugerida por Quine (1953)

5.7-Quine y el cambio teórico

Para comprender mejor la naturaleza específica de los cambios que la existencia (y persistencia) de la reversión de preferencias originada por *framing effects* han generado en la teoría de la decisión, es relevante comenzar reconstruyendo de forma simplificada la manera en que Quine concebía el cambio teórico a la luz de evidencia empírica adversa. Así como los nodos de la *red conceptual* presentan

una disposición espacial particular con respecto a la experiencia, desde una perspectiva epistemológica las *teorías* pueden ser concebidas como un entramado de enunciados dispuestos en tres niveles:

(1) Axiomas periféricos

(2) Axiomas centrales

(3) Supuestos implícitos en la formulación de los axiomas, por ejemplo supuestos lógicos y ontológicos.

Hemos visto que Quine sostenía que en presencia de anomalías podía intentarse cualquier modificación en la red teórica que eliminara el problema. “...una experiencia imprevista puede acomodarse en el sistema mediante una de varias nuevas valoraciones posibles en otros tantos sectores del sistema”. Pero nuestra natural tendencia a perturbar lo menos posible el sistema nos lleva a centrar la atención en enunciados de la periferia.

“Por eso se tiene la sensación de que esos enunciados tienen una referencia empírica más precisa que los muy teóricos enunciados de la física, de la lógica o de la ontología. Puede considerarse que éstos están situados en una zona relativamente central de la red, lo que significa meramente que presentan poca conexión preferencial con algún dato sensible determinado”⁵³.

Algo semejante ocurre a nivel de las teorías. Si la modificación de alguno de los axiomas periféricos no da resultado se procede entonces a cambiar alguno de sus axiomas centrales. Y si esto tampoco resulta puede que sea necesario introducir una transformación más radical que afecte a alguno de sus supuestos básicos.

La elección de cuáles enunciados modificar en presencia de una anomalía es una cuestión pragmática, debido a que el sistema está subdeterminado respecto de la

⁵³ Quine (1984), pp. 49 y ss.

experiencia; “ninguna experiencia concreta y particular está ligada directamente con un enunciado concreto y particular...”. En el caso de los fenómenos cuánticos, Quine pensaba que la teoría cuántica había efectuado un cambio radical respecto de la visión tradicional al nivel de su lógica presupuesta. En la lógica subyacente a la teoría cuántica se modifica una regla de inferencia lógica: el principio de tercero excluido.

Lo dicho sirve como punto de comparación para analizar el cambio efectuado por *Prospect Theory* en teoría de la decisión. Parece valer también en este caso que el *locus* donde efectuar los cambios está influido por consideraciones de economía, lo que otorga cierta discrecionalidad pragmática acerca de qué modificar.

Desde el punto de vista de EUT, la reversión de preferencias es un fenómeno sorprendente y disruptivo, y su existencia demanda la revisión de alguna de las premisas. Los *axiomas* centrales y periféricos concentran como es habitual todo el contenido informativo, y los *supuestos* incorporan conocimiento pre existente de otras teorías o disciplinas cuya validez se da por garantizada. El problema es cuál de ellas es la culpable. En el Capítulo II llamamos a este “el problema de la flecha del *modus tollens*”, y podemos ahora reconstruirlo brevemente. La siguiente es una enumeración simplificada de los componentes de cada categoría⁵⁴:

⁵⁴ Hay distintas versiones de los axiomas. Para una enumeración completa, véase Starmer (2000)

Axiomas y Supuestos de EUT

Axiomas	Supuestos
•Transitividad	• <i>Homo extensionalis</i>
•Completitud	
•Continuidad	
•Independencia	

Hemos reemplazado los supuestos de invariancia descriptiva y procedural por una caracterización más sugestiva: ambos supuestos implican que ni los procedimientos para sacarlas a la luz ni las descripciones de los resultados alteran las preferencias de los individuos. Esto implica que los seres humanos toman decisiones motivados únicamente por el contenido “objetivo” de las loterías, *i.e.* probabilidades y resultados. En el Capítulo IV llamamos a los individuos cuyas decisiones siguen este patrón, *homo extensionalis*, porque sus decisiones respetan el Principio de Extensionalidad: dos loterías con idéntica extensión pero distinta intensidad son reconocidas como la *misma* lotería y por ende su lugar en la escala de preferencia no cambia. EUT supone además el conocimiento matemático, la lógica estándar, etc.

El surgimiento temprano de los contraejemplos tradicionales –anomalías y paradojas– representó una amenaza controlada para la Teoría de la Decisión Racional, en la medida en que pudieron dirigirse exitosamente a algunos de los *axiomas* de EUT. Por esta razón los sucesivos desarrollos de la teoría intentaron modificar alguno de ellos, con el objetivo puesto en una reconciliación entre la teoría y la conducta observada. No todos los axiomas de EUT fueron considerados igualmente importantes. Prevalcía un consenso relativo a que el núcleo de EUT estaba compuesto por los axiomas de continuidad, transitividad y completitud. Juntos ellos formaban los que usualmente se denomina *preferencias bien formadas* (Starmer, 2000) o también preferencias racionales (Marqués y Weisman, 2008). No sorprende que los economistas trataran de preservar este

núcleo central e intentaran modificar EUT asumiendo los mínimos costos posibles. Machina (1982) por ejemplo, debilita el axioma de independencia, el cual no amenaza la racionalidad. Tal estrategia fue capaz de eliminar algunas anomalías, pero no *preference reversals*. Para lidiar con este fenómeno *Regret Theory* (Loomes y Sugden, 1982) invaden el corazón de EUT y dejan caer la transitividad. Otras propuestas han relajado la completitud (Levi, 1980; McCleenen, 1990) y la continuidad (Simon, 1955)

El hecho importante es que la reversión de preferencias no pertenece a la clase tradicional de contraejemplos: ninguno de los axiomas pudo tomarse como responsable de ellos. A la luz de la historia posterior se revela que sus raíces se encuentran en un sector más profundo de la teoría, constituido por los *supuestos* de EUT —específicamente, el supuesto de extensionalidad—, que está en la base misma de la modelización convencional. La tarea fue exitosamente cumplida por *Prospect Theory*. Dejar caer el supuesto del *homo extensionalis* con el objetivo de racionalizar los fenómenos, supuso modificar la ontología básica supuesta por EUT. Este es un cambio mayúsculo y sus consecuencias lógico-filosóficas son de peso. En el nivel lógico, *Prospect Theory* modela agentes intensionales, los cuales eligen no entre loterías, sino entre descripciones de loterías. En el nivel filosófico, emergen problemas relativos a la ontología, la racionalidad, y la predictibilidad. Quizás de mayor importancia práctica, suponer decisores intensionales proporciona la base para la manipulación de las decisiones individuales. Una forma particular de paternalismo puede ser defendida sobre esta base (*cfr.* Cap. VII). Las restantes secciones de este Capítulo exploran la relación entre los cambios lógicos y ontológicos. Se incluye al final un Apéndice en el cual se proporciona una reconstrucción posible de la lógica intensional respecto a loterías (pero extensional respecto a *descripciones de loterías*) seguida por los agentes descriptos en *Prospect Theory* al tomar decisiones en contextos de riesgo.

5.8-Ontología

Desde la perspectiva de un observador extensional, un caso de *preference reversals* originado en *framing effects* (por ejemplo, *Asian Disease*) sería reconstruido de la siguiente manera: en la primera elección los sujetos prefieren A por sobre B, y en la segunda B por sobre A. La segunda elección implica que A no es preferido sobre B. Asumiendo que las preferencias no han cambiado, puede entonces describirse el resultado del experimento de esta manera: A es preferido sobre B y no ocurre que A es preferido sobre B. Es decir

$P \cdot \sim P$

El experimento revela una *contradicción* en las elecciones de la mayoría de los individuos. Esto puede describirse de dos formas: sostener que los sujetos tienen preferencias contradictorias o que carecen de preferencias. Si no se desea abandonar el terreno de las preferencias (y con él el de la economía), se impone la primera opción: casos como el de *Asian Disease* revelan que existen “hechos” contradictorios en el dominio de las preferencias. El principio de Tercero Excluido garantiza que esto no puede ocurrir:

$\sim (P \cdot \sim P)$, es decir $P \vee \sim P$

Una manera de eliminar la perplejidad que causa la reversión de preferencias en el caso de *Asian Disease* es *abandonar el principio de Tercero Excluido*. Admitir que existen algo así como preferencias inconsistentes en el dominio de las elecciones riesgosas. Ante la anomalía de *Asian Disease*, pues, ésta podría ser una respuesta extrema: si el hecho contradictorio no desaparece modificando sus axiomas puede incorporárselo a la familia de hechos posibles (racionalizarlo) abandonando la restricción lógica que impide que hechos de este tipo puedan darse.

Esto muestra que los cambios en la *ontología* pueden estar relacionados con cambios en la *lógica subyacente*. Supongamos que una mañana una persona

cualquiera se desdobra, y vemos una imagen suya en su cuarto y otro en el escritorio. Esto está prohibido por el principio de Tercero Excluido. Si el hecho prohibido persiste y nos resignamos a aceptarlo en la familia de los hechos “reales”, podemos hacerlo levantando la restricción lógica. O mejor dicho, el mero hecho de aceptarlo como real implica que su imposibilidad lógica ha sido desestimada. En este caso, parece que un cambio en la ontología ha obligado a un cambio mayor en la lógica. En este nuevo mundo, existe P, no P y (P y no P). Cualquiera de estas tres cosas *puede* suceder: la persona puede estar sólo en su cuarto, sólo en su escritorio o en ambos lados. Es el mundo de Alicia en el país de las Maravillas, en el que no vale el Tercero Excluido.

Pero no es esto lo que ha hecho *Prospect Theory*. El camino elegido por la teoría es otro y es menos radical. Ha considerado que los “A” y “B” de la primera elección no son *los mismos* que los de la segunda. El dominio de los objetos de elección en el caso de riesgo ha sido ampliado. Los sujetos eligen ahora no entre loterías sino entre *descripciones* de loterías. El “hecho contradictorio” ya no existe. Una contradicción en el dominio (la presencia de un hecho contradictorio) es eliminada cambiando los “individuos” del dominio. *Prospect Theory* cambia la *ontología* de (la mayoría de) los *tomadores de decisión* que modela, y no los principios lógicos *de la teoría*.

Supongamos un dominio D de “individuos” que son loterías. Estas pueden ser nombradas o descritas de diversa manera. Dados los “individuos” y sus descripciones asociadas a cada uno de ellos el Principio de Extensionalidad garantiza que descripciones equivalentes pueden ser intercambiadas en un texto sin alterar su valor de verdad. Desde el lenguaje de EUT (al que llamaremos L_1) las preferencias de los agentes de *Prospect Theory* siguen siendo contradictorias. Desde el interior de L_1 *ha cambiado la lógica*, y los agentes intensionales son simplemente contradictorios. Desde el lenguaje L_2 de *Prospect Theory* la situación es muy diferente. *Se ha modificado la ontología*, lo que permite “racionalizar” la conducta de los agentes. No hay preferencias contradictorias

desde la perspectiva de L_2 , sino una nueva definición de “mismo objeto”, categoría que ha sido reformulada.

Generalizando lo anterior, existen dos maneras de eliminar una anomalía.. La más radical es vulnerando la lógica, y admitiendo la existencia de hechos contradictorios (coexistiendo junto a los hechos “normales”). La otra (más conservadora, en términos de las creencias lógicas) es eliminando los hechos contradictorios cambiando los individuos del dominio. En este último caso la ontología es ampliada, por así decirlo. Hay que inflarla para preservar la consistencia entre la teoría y los hechos. En el primer caso hay que modificar la lógica subyacente (hay que admitir por ejemplo que una persona puede estar en más de un lugar al mismo tiempo), en tanto que en el segundo la lógica no es alterada (solo hay que admitir que la persona del cuarto no es *la misma* que el del escritorio). Aplicando a la anomalía de *Asian Disease* estas dos formas de eliminar la perplejidad que nos produce, tenemos que si se intenta hacerlo eliminando el principio de Tercero Excluido, ello significaría que aceptaríamos un mundo en el que al tener que elegir entre A y B (loterías) pueden ocurrir tres desenlaces (hechos): los sujetos podrían elegir A sobre B, B sobre A, o ambas cosas (esto último es lo que pasa en el experimento). Si se intenta eliminar la perplejidad cambiando la manera de individuar los objeto de elección, una forma de hacerlo es como lo hace *Prospect Theory*: el A que es preferido a B no es el *mismo* “individuo” cuando B es preferido a A (porque se sustituye a las loterías por sus descripciones; acotando el dominio de descripciones a las efectuadas en términos de ganancias y pérdidas). Esta última forma cambia la ontología de EUT (en realidad, duplica el número de “individuos”, pero ya no es el mismo tipo de individuos; los “individuos” objeto de elección han cambiado), borrando del dominio hechos contradictorios.

5.9- Clases de Individuos y Racionalidad

Parece presentarse entonces una situación curiosa: *Prospect Theory* reconoce la co-existencia de dos clases de individuos, que asumen ontologías diferentes. Los individuos difieren acerca de cuáles (y cuántas) son las entidades básicas que pueblan el mundo a los efectos de adoptar decisiones en condiciones de riesgo. Quienes se comportan consistentemente, según la teoría convencional (los experimentadores e individuos “ilustrados”), “habitan” en un mundo L_1 cuyas entidades son loterías. Quienes se comportan inconsistentemente según EUT (la mayoría) habitan un mundo L_2 conformado por descripciones de loterías. Se presenta pues un curioso fenómeno de incompatibilidad, no sólo entre teorías diferentes (EUT y *Prospect Theory*), lo cual se enmarcaría perfectamente en la visión kuhneana del cambio teórico, sino entre los tomadores de decisiones *dentro* de una misma teoría. Esto es así porque *Prospect Theory* no modela el mundo tal como es, independientemente de la intervención humana, sino que lo que modela son las reacciones de los individuos ante la visión que tienen de la situación de elección en que se encuentran. La teoría establece que hay dos clases básicas de individuos: una mayoría que se comporta intensionalmente y una minoría que lo hace extensionalmente (siempre en referencia a loterías), y ello implica una asimetría fundamental, cuyas consecuencias exceden el campo puramente conceptual y se extienden de manera concreta en la posibilidad de *manipular* las decisiones de los agentes intensionales.

Por otra parte, existe un nexo entre el cambio en la ontología y la *racionalidad*: eliminar la anomalía significa en el contexto de esta teoría mostrar la decisión como racional, es decir *restaurar la racionalidad*.

Según declaración de los propios autores, *Prospect Theory* no es una teoría de la decisión racional. Y esta aseveración puede ser fundada en el hecho de que los agentes que modela son *decisores intensionales*, que carecen de preferencias *sobre loterías*. Ello hace que su comportamiento sea irracional (es decir, que

exista reversión de preferencias). Sin embargo, los agentes intensionales poseen un orden de preferencias bien definido sobre *descripciones de loterías* (que pasan a ser ahora las entidades base entre las que eligen), y maximizan este orden. En la medida en que racionalidad se reduce a maximización, los decisores intensionales son racionales.

Desde un punto de vista diferente, hay un argumento adicional e independiente del anterior que permite rescatar la racionalidad para individuos intensionales. Asumiendo que hay dos clases de individuos: observadores extensionales y una mayoría de agentes intensionales, existe una manera débil de defender la racionalidad de *ambos* grupos. El Principio de Racionalidad popperiano sostiene que las decisiones de los individuos son siempre apropiadas a la percepción que éstos tienen de la situación en que se encuentran. Los agentes extensionales modelados en EUT y los intensionales modelados en *Prospect Theory* perciben de diferente manera la situación de elección que enfrentan. Desde la perspectiva de los primeros, la elección de los segundos es irracional. Pero el dato de que operan con una lógica intensional respecto a loterías permite restaurar la razonabilidad de su conducta de elección mostrando que ellos poseen un *orden bien definido sobre descripciones de loterías*.

Esto puede interpretarse como que existe un nexo entre la ontología asumida por los individuos y la racionalidad de su comportamiento: la cuestión de si los individuos se comportan o no racionalmente puede volverse dependiente de la ontología que se asuma como real (o, una tesis más débil todavía, de la ontología que *ellos* asuman). Si se acepta una ontología reducida en que las loterías son los objetos de elección, la mayoría procede irracionalmente (y la minoría lo hace racionalmente); si, por el contrario, se acepta un universo de *descripciones de loterías*, el comportamiento de la mayoría es ahora racional y el de una minoría no lo es. Si se desea independizar la racionalidad de cuestiones ontológicas, la racionalidad puede ser vista a la manera de Popper en su *Lógica de la Situación*, como consistencia entre las acciones emprendidas y la apreciación que los

sujetos tienen de la situación de elección en que se encuentran. En este escenario, *todos* son racionales (por hipótesis). Todos se comportan racionalmente de acuerdo a la ontología que aceptan o presuponen. En el Capítulo siguiente veremos que, de hecho, es posible interpretar EUT, y *Prospect Theory* como instancias alternativas de la Lógica de la Situación popperiana.

El pasaje de EUT a *Prospect Theory* ejemplifica un tipo *particular* de cambio teórico. En primer lugar, el cambio de la física clásica a la física relativista es sólo parcialmente comparable al cambio que se produce en el pasaje de EUT a *Prospect Theory*. Como EUT es una teoría normativa aceptada como adecuada, puede usarse junto a esta, como lo hacen los modelos paternalistas, cuando el propósito no es meramente descriptivo. EUT continúa siendo utilizada como paradigma de comportamiento ideal (racional), en tanto que *Prospect Theory* permite identificar las desviaciones del comportamiento ideal. La teoría anterior, en este caso, nunca es abandonada, sino que permanece vigente como *benchmark*.

Es la preeminencia normativa de EUT la responsable de introducir una cesura con respecto al cambio conceptual revolucionario sugerido por Quine, y dotarlo de características particulares. *Prospect Theory* no desplaza completamente a EUT. Los modelos de *Behavioral Economics* recurren típicamente a ambos saberes. Emplean *Prospect Theory* para descubrir la manera en que los individuos (intensionales) habrán de elegir entre diversas descripciones de loterías, y utilizan EUT para identificar aquello que más les conviene e inducirlos a tomar ese camino: los asiste en su propio beneficio, como veremos en el capítulo siguiente. Para ser paternalista libertario hay que ser un agente extensional.

En los párrafos precedentes se sugiere que tanto los agentes de EUT como los de *Prospect Theory* pueden ser considerados *racionales* en el sentido de la Lógica de la Situación popperiana. Sin embargo, sólo EUT parece tener preeminencia

normativa, y es considerada como una teoría de la decisión *racional* a diferencia de *Prospect Theory*. La diferencia entre ambos sentidos de racionalidad, que ocasiona tantos malos entendidos en Economía, será explorada en lo que sigue.

Apéndice I: Un sistema de lógica para *Prospect Theory*⁵³

El sistema PT

El sistema está diseñado como un metalenguaje formal. Esta estrategia permite modelar agentes que no prefieren entre loterías, sino entre *descripciones* de loterías. La relación de *preferencia* será denotada por “P”

El lenguaje objeto (LO)

1. Símbolos individuales⁵⁴:

1.1. Constantes de individuo para loterías: $a, b, \dots, w, a_1, b_1, \dots, w_1, \dots$

1.2. Variables de individuo para loterías: $x, y, z, \dots, x_1, y_1, z_1, \dots$

2. Letras predicativas: P, Q, R, ... (Posiblemente sub y supra-indicadas) para propiedades de loterías.

Las constantes lógicas son las conectivas, cuantificadores y el símbolo de identidad.

Las constantes individuales constituyen nombres, en el lenguaje formal, de loterías pertenecientes al dominio de interpretación. Los símbolos predicativos representan propiedades de loterías, tales como la relación de ser más riesgosa. En LO podemos expresar, por ejemplo, identidad entre loterías, relaciones de mayor, menor o igual riesgo entre loterías, etc. Ningún símbolo predicativo será

⁵³ En Lerner, Marqués y Weisman (en preparación)

⁵⁴ Obviamos descripciones definidas en el lenguaje formal para mayor simplicidad. En el lenguaje usado (el castellano) utilizamos indistintamente los términos “nombre” o “descripción”.

interpretado en LO como la relación de ser preferido, ya que, como se ha establecido, la relación de preferencia (P) no está definida entre loterías.

La definición de término, las reglas de formación e inferencia y los axiomas lógicos son los habituales para una teoría de primer orden con identidad.

En LO se pueden incorporar axiomas propios que no involucren a “P”.

El metalenguaje (MT)

MT está construido como una extensión sintáctica de LO.

-Alfabeto adicional

La teoría contiene el siguiente vocabulario adicional (respecto de LO):

1. *Símbolos individuales:*

1.1. Constantes de individuo para nombres de loterías: $a_n, b_n, \dots, w_n, a_{1n}, b_{1n}, \dots, w_{1n}, \dots, a_{in}, b_{in}, \dots, w_{in}$ ⁵⁵, ...

1.2. Variables de individuo para nombres de loterías: $x_n, y_n, z_n, \dots, x_{1n}, y_{1n}, z_{1n}, \dots, x_{in}, y_{in}, z_{in}$ ⁵⁶, ...

2. *Símbolos de función:* $f^1_{g1}, f^1_{g2}, \dots, f^1_{11}, f^1_{12}, \dots$

“ f^1 ” se interpreta como la función “nombre de”. Así, de manera informal, interpretaríamos “ $f^1_{g1}a$ ” como “el primer nombre en términos de ganancia de la

⁵⁵ Convenimos en referirnos al n-ésimo nombre de una lotería a mediante a_n . También podríamos haber diferenciado los símbolos individuales de LO de los de MT colocándoles diferentes subíndices indicativos del subconjunto del dominio en el que toman valores. Análogamente, mediante la variable x_n nos referimos al n-ésimo nombre de una lotería indeterminada x .

Es posible referirse al nombre n ($n \in \mathbb{N}$) de una lotería, dado que el conjunto de nombres es numerable.

⁵⁶ Las variables son seleccionadas; están restringidas a nombres.

lotería a". " f^l " es una función que hace corresponder a una lotería dada, uno de sus nombres (consideramos una secuencia de nombres para cada lotería).

3. *Símbolos de predicado*: la teoría consta de un único símbolo predicativo "**P**" en MT.

Las constantes lógicas son las usuales.

Definición de término en MT

1. Las constantes y variables de individuo son términos.
2. Una letra funcional seguida de una constante o variable individual de LO es un término.
3. Ninguna otra cosa es término.

Las constantes individuales propias de MT y los términos funcionales denotan nombres (descripciones) de loterías pertenecientes al dominio de interpretación. Es decir, constituyen nombres, en el lenguaje formal, de descripciones de loterías.

-Reglas de formación de fórmulas atómicas

1. Las reglas de formación de LO, excepto la de identidad.
2. Si α y β ⁵⁷ son términos en negrita o funcionales (es decir, no pertenecientes a LO) de MT, entonces $P\alpha\beta$ es fórmula atómica.
3. Si α es un término en negrita o funcional, entonces $G\alpha$ y $L\alpha$ son fórmulas atómicas.

⁵⁷ α y β son variables metametalingüísticas para nombres de nombres de loterías.

La relación de preferencia **P** no resulta significativa afirmada entre loterías; por lo tanto, $P\alpha\beta$ no constituye una fórmula bien formada cuando α y β son símbolos de loterías. Es decir, no todo término es un posible argumento para **P**: los nombres de loterías (pertenecientes a LO) no lo son. Sólo los nombres de descripciones de loterías (es decir, los símbolos individuales pertenecientes al vocabulario adicional de MT) constituyen argumentos permitidos. Esta restricción refleja el hecho, que ya habíamos mencionado, de que las preferencias de los agentes se establecen entre descripciones de loterías, y no entre las loterías mismas.

G se interpreta informalmente como la propiedad de estar expresada en términos de ganancia y **L** como la propiedad de estar expresada en términos de pérdida. Ambas constituyen propiedades de descripciones de loterías.

Las reglas de formación e inferencia para conectivas y cuantificadores para el vocabulario adicional de MT son las usuales para una lógica de predicados de orden uno, sin identidad.

El hecho de que no pueda expresarse en el sistema *identidad* entre nombres de descripciones, obedece al hecho de que PT no se ocupa de estudiar preferencias entre expresiones lingüísticas, sino entre loterías. El comportamiento preferencial entre entidades lingüísticas cae fuera de su ámbito. Aunque no sabemos cuál sería dicho comportamiento, presumiblemente resultaría no extensional. Si quisiéramos representarlo formalmente, deberíamos acudir a una lógica con identidad intensional, en la que se restringiera la sustitución de idénticos.

Se permiten cuantificadores seleccionados, es decir, cuantificadores cuyo dominio de cuantificación está restringido (en nuestro caso, al subconjunto de loterías del dominio D o al subconjunto de nombres de D) (Hodges, 1983).

De este modo, podemos expresar, por ejemplo:

$$\forall x P f_{g1}^l x f_{l1}^l a$$

que se lee: existe alguna lotería cuyo nombre₁ en términos de ganancia es preferido al nombre₁ en términos de pérdida de la lotería a. En este caso, “x” toma valores en el subconjunto de loterías del dominio, mientras que en el caso

$$\forall x P_x f'_{11}a$$

que se lee: existe algún nombre de lotería que es preferido al nombre₁ en términos de pérdida de la lotería a, “x” toma valores en el subconjunto de nombres de loterías del dominio.

-Esquema Axioma de inversión

La teoría afirma ciertas regularidades o patrones, para cuya representación es necesario contar con nombres de entidades lingüísticas como los nombres de descripciones de loterías. Por ejemplo, para el caso de reversión de preferencias de *Asian Disease*, reseñado en el Capítulo 4.

Éste resultado establece que cuando las alternativas se expresan en términos de **G** las decisiones mayoritarias reflejan aversión al riesgo, mientras que cuando idénticas alternativas se expresan en términos de **L** las decisiones mayoritarias reflejan propensión al riesgo.

Podríamos atribuir a este enunciado el estatuto de axioma propio de PT. Su representación formal sería:

$$\alpha = \gamma \ \& \ \beta = \delta \ \& \ G\alpha \ \& \ G\beta \ \& \ L\gamma \ \& \ L\delta \ \rightarrow (\beta > \alpha \ \& \ \delta > \gamma \ \rightarrow P\alpha \ \beta \ \& \ P\delta \ \gamma)$$

donde:

- α , γ , β y δ son variables para nombres de loterías (es decir, para símbolos individuales de LO).

- α , γ , β y δ son variables para nombres de loterías (es decir, para símbolos individuales adicionales de MT).

- \rightarrow es una relación definida en el subdominio de loterías.

- P , G y L están definidas en el subdominio de descripciones de loterías.

MT constituye una extensión de LO; en efecto, todo teorema de LO es también teorema de MT.

Semántica

El dominio de interpretación D de la teoría es el conjunto unión de dos conjuntos: un conjunto no vacío de loterías D_l y un conjunto no vacío de descripciones de loterías D_n . Este último, a su vez, incluye dos subconjuntos: el subconjunto ordenado de nombres de loterías en términos de G y el subconjunto ordenado de nombres de loterías en términos de L .

En toda interpretación adecuada de la teoría se asigna a P el conjunto de todos los pares ordenados (α, β) de nombres de loterías pertenecientes a D_n tales que α es preferido a β .

$P\alpha\beta$ es verdadero si, y sólo si, (α, β) pertenece al conjunto de pares ordenados de D_n que la interpretación asigna a P .

Se procede de manera análoga para los restantes predicados de la teoría.

Las condiciones de verdad para fórmulas moleculares y cuantificadas son las usuales.

Observaciones

El formalismo presentado resulta adecuado para modelar casos como el de *Asian Disease*. Hemos diseñado para ello una teoría formal en la que la relación de

preferencia es una relación metalingüística entre determinados objetos del dominio: descripciones de loterías.

Esta aproximación difiere de otras como la de Bourgeois-Gironde y Giraud (2009), en la que los objetos de preferencia son *estados de cosas* denotados por enunciados. En este sentido, el lenguaje de su propuesta es similar al adoptado por G. H. von Wright (1967) en “La lógica de la preferencia”.

En nuestro sistema, por el contrario, los objetos de preferencia son objetos determinados del dominio, denotados por letras de individuo. Este tratamiento se asemeja más a los actuales modelos de preferencia.

Pero ahora no sólo hay loterías en el dominio de interpretación de la teoría, sino también descripciones de loterías; esta extensión del dominio nos permite representar axiomas de PT que dependen de cómo se formulan las opciones (y no de las opciones mismas). Y este cambio en el dominio de la teoría implica un *cambio ontológico*.

En nuestro sistema, “P---“ constituye un contexto no extensional, dentro del cual no es válida la regla de sustitución, incluso (presumiblemente) respecto de los nuevos tipos de argumentos: nombres de descripciones de loterías. Sin embargo, como ya hemos señalado, no es necesario apelar a una lógica intensional. La razón es que simplemente no necesitamos representar identidad entre nombres de las nuevas entidades, por un motivo instrumental: la preferencia entre descripciones no es incumbencia de PT.

La teoría supone (está basada en) lógica y parte de teoría de la probabilidad (esta última formalizada en el subsistema LO). En LO podemos formalizar, por ejemplo, relaciones entre loterías. Como vimos, la relación de preferencia no está definida en el dominio de LO. El estrato MT, con su vocabulario y reglas adicionales, permite formalizar las preferencias de agentes PT.

Finalmente, dado que la teoría de Kahneman y Tversky habla de un alto porcentaje de decididores, y no de la totalidad de los mismos, resultaría adecuada una formulación *default* de los axiomas de la teoría. De este modo, la teoría reflejaría cómo deciden *en general* los individuos en ciertos contextos, admitiendo excepciones. Esta formulación, por supuesto, invalidaría la semántica propuesta.

Capítulo VI

Racionalidad

1-Introducción

La polisemia del término “racionalidad”⁶⁰ suele acarrear dificultades al momento de analizar las implicaciones de las distintas teorías de la decisión. Sus diferentes significados muchas veces resultan en notables confusiones. Para oscurecer aún más el panorama, el adjetivo “racional” arrastra consigo notas *normativas* además de descriptivas, lo que origina malentendidos en relación al nivel del discurso en el que se está argumentando. Un análisis previo es necesario para disminuir la posibilidad de malos entendidos, y echar luz sobre los alcances de las tesis defendidas.

Este capítulo mostrará cómo puede una teoría *descriptiva* de la decisión incorporar la racionalidad, y en cuál de los sentidos específicos del término puede hacerlo. Para esto resulta necesario establecer el significado de preferencias racionales, creencias racionales y decisión racional, por un lado, y por el otro analizar el uso, sancionado por la tradición, de la *maximización* (tanto en su sentido hermenéutico como en su versión de hipótesis empírica) como sinónimo laxo de *racionalidad*. Son cinco usos del término racionalidad *diferentes* (e interrelacionados), un hecho que conviene no perder de vista.

De acuerdo con Kahneman (1981), la teoría convencional había reducido racionalidad a consistencia ¿Qué significa consistencia, y cuál su relación con la racionalidad económica? O en términos más amplios, ¿qué significa “racionalidad” cuando se habla de una Teoría de la Decisión *Racional* como EUT?

Para responder a estas preguntas comenzaremos por la concepción de racionalidad implícita en la teoría de la decisión racional en condiciones de certeza, conocida como Teoría de la Utilidad Ordinal (TUO).

⁶⁰ Ver Nudler, O.: *La Racionalidad, su poder y sus límites*, Buenos Aires, Paidós, 1996.

2-Racionalidad en condicione de *certeza*: Teoría de la utilidad ordinal

La teoría económica postula que las decisiones humanas están motivadas por deseos y creencias⁶¹. Para que una decisión sea *racional* es necesario (aunque no suficiente) que los deseos y las creencias sean racionales. Por supuesto, esto no hace sino desplazar el problema: ¿qué significa tener deseos y creencias racionales?

Los deseos están representados por preferencias sobre objetos. No hay ningún tipo de restricción sobre los objetos particulares de deseo. Este es un punto que merece enfatizarse, con el objeto de alejar un malentendido frecuente en relación a los límites artificialmente estrechos que la economía impondría a los deseos humanos. En la teoría de la decisión convencional, las preferencias son *libres*; uno puede incluso preferir conductas difícilmente defendibles como racionales desde el punto de vista del sentido común, y la economía no tiene nada que decir al respecto. Esto no quiere decir que *cualquier* conducta sea racional. Las restricciones existen, pero no están puestas del lado del *contenido* de las preferencias, sino en la *estructura* que debe respetar el *conjunto de preferencias* de un individuo. No toda estructura de preferencias es válida, lo que introduce límites al tiempo que dota de contenido empírico a la teoría convencional. Para serlo debe respetar cuatro axiomas:

a-Transitividad: *Las preferencias son transitivas si y sólo si siempre que el individuo prefiera el objeto A al B y el objeto B al C, entonces prefiere el objeto A al C.*

Este axioma en particular ha sido defendido como un requisito de racionalidad fundamental por medio de un argumento conocido en la literatura económica

⁶¹ Hausman (1992, p.6)

como “*money pump argument*”. Supóngase que alguien tiene preferencias no transitivas: en tal caso podría embarcarse voluntariamente en una serie de intercambios, cada uno atractivo para el sujeto, que lo dejarían al final de la serie con menos bienes que al principio. Las preferencias intransitivas convierten a quien las sustenta en una víctima potencial de sujetos racionales⁶².

b-Completitud: *Las preferencias son completas si y solo si el individuo puede determinar ante cualquier par de bienes A y B si prefiere A a B, o prefiere B a A o es indiferente entre B y A.*

La presentación brindada hasta aquí es un poco menos exigente que la tradicional (Hausman, 1992). Requiere que los electores decidan ante cada par de bienes con los cuales lo enfrentemos si prefieren uno a otro o son indiferentes ante ellos. La completitud sin embargo es un requisito menos obvio de racionalidad; alguien podría negarse a decidir entre dos “bienes” sin dejar de ser racional. En el ejemplo clásico de Hausman, si preguntaran cuál de sus hijos preferiría que sea seleccionado para morir, una persona razonable bien puede negarse a considerar siquiera la cuestión.

Una vez que tenemos transitividad y completitud dos cosas suceden: en primer lugar, *el conjunto de preferencias es racional*. Tales son las *preferencias racionales*, en su versión mínima, que está en la base de múltiples modelos económicos. En segundo lugar, es posible construir un orden⁶³: con preferencias

⁶² Por ejemplo, supongamos que X prefiere A a B, B a C y C a A, y que al comienzo del día de mercado empiece con 10 pesos y una unidad de A como únicos activos. En estas condiciones, estaría dispuesta a intercambiar su unidad de A por una de B, y pagar una cantidad arbitrariamente pequeña pero positiva –digamos, 1\$–; si ahora le ofrecemos una unidad de C, la cambiaría por su unidad de B y nos pagaría 1\$ más; por último, entregaría su unidad de C a cambio de una unidad de A, pagando nuevamente 1\$. Al final del día de mercado, el sujeto con preferencias intransitivas tendría en su haber 7 pesos y una unidad de A. Pocos llamarían a este sujeto racional.

⁶³ Un requisito necesario cuando trabajamos con bienes infinitos es el de *continuidad*. Las preferencias en tal caso deben satisfacer un axioma adicional, denominado *axioma de continuidad*. En general no se lo considera más que una concesión técnica a la tratabilidad matemática.

racionales y una gran cuota de paciencia podemos escribir una lista ubicando en primer lugar los bienes más preferidos y en último los menos preferidos. Los bienes a los que se es indiferente ocuparían el mismo renglón en la lista. Si a continuación escribimos a la derecha de cada objeto un número arbitrario en magnitud pero determinado en cuanto al orden, y lo llamamos “utilidad”, tendremos una *función de utilidad ordinal*. Las cifras pueden ser cualesquiera; la única exigencia es que el número más grande en magnitud se le asigne al primer elemento de la lista, y el más pequeño al menos preferido (los elementos que colocamos en el mismo renglón compartirán el mismo número), pero no es necesario que guarden otra relación. Por ejemplo, se le podría asignar a A, B y C los números arbitrarios 1000, 45, $45/2$, y este sería una función perfectamente aceptable.

En la tradición utilitarista de Bentham (1789) el término “utilidad” refiere a una propiedad presente *“in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness...or...to prevent the happening of mischief, pain, evil, or unhappiness”* Resulta importante tener en mente que la concepción benthamita está por completo *ausente* en el término utilidad tal como se lo usa aquí. Como se ha dicho, la teoría microeconómica se compromete con que la utilidad es un índice de preferencias, pero nada más; en particular, no hay mención alguna a su naturaleza interna, o a eventuales estados psicológicos o hedónicos de los agentes. El punto merece ser destacado porque suele ser pasado frecuentemente por alto, y se asume de manera errónea que la utilidad maximizada por los individuos en Teoría de la Decisión es la vieja utilidad benthamita.

Es conveniente notar que todavía no estamos ante una *decisión* racional: simplemente tenemos un orden “bien formado” de bienes, cada uno con un número asignado a su derecha al que hemos llamado *utilidad*. Para seleccionar algún elemento de la lista se necesita de una regla, que es la siguiente:

*Regla de Maximización: "Los agentes eligen el objeto superior de la lista, es decir aquel cuya utilidad es máxima"*⁶⁴

Ahora es posible definir aquello que tantos modelos económicos denominan "decisión racional", o racionalidad: la decisión tomada por un individuo es racional si y solo si se cumplen dos condiciones:

a-El individuo posee preferencias racionales

b-El individuo maximiza su utilidad ordinal

En esta *definición doble de racionalidad* un individuo puede ser racional y dadivoso, o altruista; más aun, no se dice nada acerca del *contenido* de los fines perseguidos⁶⁵. Tal ausencia de constricciones puede tener consecuencias éticas desagradables (es incapaz de filtrar preferencias "sádicas" sobre el dolor ajeno, por ejemplo, en el sentido que, si quien las sustenta cumple con las dos condiciones arriba enunciadas, sería racional al menos en este sentido), pero el desagrado se desvanece si uno cae en la cuenta que es el producto de una falacia de atinencia: que sea racional en el sentido económico no implica que lo sea en un sentido más amplio. En particular, *racional* aquí no significa recomendable o defendible desde el punto de vista ético, como parece seguirse del uso habitual⁶⁶.

⁶⁴ Una regla menos transparente aunque mejor formulada y más conocida es la que utiliza Hausman (1992), en la cual una decisión es maximizadora si no existe una opción más preferida a aquella que el sujeto elige.

⁶⁵ Llamaremos a la racionalidad que impone restricciones al contenido de las preferencias, *racionalidad substantiva II*.

⁶⁶ Por supuesto, en un sentido definido los términos de una teoría desarrollada *siempre* se alejan del uso habitual. Parece ser un requisito necesario, porque la vaguedad y la sugestiva imprecisión que caracterizan a los lenguajes naturales, y a las que deben tanto su invaluable riqueza expresiva como su remarcable capacidad de sugerir distintas asociaciones novedosas, constituyen una limitación cuando se trata de una tarea analítica, la cual demanda precisión y claridad conceptuales. Para lograrlas las teorías recortan significados y redefinen términos, lo cual suele considerarse no sólo parte habitual de la práctica científica legítima, sino que además parece constitutiva de *toda* empresa científica. En tal sentido la teoría económica no es una excepción, y utiliza términos como "maximización" o "racionalidad" con un significado técnico, alejado del uso habitual, de la misma manera en que la física lo hace con términos

3-Certeza e Incertidumbre

Hasta ahora no hemos dicho nada acerca de las creencias. Ha sido así porque la teoría de la utilidad ordinal está construida para funcionar en condiciones de conocimiento perfecto, y en tal caso las creencias coinciden con la realidad (es decir, no hay creencias erróneas) El valor de TUO no tiene que ver con su realismo, que en este punto es nulo, ni con el contenido empírico, porque no da lugar a predicción alguna, sino que se relaciona con lo que podríamos llamar un valor *arquitectónico*. Tomando como base la teoría de la decisión en condiciones de certeza se puede derivar analíticamente curvas de indiferencia, a partir de las cuales se construyen funciones de demanda individual y finalmente de demanda de mercado, que junto con las de oferta agregada determinan los precios de equilibrio. De esta manera, el edificio analítico que culmina en las alturas del equilibrio general tiene por fundamento electores racionales en el sentido de TUO. Algo muy distinto ocurre cuando pasamos a condiciones de riesgo o incertidumbre, en cuyo caso las creencias ganan protagonismo, y se incorporan al análisis bajo la forma de *probabilidades*.

Cuando no estamos seguros de un evento, nuestra creencia es probabilística. Por ejemplo, no tenemos certeza de cuál es el número a obtener si arrojamos un dado “normal”; pero tenemos cierto conocimiento acerca de cuál es la probabilidad de obtener cualquiera de las seis posibles cifras. Tampoco tenemos certeza acerca del resultado del próximo partido de fútbol que dispute el equipo X, pero podemos concederle una probabilidad de ganar el encuentro de, digamos, el 60 %. Ambos números probabilísticos son *distintos*, y resulta evidente que el grado de desconocimiento que enfrenta el sujeto en cada caso es de naturaleza diferente. Para dar cuenta de la distinción, suele discriminarse entre riesgo e incertidumbre.

como “masa”, “movimiento”, etc. Para una posición alternativa, véase Crespo y García Huidobro (2010, mimeo)

La distinción tiene su origen en *Riesgo, Incertidumbre y Beneficio* (1921), de Frank Knight. El riesgo se asocia allí con situaciones de elección en que las probabilidades son *objetivas* –son datos para el sujeto, como en los típicos experimentos con urnas o dados, y el segundo con probabilidades *subjetivas*, situaciones donde se desconoce al menos parcialmente la distribución de probabilidad.

De la disputa acerca de la naturaleza de la probabilidad surgió la escuela subjetivista, desarrollada por Frank Ramsey (1931), Bruno de Finetti⁶⁷ (1937) y Leonard Savage (1954). Para los subjetivistas las probabilidades son *grados de creencia*, aplicables tanto a eventos repetibles (v.gr., la tirada de un dado) como a eventos únicos (la perspectiva de la destrucción del medio ambiente) Cualquier asignación de probabilidad es permitida, siempre que en conjunto respeten ciertos axiomas básicos: la probabilidad asignada a eventos excluyentes y exhaustivos debe sumar uno, eventos conjuntivos multiplican las probabilidades y eventos disyuntivos las suman, etc. Una vez satisfechos, estos axiomas implican que las probabilidades subjetivas coinciden matemáticamente con las probabilidades estándar⁶⁸.

Tomando partido estratégicamente por la interpretación subjetivista de la probabilidad, es posible *reducir* las circunstancias de riesgo e incertidumbre a

⁶⁷ De acuerdo con Mongin (1997, p.3) “*De Finetti was particularly emphatic in claiming that probability does not exist in any substantial sense. As he conceived of it, probability does not even necessarily exist in the subject’s mind: it might just be the numerical expression, as perceived by an outside observer, of the property that the subject behaves coherently when choosing between uncertain prospects*”

⁶⁸ Una denominación algo más laxa permite clasificar las circunstancias de *riesgo* como aquellas donde las probabilidades están dadas o son objetivas (al menos en un sentido intersubjetivo; todo sujeto razonable le asignaría a la probabilidad de obtener “cara” en una tirada de una moneda normal un 50%); las circunstancias de *incertidumbre* son aquellas en las que dos sujetos igualmente razonables podrían asignar probabilidades muy distintas. Por ejemplo, la probabilidad de que el precio de un bono suba puede ser distinta para dos analistas, sin que haya *a priori* ninguna forma de determinar si alguno de ellos está equivocado.

una sola, y simplificar de esta manera el problema. Es lo que haremos en lo que sigue.

En circunstancias de riesgo o incertidumbre, la decisión racional es aquella motivada por preferencias racionales y creencias (probabilísticas) racionales. Hemos analizado el significado de *preferencias* racionales, y es hora de hacer lo propio con las segundas. Afortunadamente resulta mucho más sencillo. En condiciones de riesgo o incertidumbre, las creencias son probabilísticas, y una *creencia* probabilística es *racional* cuando obedece a los axiomas clásicos de la teoría de la probabilidad.

Existe una defensa pragmática de la racionalidad de las creencias probabilísticas conocida como *Dutch Book Argument*, que guarda cierto paralelismo con el *Money Pump Argument* señalado hace unos momentos para el caso de transitividad. En el ámbito de la teoría de la decisión económica, un *Dutch Book Argument* constituye una serie de apuestas que garantizan la pérdida a quien posea creencias irracionales. Como en el caso anterior, las creencias irracionales en sentido probabilístico convierte a quien las posea en una víctima potencial de agentes con creencias racionales⁶⁹.

4-Racionalidad en *Expected Utility Theory*

EUT es una teoría de la decisión *racional*. Puesto brevemente, afirma que una decisión en contextos de riesgo será racional si y sólo si se cumplen dos condiciones:

⁶⁹ Por ejemplo, supongamos que alguien le asigne al evento “llueve mañana” una probabilidad de 0.6, y al evento “no llueve mañana” la misma probabilidad. Claramente esto viola uno de los axiomas básicos de la probabilidad, según el cual las probabilidades atribuidas a dos eventos exhaustivos y excluyentes debe sumar 1. Si nuestro hombre tuviera creencias racionales en el sentido utilizado aquí, debería haberle asignado al segundo evento una probabilidad de 0.4. Pero dada su asignación errónea, nuestro individuo aceptaría una apuesta consistente en pagar 10\$ si llueve y recibir 9\$ si no llueve con el agente A, y pagar 10\$ si no llueve y 9\$ en caso que llueva con el agente B. Sea cual sea el resultado meteorológico, el individuo perderá con certeza 1\$.

a-Las preferencias son *racionales* (es decir transitivas, completas, continuas e independientes)

b-El sujeto elige la lotería que posee máxima utilidad esperada.

Tradicionalmente se han emprendido dos clases de defensas de la racionalidad implícita en EUT, una directa y una indirecta. La primera argumenta a favor del carácter normativo de la Hipótesis de Utilidad Esperada, mientras que la segunda emprende una defensa normativa de los axiomas con los cuales la hipótesis es equivalente. En ninguno de los casos se analiza el carácter maximizador de los individuos (la condición b), porque es la instanciación de un principio mucho más general, de carácter metafísico, que se encuentra (para decirlo en términos lakatosianos) en el núcleo duro del programa de investigación neoclásico (Boland, 1981)

La interpretación normativa *directa* de la Hipótesis de la Utilidad Esperada es de larga data. Bernoulli ya había dejado entrever algunas indicaciones referidas a su valor normativo, aunque pone el acento en su valor descriptivo. La interpretación normativa de la hipótesis es señalada también por Neumann y Morgenstern (Schoemaker, 1982, p.531). Más claramente aun, Savage (1954), Marshak (1950) y Robert Strotz (1953) señalan que EUT es menos una teoría descriptiva, acerca de cómo la gente decide de hecho, que una teoría *normativa*, acerca de cómo se debe adoptar decisiones racionalmente⁷⁰.

El descubrimiento de Morgenstern y Von Neumann abrió la puerta para una defensa indirecta de la normatividad de la Hipótesis de la Utilidad Esperada, al

⁷⁰ “*my own feeling is that it would be a strange man indeed who would persist in violating these precepts once he understood clearly in what way he was violating them*” (Strotz, 1953, p. 393)

mostrar su equivalencia con un conjunto de axiomas⁷¹ racionalmente defendibles: transitividad, completitud, continuidad e independencia.

Starmer (2000) ilustra claramente este punto:

“[EUT] showed that the expected utility hypothesis could be derived from a set of of apparently appealing axioms of preference (...) To the extent that its axioms can be justified as sound principles of rational choice to which any reasonable person would subscribe, they provide grounds for interpreting EUT normatively (as a model of how people ought to choose) and prescriptively (as a practical aid to choice)” (p. 334)

La defensa normativa de los tres primeros ya ha sido expuesta. ¿Por qué puede ser el *Axioma de Independencia* un requisito para la decisión racional? Friedman y Savage (1952) han tratado de articular una respuesta convincente:

“(...) this postulate is implied by a principle that we believe practically unique among maxims of wise action in the face of uncertainty, in strength of its intuitive appeal. The principle is universally known and recognized; and the Greeks must surely have had a name for it, though current English seems not to.

To illustrate the principle (...) suppose a physician now knows that his patient has one of several diseases for each of which the physician would prescribe immediate bed rest. We assert that under this circumstances the physician should and, unless confused, will prescribe immediate bed rest whether he is now, later, or never, able to make an exact diagnosis”⁷²

De manera independiente a las circunstancias concomitantes, es *razonable* mantener la preferencia de un evento sobre otro. Con esto se completa una

⁷¹ Algunas presentaciones incluyen además un “postulado de reducción”, como criterio para identificar loterías (Hausman, 1992, p. 26)

⁷² Friedman y Savage (1952), p. 469. Mongin (1997) desestima esta defensa.

estrategia para defender la racionalidad de EUT. Los cuatro axiomas de los cuales se deriva el principio de la utilidad esperada son razonables tomados individualmente, por lo cual *la Hipótesis de Utilidad Esperada también lo es.*

5-Racionalidad Económica vs Racionalidad Clásica

La racionalidad presente en EUT (preferencias bien formadas y maximización) tiene la característica de ser consistente. No vulnera ninguna de las reglas lógicas, se puede representar matemáticamente y brinda una regla única de elección. Estos rasgos la acercan a la racionalidad clásica, al menos en el sentido amplio en el que puede definirse siguiendo la tradición de Descartes y Leibniz:

“¿En qué consiste, pues, tal concepción clásica de la racionalidad? Podríamos decir, a modo de breve caracterización, que de acuerdo con ella la racionalidad es una capacidad de la mente humana que posee, entre otros, dos atributos básicos: es, en primer lugar, única, o sea que no puede haber una pluralidad de racionalidades, y, en segundo lugar, el lazo que establece entre sus datos de partida y sus conclusiones es un lazo necesario. Por lo tanto, dos individuos cualesquiera que tengan la misma información sobre un problema dado deberían llegar, si siguen un procedimiento cabalmente racional, a la misma solución. Ello es así porque la racionalidad opera de acuerdo con ciertas reglas universales, en particular las reglas de la lógica.”⁷³

Sin embargo, en la Teoría de la Decisión Racional no es cierto que dos individuos con la *misma* información lleguen a la *misma* solución, a menos que el término “información” sea tan amplio que incluya los *deseos* subjetivos de los agentes. Pero esto sería un tanto forzado, porque la acepción más extendida del término lo hace referir a fenómenos objetivos. En tal sentido, si por información se entienden los *datos* del problema de elección –probabilidades y resultados

⁷³ Nudler, O. *Op. Cit.*, p. xii.

posibles- entonces dos individuos pueden llegar a decisiones diferentes, ambas igualmente racionales, y ambas plenamente compatibles con EUT: el resultado no es único.⁷⁴

A causa de la falta de unidad del resultado, el concepto de racionalidad de EUT no puede ser el de la racionalidad clásica. Quizás por esto hay quienes buscan enmarcar a la racionalidad implícita en los modelos económicos como racionalidad coherentista.

6-Racionalidad Económica ¿Formal o Sustantiva?

Hemos visto que la definición de decisión racional en la economía convencional se define a partir de dos notas muy diferentes:

- a) Preferencias racionales
- b) Maximización de la Utilidad Esperada

Por otra parte, hay quienes denominan a la racionalidad económica como meramente instrumental, o incluso puramente formal, tautológica y desprendida de fines. Ahora bien, si esto fuera cierto no se entiende cómo es posible que se hayan encontrado *anomalías* y *contraejemplos* en la historia reciente. O el concepto de racionalidad es tautológico y vacío, o posee contenido empírico y por ende contraejemplos posibles.

En la vida diaria usamos la racionalidad con dos propósitos muy diferentes, aunque interrelacionados. El primero es el de dar sentido a la conducta de otras personas. Lo llamaremos “uso hermenéutico” de la racionalidad. Para interpretar o entender las acciones de la gente que nos rodea *suponemos que son racionales*; ante un acto que nos sorprende, ensayamos y vamos descartando diversos intentos sucesivos de comprensión. En el *Apéndice* mostraremos que en rigor

⁷⁴ Esto es lo que, posiblemente, Thomas Kuhn ha querido mostrar para el caso de la elección entre teorías alternativas: dos científicos pueden decidir de manera opuesta frente a teorías en competencia, y las dos decisiones son *racionales* en el sentido de EUT. (Weisman, 2009)

tales interpretaciones son *aplicaciones directas* del Meta Modelo conocido como Lógica de la Situación popperiana, pero por el momento podemos avanzar sin introducir nueva terminología. Si al final del día ningún intento queda en pie, sea porque el acto es realmente desusado, sea por nuestra incapacidad de generar interpretaciones consistentes, simplemente renunciamos a comprender. En tales casos suele afirmarse que “X está loco”, lo cual es una trabajosa manera de decir que su conducta es incomprensible para nosotros.

El segundo propósito es menos evidente, pero no menos habitual. Se utiliza generalmente cuando se pretende *anticipar* la conducta de alguien. Podemos denominarlo “uso predictivo” de la racionalidad. Supongamos que seguimos por televisión las andanzas de un dirigente político, que en este momento formula declaraciones inusualmente amistosas en relación a un candidato del arco opositor. Una vez que le damos *sentido* a esta conducta, inventando fines potenciales que hacen compatible la conducta observada con la racionalidad de esa persona, podemos *anticipar* conductas: si los fines imputados al político son, por ejemplo, tejer una alianza con la oposición, podemos prever ciertas acciones. En caso de que la conducta observada sea completamente diferente a la anticipada, lo más probable es que se reconozcan que los fines del sujeto eran otros. Un filósofo de la ciencia expresaría lo anterior diciendo que el uso predictivo de la racionalidad da lugar a reconstrucciones empíricas corregibles.

El ejemplo anterior muestra que hay *dos usos diferentes para la racionalidad*: el de construir interpretaciones y el de anticipar conductas. El uso predictivo supone el uso hermenéutico. Para anticipar conductas se necesita no solamente atribuir racionalidad sino además adjudicarla al sujeto algún fin específico. Una vez establecido este fin, y dando por supuesto que esa persona es racional, se analiza el contexto para identificar de antemano la próxima jugada.

En la práctica habitual, muchos economistas comparten de manera implícita este doble uso de la racionalidad. En primer lugar, toman *maximización* y

racionalidad como sinónimos laxos⁷⁵. En segundo lugar, utilizan la maximización como un *principio metodológico* que les permite comprender conductas, y *también* construyen modelos adjudicándoles fines concretos a los agentes para anticipar su conducta. En ambos casos se afirma que hay agentes maximizadores o racionales *en dos sentidos sumamente diferentes*. Como hemos dicho, confundirlos conduce a enredos de diversa magnitud. En lo que sigue se intentará desenredar algunos de ellos, relativos al status de *Prospect Theory* como teoría de la decisión.

7-Prospect Theory como teoría de la decisión no racional

Prospect Theory no es una teoría de la decisión racional; más aun, no *puede* serlo. Un requisito mínimo de racionalidad, algo que debe estar en el núcleo de cualquier teoría posible de la racionalidad es la *consistencia*, requisito sin duda mínimo pero lamentablemente ausente del campo de las decisiones humanas.

Diversos experimentos muestran que los seres humanos violan el principio de extensionalidad⁷⁶. Como la extensionalidad es condición necesaria para cualquier decisión *racional*, ninguna teoría de la decisión que quiera describir la conducta de decisores extensionales puede ser *racional*. Como vimos, hay un amplio consenso en relación a que *Prospect Theory* no es una teoría de la decisión racional y que además no *puede* serlo:

“Because invariance [is] normatively essential and descriptively invalid, a theory of rational decision cannot provide an adequate description of choice behavior.(..)Consequently, the dream of constructing a theory that is acceptable both descriptively and normatively appears unrealizable” (Kahneman,1986,p.55)

Pero ¿qué significa que algo sea una teoría de la decisión *racional*? ¿Qué significa el término “racional” en este contexto?

⁷⁵ Cfr. Crespo y García Huidobro (2010) para una consideración crítica de tal práctica.

⁷⁶ Ver Capítulo IV.

8-Racionalidad como maximización en Economía *Mainstream*

Boland⁷⁷ ha dicho que es fútil criticar el principio neoclásico de maximización, y sugiere que es irrefutable. Por su parte, algunos críticos (v.gr. Caldwell) señala que algunos economistas lo toman como una hipótesis tautológica y otros como una hipótesis empírica, y hay quienes, como Simon (1955) o Kahneman (2001), argumentan que no solo es refutable, sino que *está* refutado. Obviamente no pueden tener todos razón, a menos que se refieran a cosas distintas. En la presente sección sostendremos que, efectivamente, tal es el caso.

La confusión surge porque hay dos principios de maximización en juego. El primero es un principio puramente vacío y formal, que históricamente hunde sus raíces en los clásicos, y más específicamente en el imperativo smithiano de explicar las conductas de mercado no por simpatía o buenos sentimientos, sino como el resultado de acciones auto interesadas. En sus orígenes fue un imperativo metodológico, relativo a las *explicaciones* a construir, no un imperativo ontológico respecto a la *naturaleza* humana⁷⁸. Retomado por los marginalistas, el auto interés devino en maximización, lo cual abrió la puerta de

⁷⁷ De acuerdo con Boland (1981, p 5) el Principio Conductual de Maximización es tiene un carácter metafísico y cumple una función metodológica: "*In the case of neoclassical economics, the maximization hypothesis plays [a] methodological role. Maximization is considered fundamental to everything; even an assumed equilibrium need not actually be put beyond question, as disequilibrium in a market is merely a consequence of the failure of all decision makers to maximize. Thus, those economists who put maximization beyond question cannot 'see' any disequilibria*"

⁷⁸ Este punto puede ser controvertido; un par de líneas adicionales. La hipótesis del auto interés no afirmaba originalmente que los seres humanos sean egoístas *per se*, sino que las *interpretaciones* de su conducta debían suponer el mínimo grado de simpatía posible como buena práctica metodológica. Visto así refleja un principio de parsimonia razonable, en la medida en que erige una constricción a lo que de otra manera sería una libertad descontrolada para crear modelos *ad hoc*. En otros términos, siempre se podrían atribuir los fenómenos económicos deseables, como el desarrollo, a la simpatía y los negativos a la antipatía humanas, lo cual puede ser útil para condenar moralmente las prácticas económicas de una sociedad (como Proudhon en *Filosofía de la Miseria*), pero no para comprenderla (como bien le reprocha Marx en *Miseria de la Filosofía*)

manera estratégica a la utilización de herramientas matemáticas. La justificación conceptual detrás de la *maximización* como sucedáneo del auto interés ha sido posiblemente la siguiente: el célebre carnicero o panadero que menciona Adam Smith buscan su beneficio guiados por su *propio interés*, lo que los lleva a tratar de obtener la máxima ganancia -es decir a maximizar. La maximización es la forma *racional* de seguir el propio interés o alternativamente, la búsqueda del propio interés está detrás de la maximización.

El principio de maximización así entendido no postula ningún “fin” particular: establece más bien una instrucción dirigida al teórico que puede parafrasearse de la siguiente manera

“Explique toda conducta humana como si fuera el resultado de la maximización de algo”

o bien

“Construya modelos que supongan agentes maximizadores de alguna cosa”

En términos de Boland la instrucción deriva del carácter metafísico de la hipótesis, que está en el núcleo del programa de investigación neoclásico y por lo tanto es irrefutable por decisión metodológica. En su reconstrucción el principio afirma:

Principio Conductual de Maximización (PCM): *Todos los agentes maximizan algo*

El carácter *metafísico* de PCM explica que se pueda aplicar a *cualquier* conducta humana. Para hacerlo sólo basta ingenio suficiente como para postular fines adecuados, completando el casillero vacío “algo” de manera conveniente. Como el PCM no postula ninguno, no sorprende que se acomode a cualquiera de ellos. La persona más abnegada maximiza su “utilidad altruista”, así como el individuo que se arroja por la ventana porque cree que el fin del mundo se acerca

maximiza su utilidad percibida. *Cualquier* conducta es económicamente racional desde el punto de vista del Principio Conductual de Maximización, lo que equivale a decir que es irrefutable. En palabras de Gary Becker:

"[My] analysis assumes that individuals maximize welfare as they conceive it, whether they be selfish, altruistic, loyal, or masochistic. (...) Unlike Marxian analysis, the economic approach I refer to does not assume that individuals are motivated solely by selfishness or gain. It is a method of analysis" (Becker, G. 1992)

Resulta claro que tanto Boland como Becker hacen referencia al principio de maximización *metodológico*, PCM, y en tal sentido son herederos directos tanto de la tradición de los clásicos como del uso hermenéutico de la racionalidad que solemos hacer habitualmente para entender la conducta de la gente que nos rodea. Este uso es diferente del segundo, que entra en escena una vez que se fijan los fines.

En efecto, el *segundo* principio de maximización surge cuando se especifica algún fin: por ejemplo, cuando se afirma que los agentes maximizan su utilidad esperada, como en EUT. Una vez que se especifican fines, *el principio adquiere contenido sustantivo*⁷⁹.

Como hemos visto, en el contexto de EUT los agentes son racionales si y sólo si:

1-Poseen preferencias racionales

2-Maximizan su utilidad esperada

Ambas notas pueden reducirse a la segunda, porque maximizar la utilidad esperada supone la existencia de una función de utilidad esperada, la cual surge a

⁷⁹ Llamaremos a este sentido de racionalidad sustantiva, *racionalidad sustantiva I*. Es claro que es diferente, y más débil, a la racionalidad sustantiva II vista anteriormente. No impone restricciones sobre fines, como hace ésta, sino sobre la estructura de preferencias: de allí extrae su contenido empírico.

partir de preferencias completas, continuas, transitivas, (e independientes, en EUT). Las restricciones impuestas sobre las preferencias están pensadas para garantizar un orden bien formado, algo así como una lista de objetos (loterías en el caso de EUT) reveladora de las preferencias del individuo. Sobre ese listado, el principio de maximización establece que se elije el objeto más alto disponible. En términos explícitos, el Principio Sustantivo de Maximización dice que en circunstancias de riesgo o incertidumbre:

Principio Sustantivo de Maximización (PSM): *Todos los agentes maximizan su utilidad esperada*

A diferencia del Principio *Conductual* de Maximización, este principio es sustantivo, tiene contenido empírico, se utiliza para anticipar la conducta de los individuos –y no solamente para darle sentido- y *por lo tanto es refutable*. Su función es muy diferente al primero.

PCM ≠ PSM

Así como la racionalidad tiene dos usos –el uso hermenéutico y el uso predictivo- la maximización se utiliza en dos sentidos. Recordemos que muchos economistas utilizan el término “maximización” como sinónimo de racionalidad, una costumbre venerable por su origen, pero susceptible de engendrar confusiones. Ahora puede verse que tanto Boland como sus críticos tienen razón: el Principio de Maximización (*conductual*) es metafísico e irrefutable, y el Principio de Maximización (*sustantivo*) es refutable y de hecho está refutado para los partidarios de Kahneman y Tversky. No sorprende entonces la heterogeneidad de respuestas que se ha generado en torno a esta hipótesis:

“(...) Tanto defensores como opositores de la hipótesis han discutido extensamente acerca del status del postulado de maximización. A veces la hipótesis es concebida como poseyendo contenido empírico y a veces es

formulada como una tautología; a veces constituye la pieza central de un sistema teórico y a veces un artificio apenas mencionable” (Caldwell, 1981)

La aparición de ambos usos también es patente en los siguientes pasajes de Kahneman (1986):

*“The assumption of rationality has a favored position in economics. It is accorded all the methodological privileges of a **self-evident truth**, (...), a **tautology** (...). Furthermore, the assumption of rationality is protected by a formidable set of defenses (...) **that restrict the significance of any observed violation of the model.**” (subrayado propio)*

Es fácil entrever varios enredos filosóficos aquí. El pasaje sugiere que el supuesto de racionalidad posee los privilegios metodológicos de ser una verdad tautológica, y además los contraejemplos suelen chocar contra las defensas erigidas por los epígonos de la economía neoclásica. Pero cualquier estudiante de lógica sabe que nada que posea los privilegios metodológicos de una tautología puede tener contraejemplos empíricos, porque su verdad es *a priori*. Evidentemente no es posible que ambas afirmaciones se refieran a la misma hipótesis. La dificultad desaparece cuando se reconocen dos hipótesis de maximización/racionalidad distintas, una con privilegios metodológicos (tautológica *avant la lettre*) y otra con contenido empírico y por ende capaz de sustentar contraejemplos.

Otro de los múltiples ejemplos se encuentra en Kahneman, Wakker y Sarin (1997) En el contexto de la diferencia entre la utilidad revelada en las decisiones –*decision utility*– y la utilidad experimentada como producto de esas decisiones –*experienced utility*– se afirma:

“The elimination of experienced utility from economic thought was justified by important considerations, but it was not costless. Perhaps the heaviest cost was that the exclusive concern with decision utility removed some important

problems from the reach of empirical research. In particular, the proposition that people maximize utility was granted the status of a maintained hypothesis, which is used to constrain the interpretation of other facts but is not itself subject to test. Admitting experienced utility as a measure of outcomes turns utility maximization into an empirical proposition, which will probably be found to provide a good approximation to truth in many situations and to fail severely in others” (pág. 397, resaltado propio)

Es cierto que *experienced utility* convierte a la maximización de la utilidad (en sentido metodológico, es decir PCM) en una hipótesis **empírica**, porque instancia el “algo” que se maximiza –la utilidad experimentada- y de esa manera permite anticipar conductas. Pero *exactamente lo mismo sucede con EUT*, al instanciar el “algo” que se maximiza como utilidad esperada. Ambas están al mismo nivel, como instanciaciones del Principio Conductual de Maximización que permiten arribar a hipótesis con contenido empírico, aplicables en algunos contextos y capaces de “fallar severamente en otros”. Como en el uso predictivo de la racionalidad, una vez que se fijan los fines al modelo, éste se torna empíricamente corregible, pudiendo fracasar severamente como sugiere el pasaje.

La contraposición entre quienes deploran el uso hermenéutico de la maximización y quienes lo celebran es clara en el siguiente pasaje de Schoemaker (1982, p.539 y ss). Refiriéndose a Scitovsky (1976) y Gary Becker (1976), afirma:

“Tibor Scitovsky remarked that the concept of utility maximization “set back by generations all scientific inquiry into consumer behavior, for it seemed to rule out –as a logical impossibility- any conflict between what a man chooses to get and what will best satisfy him”. Gary Becker on the other hand views this as a strength: he acknowledges the potential for tautology or circularity in economics, especially if unobservable transactions costs are permitted ex post: “of course, postulating the existence of costs closes or ‘completes’ the economic

approach in the same, almost tautological, way that postulating the existence of (sometimes unobserved) uses of energy completes the energy system and preserves the law of the conservation of energy... the critical question is whether a system is completed in a useful way" i.e, whether it yields "a bundle of empty tautologies" or provides the basis for predicting behavioral responses to various changes"

Hasta ahora hemos argumentado que los dos principios que nos permitían darle sentido y anticipar la conducta de las personas se utilizan también en contextos más académicos, aunque lamentablemente se los suele confundir. EUT es una teoría de la decisión racional que interpretada descriptivamente intenta anticipar conductas. Avancemos un paso más. ¿Qué hace que EUT sea una teoría de la decisión *racional*? Hemos visto que la respuesta convencional es que su racionalidad tiene un doble origen: preferencias racionales y maximización. De acuerdo al presente análisis, estos dos factores pueden trasvasarse en una tríada de elementos interrelacionados:

1-Presencia de PCM: explica la conducta de los agentes como maximizando *algo*

2-Presencia de PSM: es capaz de anticipar la conducta de los agentes suponiendo que maximizan algo *específico*, es decir la la utilidad esperada

3-Presencia de preferencias racionales: postula un orden transitivo, completo, continuo e independiente que hace posible la maximización en términos matemáticos.

Así, cuando se afirma que EUT es una teoría de la decisión racional se dice que modela individuos maximizadores (en los sentidos 1 y 2) y que poseen preferencias "ordenadas".

9-Racionalidad en *Prospect Theory*

Lo llamativo es que las tres notas que suelen vincularse con racionalidad están también presentes en *Prospect Theory*.

- a) Maximización: *Behavioral Economics* suele utilizar individuos maximizadores. Por ejemplo, en *Behavioral Life Cycle Hypothesis*, el “planner” es un sujeto maximizador:

“The doer is assumed to be pathologically myopic, concerned only with current period consumption. At date t the doer is assumed to possess a subutility function $u_t(c_t)$. We assume diminishing marginal utility ($U_t(-)$ is concave in c_t), and also nonsatiation (U_t is strictly increasing in c_t). In contrast, the planner is concerned with maximizing a function of lifetime doer utilities” (Thaler and Sheffrin, pág. 612)

En otros modelos la conducta maximizadora es aún más explícita. Por ejemplo, en *Equity Premium Puzzle*, la maximización es clara:

“The “prospective utility” of a gamble, G , which pays off x_i with probability p_i is given $V(G) = \sum \pi_i v(x_i)$, by where π_i is the decision weight associated assigned to outcome i .”

*“Therefore, every year Y will solve her asset allocation problem by choosing the portfolio that **maximizes her prospective utility** one year away, just as X does”* (Benartzi and Thaler, 1995, pág. 79 y ss. Subrayado propio)

- b) Preferencias bien formadas: ¿Son las preferencias sobre descripciones de objetos bien formadas? En la medida en que la función valor de *Prospect Theory*, graficada más abajo, es una función continua y derivable, la respuesta es positiva.

Si lo anterior se acepta, entonces podemos pasar revista a las características de *Prospect Theory*:

1-Tiene la maximización como principio conductual (PCM)? Si

2-Tiene la maximización como hipótesis sustantiva (PSM) ? Si

3-Tiene preferencias transitivas, completas, y continuas? Si⁸⁰

Entonces hay racionalidad en *Prospect Theory*. Es una instanciación diferente a la de EUT, y ahora podemos ver claramente que la diferencia reside en el *objeto* de la maximización⁸¹.

Aunque no hay ninguna razón para quedarse con el uso estrecho del término, es posible que en la medida en que los economistas continúen tomando a la maximización como un sinónimo laxo de racionalidad, la confusión es inevitable. Si los agentes de *Prospect Theory* realmente maximizan la *Utilidad Prospectiva*, entonces desde el punto de vista convencional puede considerarse una Teoría de la Decisión *Racional*, a pesar de que:

“It is noteworthy that the everyday usage of the term “rationality” does not correspond to the economists’ definition as transitivity and completeness, that is, maximization of something. The common understanding is instead the complete exploitation of information, sound reasoning, and so forth” (Arrow, 1986, p.S390-1)

El Principio de Extensionalidad es una condición necesaria para cualquier teoría de la decisión *racional*. Los agentes que violan extensionalidad no son racionales *en términos de EUT*, ni pueden serlo, porque sus preferencias no guardan un

⁸⁰ Los alcances de esta demostración exceden los términos de esta tesis. En Weisman (2011, en elaboración) se presenta la argumentación completa.

⁸¹ Ver *Apéndice*.

orden bien formado, por lo cual no existe una función de utilidad esperada para maximizar. Pero el PMC obliga a interpretar *cualquier* conducta observada como un caso de maximización, y además *garantiza* que puede hacérselo: obedeciendo implícitamente este imperativo, Kahneman y Tversky demuestran que la conducta anómala puede modelarse como un caso de maximización: la maximización de la utilidad *experimentada*, o de la utilidad *prospectiva*.

10-Posible carácter normativo de *Prospect Theory*

Existe un argumento *independiente* al de la maximización a favor del carácter normativo de *Prospect Theory*. Los individuos modelados allí buscan maximizar la utilidad experimentada, y su carácter hedónico brinda espacio suficiente como para montar una defensa normativa alternativa a la convencional⁸². Después de todo, la idea de que las acciones futuras se llevan a cabo para estar más satisfechos en relación al estado presente, es una de las tradiciones asociadas con el término racionalidad.

Notablemente, Kahneman ya entrevía en 1981 la posibilidad de que la psicología de la decisión tuviera esta carga *normativa*. Para dar el salto, es necesario complementar el significado heredado de racionalidad como maximización con consideraciones relativas a la *naturaleza* del objeto a maximizar. No todo objeto de maximización es racionalmente defendible. En particular, la utilidad convencional no lo es, al menos en algunos contextos (en efecto, ¿por qué los individuos deberían maximizar la utilidad esperada? ¿Qué clase de ventaja les reporta, aparte de la consistencia?), mientras que la utilidad *experimentada* sí es defendible normativamente:

“The psychology of choice is also relevant to the normative question of how decisions ought to be made. [...] the modern theory of rational choice has

⁸² La defensa supone una racionalidad substantiva en sentido II (donde el objeto a maximizar es *normativamente* defendible) como diferente (y superior) a la racionalidad substantiva en sentido I. (En Weisman, 2011, en preparación)

adopted the coherence of specific preferences as the sole criterion of rationality. [...] Consistency is only one aspect of the lay notion of rational behavior. As noted by March, the common conception of rationality also requires that preferences or utilities for particular outcomes should be predictive of the experiences of satisfaction or displeasure associated with their occurrence. [...] In particular, predictive considerations may be applied to select the decision frame that best represents the hedonic experience of outcomes” (Kahneman y Tversky, 1981)

La estrategia de Kahneman pasa por postular la existencia de un concepto de racionalidad afincado en la tradición de David Hume: son racionales quienes maximizan utilidad *experimentada*, o *felicidad*. Volveremos sobre ello en el capítulo siguiente.

Apéndice: Las teorías de la decisión como instanciaciones de la Lógica de la Situación popperiana.

En la sección precedente se distinguió entre dos clases de *hipótesis*, una formal o conductual, de carácter metafísico, y una sustantiva, de carácter empírico. La dualidad puede rastrearse hasta dos clases de modelos, una conformada por esquemas explicativos, y la otra compuesta por modelos empíricos que surgen por instanciación de aquéllos.

En rigor, la distinción está implícita en el Capítulo 3. Allí argumentamos que el modelo *Behavioral Model* propuesto en Simon (1955), no es un *modelo* empírico, lo cual permitió redimensionar sus aportes en cuanto teoría de la decisión racional alternativa a la Teoría de la Utilidad Esperada (EUT) y su status dentro del campo de la teoría de la decisión. El análisis sugería que, contra la opinión de muchos especialistas en teoría de la decisión, *Behavioral Model* posee *menor* (quizás *nulo*) contenido empírico que EUT. Para explicar este hecho se distinguió entre lo que fue denominado “esquema explicativo” y modelo empírico, y el peso de la argumentación se dirigió a mostrar que el modelo de Simon no era un modelo *empírico* capaz de proporcionar explicaciones y predicciones de acciones, sino más bien un esquema explicativo o interpretativo incapaz de proporcionar consecuencias empíricas.

En este apéndice se desarrollarán algunas de las ideas expuestas anteriormente de manera incipiente. En particular se sostendrán las siguientes tesis:

1) Ciertos esquemas argumentativos funcionan como un conjunto de instrucciones para la construcción de modelos económicos (y sociales) dotados de fuerza explicativa y predictiva. Designaremos a estos esquemas como Meta Modelos (MM) y a los modelos que permiten construir Modelos Empíricos (ME). La Lógica de la Situación popperiana es uno de aquellos Meta Modelos.

2) Aunque los Meta Modelos no poseen contenido propio pueden ser *aplicados* o *especificados*. En el primer caso, se obtienen aplicaciones o instanciaciones

directas; el segundo, si se lo hace de manera conveniente, pueden dar lugar a Modelos Empíricos. En tal sentido, los MM deben ser interpretados como conformando una *taxonomía*: un conjunto de casilleros vacíos que identifican el tipo de variables relevantes, pero que dejan inespecificado qué contenido (y valores) concretos debe darse a dichas variables y cómo deben ser relacionadas entre sí.

.3) La Lógica de la Situación popperiana puede ser instanciada de diversas maneras capaces de arrojar explicaciones y predicciones de acciones. En particular, mostraremos que tanto EUT como *Prospect Theory* pueden ser concebidas como *instanciaciones* de la Lógica de la Situación.

4) La distinción propuesta pueden aportar al esclarecimiento de algunos debates sumamente oscuros, como el relativo al *status* del Principio de Racionalidad.

A. Meta Modelos (MM) y Modelos Empíricos (ME)

El abordaje de los fenómenos sociales involucra lo que llamaremos Meta Modelos. Los Meta Modelos no son modelos en sentido estricto, sino *esquemas* que orientan (muchas veces de manera implícita) acerca de cómo construirlos. Es conveniente concebirlos como un conjunto de categorías (casilleros) donde volcar información del tipo relevante. El esquema no proporciona una teoría, en el sentido tradicional del término, sino que identifica el tipo de variables relevantes para resolver ciertas clases de problemas pertenecientes a un dominio particular. Por ejemplo, sostener que las elecciones de los individuos deben ser explicadas por sus creencias y deseos, como lo hace la *Folk Psychology*, no es ofrecer una explicación propiamente dicha de un fenómeno social particular, sino más bien un esquema para interpretar las decisiones humanas. Sólo indica que las elecciones se vuelven inteligibles modelando a los individuos mediante dos tipos de categorías. Pero no indica cómo especificar dichas categorías ni proporciona un mecanismo preciso de su interacción.

Los Meta Modelos pueden ser *aplicados* a casos específicos de conducta observada (el conocido caso de Richard es un ejemplo de aplicación) o también *especificados* de diversa manera y sus resultados pueden diferir en cuanto a su capacidad para brindar explicaciones y predicciones de fenómenos observables. Llamamos Modelos *Empíricos* (ME) a aquellas *especificaciones* de Meta Modelos capaces de predecir (correcta o incorrectamente) patrones observables de comportamiento. En los Modelos Empíricos acerca del comportamiento de elección los individuos descritos en el modelo pueden ser decididamente *ideales* o presentar distintos grados de aproximación a la “realidad”. Y algo semejante puede decirse del contexto. Por ejemplo, EUT es un modelo que describe (y anticipa) las decisiones que adoptan los *homo economicus* al elegir entre loterías alternativas, en que las probabilidades asociadas a los resultados son perfectamente conocidas. Para evitar equívocos podría llamárselo *modelo ideal* de la decisión. A su vez, *Prospect Theory* es un modelo que intenta describir y anticipar las decisiones riesgosas de los seres humanos reales. Pese a sus diferencias ambos son legítimos Modelos Empíricos porque de ellos se obtienen predicciones y explicaciones de la conducta de elección observada.

Los Meta Modelos, del tipo de la Lógica de la Situación, *no* son modelos ideales ni “realistas”; no son descripciones del comportamiento de agentes de ningún tipo. Conviene pensarlos como *normativos* en el sentido tradicional del término: es decir, como un *conjunto de reglas o instrucciones* para construir modelos empíricos. Esta flexibilidad de los Meta Modelos hace posible que la Lógica de la Situación permita la construcción de especificaciones como EUT y *Prospect Theory*. Ambos deben ser considerados *empíricos* porque de ambos se siguen predicciones (consecuencias) empíricas. Hay pues una diferencia *lógica* importante entre los Meta Modelos y los Modelos Empíricos.

La diferencia entre ellos es también *pragmática*. Un Meta Modelo es un conjunto de instrucciones dirigidas a los *teóricos* acerca de cómo construir modelos empíricos. Cada extensión normativa de Modelos Empíricos es también un

conjunto de instrucciones, pero no orientan acerca de cómo modelar ni están dirigidas a los teóricos; instruyen otras cosas y a otros actores. La extensión normativa de EUT, por ejemplo, instruye a los tomadores de decisiones para seleccionar entre loterías alternativas. Por otra parte, no todo Modelo Empírico posee una extensión normativa. *Prospect Theory*, por ejemplo, es considerado tradicionalmente un modelo *puramente descriptivo* de las elecciones humanas.

También existen diferencias *metodológicas* en cuanto a la forma de evaluarlos. Los Modelos Empíricos suelen ser juzgados por la adecuación predictiva que poseen, mientras que los Meta Modelos se evalúan o por su adecuación empírica indirecta (es decir, por su capacidad heurística para dar lugar a Modelos Empíricos adecuados desde el punto de vista predictivo), o bien por su capacidad de brindar inteligibilidad o echar luz sobre los fenómenos observados.

B-Meta Modelos. El caso de La Lógica de la Situación

En esta sección se examinará un Meta Modelo con capacidad para sugerir la construcción de Modelos Empíricos: la Lógica de la Situación popperiana. Para corroborar este status se ilustrará mediante una aplicación ya conocida y se mostrarán además dos especificaciones suyas: EUT y *Prospect Theory*.

Popper muestra que en las ciencias sociales existe algo así como un dispositivo general de explicación de acciones, al que llama la *Lógica de la Situación* (LS), presupuesto en cualquier modelo teórico particular diseñado para explicar acciones⁸³. La Lógica de la Situación no es más que una formulación precisa del esquema explicativo conocido como *Folk Psychology*. Según ella, para explicar (o predecir) las acciones típicas de los individuos o conjuntos de individuos es necesario construir un modelo en donde se especifique el *contexto* de la acción y se modele a los individuos, precisando su *conocimiento* y postulando los *finés* que pueden perseguir.

⁸³Véase “*Models, Instruments and Truth*”, en Popper, 1996.

El entorno es de naturaleza física y social, y el conocimiento y los fines les son atribuidos a los individuos por el modelador. La Lógica de la Situación puede ser entendida, entonces como un esquema general (o taxonomía) en el que se especifican casilleros vacíos que corresponden a las variables consideradas relevantes para la explicación o predicción de acciones. Se supone que en cada casillero sólo se incorpora la información relevante para el caso particular que se modela. Por ende, instruye al modelador acerca del tipo de variable relevante que debe tomarse en consideración, pero deja en sus manos decidir la especificación particular de cada variable.

Antes de considerar las especificaciones concretas, resulta ilustrativo analizar un ejemplo de aplicación brindado por Popper en “Models, Instruments and Truth”⁸⁴:

“One of my standard examples is a pedestrian, let us call him Richard, who want to catch a train and is in a hurry to cross a road crowded with moving and parked motor cars and other traffic. Let us assume that what we wish to explain are Richard’s somewhat erratic movements in making his way across the road. What are the obvious situational elements to which we shall have to refer? There are first the various parked motor cars... Then there are moving cars and moving people. They set ... limits to Richard’s possible movements, provided we assume that among his many aims there is the aim of avoiding a collision. But there are further elements in the situation which are equally relevant to the explanation of Richard’s movements: the rule of the road, police regulations, traffic signals, zebra crossing, and other such social institutions..... (Besides) the central point of situational analysis is that we need, in order to ‘animate’ it, no more than the assumption that the various persons or agents involved act adequately, or appropriately –that is to say, in accordance with the situation.... (This

⁸⁴ En Popper, K., The Myth of the Framework, Edited by M. A. Notturmo. London and New York, Routledge, 1996.

assumption) is known in the literature under the name 'rationality principle'.....

“

Construyendo una explicación en este sentido “we may be able to explain or to predict, Richard’s movement as he crosses the street”. La aplicación constituye un *modelo* que nos ayuda a entender por qué avanza y retrocede, espera, etc. En otros términos, la explicación saca a la luz lo que estaba implícito en la situación social (es decir, la “lógica de la situación”)

En la medida en que los movimientos de Richard son consistentes con los esperados por el rudimentario modelo construido, se considera que éste explica su conducta satisfactoriamente, y no hay necesidad de modificarlo. Pero supongamos ahora que nuestro hombre inesperadamente se detiene en medio de la calle y golpea un auto, gritando “Ladrones!”. En este caso, nuestra reconstrucción de la situación –según la cual Richard intentaba cruzar la calle de manera segura- estaría *refutada*, y una nueva reconstrucción o modelo debería ser propuesta con el objetivo de recuperar inteligibilidad, mostrando el bien conocido caso de una conjetura falsada reemplazada por una conjetura falsable y aún no falsada.

Es interesante notar que el caso de Richard es solamente una *aplicación directa* de la Lógica de la Situación. Son las instanciaciones –no los casilleros vacíos que, por decirlo así, las contienen- las que son *conjeturas* o hipótesis en el sentido popperiano del término, y pueden ser controladas y eventualmente refutadas por la experiencia⁸⁵. Por su parte la Lógica de la Situación en sí misma no es una conjetura, y por lo tanto no puede ser reemplazada, ni controlada, por medios empíricos directos.

⁸⁵ “Popper considera que una vez formulados, los análisis situacionales son empíricamente criticables y susceptibles de mejoramiento según los patrones del racionalismo crítico, esto es, pueden ser sometidos a refutación” (Gaeta, R., Gentile, N. y Lucero, S., *op cit*, pág. 163)

La Lógica de la Situación está compuesta por los siguientes “casilleros vacíos”:

	<i>Categorías</i>	<i>Principio de Animación</i>
LS (Meta Modelo)	Restricciones físicas	Principio de Racionalidad: “Los agentes actúan de acuerdo a su percepción de la situación”
	Restricciones sociales	
	Conocimiento	
	Objetivos	

Intencionalmente, violentamos ligeramente el rótulo de las categorías en un sentido que Popper no hubiera aprobado. Lo hacemos para facilitar la descripción de las especificaciones de este esquema, empleando un lenguaje que es más familiar para la economía. Popper no denomina como “restricciones” a los componentes del aspecto objetivo de la situación. Sin embargo, esta denominación es compatible con su espíritu, ya que las condiciones físicas como las sociales del entorno que deben ser incorporadas en el modelo son aquellas que funcionan como un obstáculo a la consecución de los fines a que el agente aspira. Más central todavía para esta exposición, hemos visto que la Lógica de la Situación es en rigor un Meta Modelo, no un Modelo Empírico. Pero pueden construirse *modelos particulares* introduciendo especificaciones en cada uno de sus casilleros (es decir, especificando sus variables relevantes). El aspecto central de nuestra interpretación es, efectivamente, que el Meta Modelo instruye al modelador acerca de cuales son las casillas a las que debe incorporar contenido para explicar las acciones particulares: debe modelar el medio (físico y social) en que transcurre la acción y dotar a los individuos de creencias (conocimiento) y fines.

Sin embargo, como Popper aclara, el modelo que se construya siguiendo estas instrucciones debe ser “animado”. La animación de un modelo físico, como el de Newton, se hace incorporando leyes de movimiento. La animación de un modelo

social, en cambio, se obtiene adicionándole el Principio de Racionalidad (PR), que puede enunciarse de esta manera: “Los individuos actúan siempre de manera apropiada a su apreciación de la situación en que se encuentran”. Este principio no forma parte del modelo, sino que es un agrega cuya necesidad se aclarará más adelante.

B.1. Modelo Empírico I: Teoría de la Utilidad Esperada (EUT)

EUT es un Modelo Empírico que resulta de una especificación del Meta Modelo LS y describe la manera en que individuos idealmente racionales deciden en condiciones de riesgo. EUT permite predecir qué patrón de elecciones adoptarán dichos agentes. Se asume que los individuos “ven” la situación de elección tal como es expresamente formulada por intermedio de EUT. Para ejemplificar, supongamos que los agentes tienen que elegir entre las dos parejas de alternativas siguientes:

Elección 1

A: (6000, 1) B: (8000, 0.8)

Los individuos son libres de elegir entre A (en cuyo caso ganarán 6000\$ con certeza) o B (y en tal caso tiene un 80 % de probabilidades de ganar 8000\$); nada en EUT impide que los agentes elijan cualquiera de las opciones. Pero una vez que la decisión ha sido tomada, la teoría predice el comportamiento de los agentes racionales ante el siguiente par de opciones:

Elección 2:

C: (6000, 0.25) y D: (8000, 0.2)

Situación de elección A

Los sujetos deben comenzar por la primera elección y continuar con la segunda. EUT permite modelar esta situación de elección siguiendo los lineamientos de la Lógica de la Situación. Es decir, identificando las restricciones (físicas e

institucionales, si las hay) y especificando las “expectativas” y preferencias de los individuos⁸⁶.

¿Cómo percibe el sujeto las situaciones de elección que enfrenta? No hay una única manera, “obvia” u “objetiva”, de hacerlo. La manera más simple de modelar su percepción de la situación de elección mediante EUT, es asumir que la percepción que tiene el sujeto de las opciones y probabilidades que enfrenta no es distorsionada. Por ejemplo, un resultado de \$6000 es estimado como teniendo una utilidad de $U(6000\$)$, y una probabilidad de 0.25 es “percibida” como una probabilidad de 0.25. En estas condiciones el problema que enfrenta el sujeto debe ser representado mediante las herramientas de EUT, y el resultado anticipado por la teoría es que si se eligió A (B) en la primer elección, se elegirá C (D) en la segunda. No todo vale.

Supongamos ahora que el patrón de elección *observado* es inconsistente con el predicho por EUT. Evidentemente, debemos concluir que el modelo no es capaz de revelar “la lógica” de la respuesta ofrecida por los individuos a la situación en que se encuentran. Es en este punto en donde se revela la importancia del *Principio de Racionalidad*, el cual asegura que los agentes deben percibir su problema de elección de una manera diferente a la modelada por EUT. En consecuencia, si se desea entenderlos es necesario reconstruir *su* manera de ver las cosas. Se debe pues reformular el Modelo Empírico de manera que sea

⁸⁶ LS contiene un casillero rotulado como “fines”, que debe ser llenado de alguna manera. Sin embargo, no hay candidatos obvios para ser incluidos en él. Una salida podría ser que las preferencias del individuo sean consideradas como correspondiendo a sus fines. Otra interpretación de emergencia es postular que, independientemente de las preferencias que pueda tener, el sujeto tiene un fin: se propone maximizar su utilidad. Ambos recursos son dudosos, pero no entraremos aquí en los detalles espinosos de esta identificación. En todo caso, podría decirse que en algunos casos las “acciones” (decisiones) de los individuos no son fácilmente describibles en términos de medios y fines. Es posible que al menos algunas elecciones tengan más el carácter de una reacción instantánea que de un cálculo deliberado de adecuación entre medios y fines.

compatible con los patrones de elección conocidos. *Prospect Theory* es una propuesta (no necesariamente única) con este propósito.

B.2. Modelo Empírico II: *Prospect Theory* (PT)

Prospect Theory es una especificación alternativa (en este caso “realista”) de la Lógica de la Situación. Para nuestros propósitos diremos que *Prospect Theory* propone herramientas conceptuales diferentes a EUT para formular la situación de elección que enfrentan los agentes y la manera en que resuelven su problema de elección. Sostiene que los individuos “editan” la información que reciben y la transforman de manera sistemática. Los sujetos poseen (o actúan como si poseyeran) tres clases de herramienta teórica:

- a) Un criterio (tácito) para distinguir los prospectos en dos clases: regulares y estrictos (positivos o negativos);
- b) Una ecuación diferente para representar cada tipo de prospectos;
- c) Dos funciones: de valor y de ponderación de las probabilidades, que transforman los valores monetarios objetivos de los resultados y las magnitudes objetivas de las probabilidades, y valen en toda situación de elección.

No es necesario discutir aquí si Kahneman y Tversky pretenden que los agentes usen de manera consciente y deliberada estas herramientas o si uno puede suponer que proceden *como si* lo hicieran. Lo importante es que si se parte ahora de la Situación de elección A, referida anteriormente, *Prospect Theory* permite una formulación de las opciones diferente a la ofrecida por EUT y, dada la función de valor (su forma y su pendiente), es posible “predecir” el patrón de elección de los individuos.

En primer lugar, las opciones son representadas ahora de esta forma:

$$A: \pi(0.25) v(6000) \quad B: \pi(0.25) [v(4000) + v(2000)] \quad (I)$$

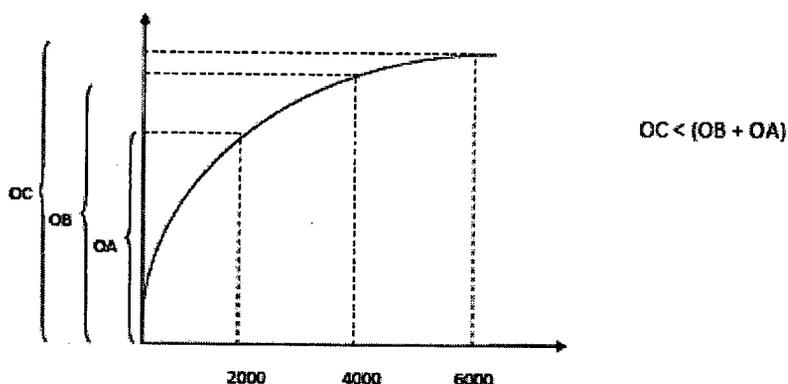
Si agregamos ahora al análisis la función de valor, que es cóncava en el dominio positivo, puede verse que, representando el cambio por una rayita “-“, y haciendo

$$v(0-6000) = OC$$

$$v(0-4000) = OB$$

$$v(0-2000) = OA$$

Resulta que $OC < OB + OA$



Podemos entonces introducir el signo “<” en (I) de la siguiente manera:

$$A: \pi(0.25) v(6000) < B: \pi(0.25) [v(4000) + v(2000)]$$

indicando con ello que B será preferida a A por la mayoría de los sujetos. Por ende, a partir de la formulación de las alternativas en términos de *Prospect Theory* y de la concavidad de la curva se puede deducir (predecir) el sentido (patrón) de preferencia que será observado.

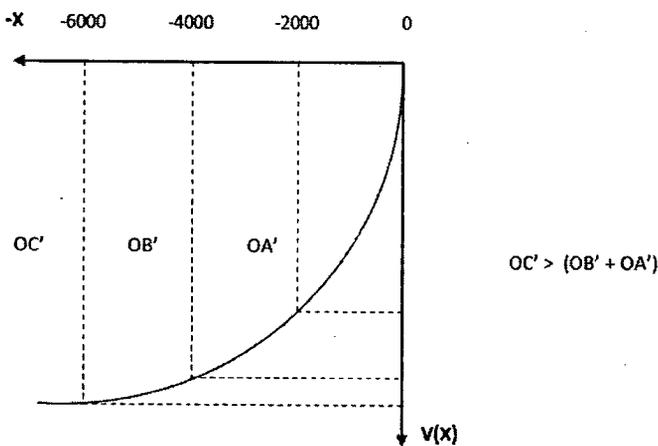
Prospect Theory resulta explicativa allí donde EUT ha fracasado: muestra que un único cambio (ganancia) de 0 a 6000 es menos valioso para los individuos que dos cambios sucesivos, de 0 a 4000, en primer lugar, y (girando el punto de referencia nuevamente al origen) de 0 a 2000, en segundo lugar (aunque el orden no es importante). Ello se desprende de la forma (cóncava) de la función de valor

en el dominio positivo. En consecuencia, especificando la Lógica de la Situación mediante *Prospect Theory* se logra mostrar que la preferencia de B por sobre A es “lógica”, dada la apreciación que hacen los individuos de la situación.

Veamos ahora qué ocurre en el caso de la elección entre loterías negativas. La formulación de la situación de elección en términos de PT es la siguiente:

$$C: \pi(0.25) v(-6000) \quad D: \pi(0.25) [v(-4000) + v(-2000)] \quad (II)$$

Introduciendo la función de valor relevante para pérdidas,



Debería verse que haciendo

$$v(0 - (-6000)) = OC'$$

$$v(0 - (-4000)) = OB'$$

$$v(0 - (-2000)) = OA'$$

Resulta que $OC' > OB' + OA'$

En consecuencia, introduciendo este resultado en (II) “predecimos” que C será más preferido que D por una mayoría de individuos (es decir, $C > D$). Valen para este caso idénticas consideraciones que en el caso anterior.

C-Comentario sobre los Modelos Empíricos descriptos

Ambas especificaciones siguen los lineamientos de la Lógica de la Situación. Resulta interesante notar que el “pasaje” de EUT a *Prospect Theory* puede ser también elucidado por su intermedio: ante el desajuste entre el modelo inicial (EUT) y el patrón de elección observado, se intenta reparar la coherencia (algo que es exigido por el Principio de Racionalidad) reformulando el modelo. El caso es análogo a la aplicación directa que da lugar al ejemplo de Richard. La Lógica de la Situación es no sólo un conjunto de instrucciones para conformar modelos, sino que contiene también instrucciones acerca de cómo *modificar* las construcciones originales en presencia de evidencia adversa. *Prospect Theory* modifica tanto el “conocimiento” como los “fines” de los sujetos. Cambia la percepción que los agentes tienen de la situación de elección a la que se enfrentan y operan con una función de valor y de ponderación que son diversas a las presupuestas en EUT. Naturalmente, aquellos ya no se “proponen” maximizar su utilidad esperada, sino alguna otra cosa (a la que Thaler denomina “*prospective utility*”). Reconstruida de esta manera “su apreciación de la situación”, se ve que las elecciones de los individuos son “lógicas”, dadas las circunstancias.

	Lógica de la Situación	EUT	Prospect Theory
Categorías	FINES	Maximizar la Utilidad Esperada	Maximizar la Utilidad Prospectiva
	CONOCIMIENTO	Conocimiento perfecto	Conocimiento imperfecto
Principio de Racionalidad	Los agentes actúan de manera adecuada a la situación	<i>Los agentes actúan maximizando la utilidad esperada</i>	<i>Los agentes actúan maximizando la utilidad prospectiva</i>

D- Deshaciendo un enredo

Uno de los puntos oscuros de *Models, Instruments and Truth* es el del status del Principio de Racionalidad⁸⁷, algo que el mismo Popper reconoce. El principio de “Actuar de manera adecuada a la situación” o bien posee una naturaleza metodológica, o bien empírica. En el primer caso, no puede ser falso, como ha afirmado Popper con anterioridad; si su naturaleza es empírica, entonces debe formar parte de teorías empíricas y sometido a tests junto a ellas. La respuesta de Popper es que es una tesis empírica, pero como forma parte de (casi) todo modelo social es “buena práctica metodológica” mantenerlo a salvo del *modus tollens*.

Algunos epistemólogos de la economía, como Wade Hands y Bruce Caldwell han llamado la atención sobre el curioso hecho de que el Principio de Racionalidad es una tesis empírica que, paradójicamente, no es falsable. Ambos acuerdan en que esto es incompatible con el falsacionismo popperiano, lo cual ha levantado sospechas respecto a la presencia de un pluralismo metodológico implícito en Popper. Hands es uno de los pioneros en llamar la atención sobre las argumentadas diferencias entre la Lógica de la Situación y el resto de la metodología popperiana:

“The problem with this strict falsificationist view of Popper is that it is inconsistent with what Popper and the Popperians within philosophy of science have actually written about economics and other social sciences. In the few places where Popper directly refers to economics, he is almost never discussing his falsificationist approach to natural science. Instead, economics is discussed

⁸⁷ Ver Popper, *Models, Instruments and Truth*, en The myth of the framework, Edited by M. A. Notturmo. London and New York, Routledge, 2006, pág. 177.

in the context of his 'situational analysis' or 'situational logic' approach to historical and social explanation”⁸⁸

Hands considera que la Lógica de la Situación ofrece una manera diferente de aproximarse a las ciencias sociales, y propone distinguir entre dos “Poppers”, Popper_n (para las ciencias naturales) y Popper_s (para las ciencias sociales).

De aceptarse, la distinción entre Meta Modelos y Modelos Empíricos permite sortear este problema. De acuerdo con ella, el *Principio de Racionalidad* tiene un doble carácter. Como parte del Meta Modelo LS es simplemente un casillero vacío. Convenientemente instanciado se transforma en una *tesis empírica*, y es por lo tanto falsable. En rigor, deberíamos distinguir entre un Principio de Racionalidad Formal y un Principio de Racionalidad Empírico, el cual se obtiene instanciando el primero. Una analogía con la lógica formal puede aclarar este punto. La forma proposicional “Todo A es B” no es un enunciado, sino que es una suerte de esqueleto de enunciado que se convierte en un enunciado propiamente dicho cuando se reemplazan los casilleros vacíos (A y B) por los elementos adecuados. Por ejemplo, instanciando “A” como “hombres” y “B” como “mortales” aparece una proposición (“Todos los hombres son mortales”) que puede ser falsada. Por su parte, la Lógica de la Situación es una *forma general de modelos* con tres casilleros vacíos: fines, conocimiento, y el Principio de Racionalidad (formal). Cuando se completan esos casilleros tenemos un modelo empírico, y dentro de él existe un *Principio de Racionalidad Empírico*.

En EUT, la instanciación del Principio de Racionalidad es el principio de maximización de la utilidad esperada (“Los hombres actúan maximizando la utilidad esperada”). Este es un principio empírico, falsable e incidentalmente falsado a la fecha. *Prospect Theory*, por su parte, instancia el Principio de Racionalidad de la siguiente manera: “Los hombres actúan maximizando la

⁸⁸ Hands, Douglass Wade, *Karl popper and Economic Methodology -A New Look, Economics and Philosophy*, 1 (1), 1985, pág. 84.

utilidad prospectiva”, y hasta la fecha esta conjetura ha resultado empíricamente satisfactoria. Por el contrario, decir “Los hombres actúan de manera *adecuada* a la situación”, tal como lo expresa Popper en la *Lógica de la Situación* no es un principio empírico, como tampoco lo es “Los hombres maximizan alguna cosa”, forma que le da Boland a lo que hemos denominado en el Capítulo 3 el *Principio Conductual de Maximización*.

Capítulo VII

Manipulación de las decisiones individuales y Paternalismo

1-Introducción

Según EUT las preferencias de los sujetos se ejercen sobre bienes y servicios y las variables causales de la decisión son aquellas dos que ya había identificado Bernoulli: los resultados posibles de una acción y las probabilidades asociadas a dichos resultados. Conociendo ambos elementos el agente es capaz de tomar decisiones racionales. Por lo mismo, son las únicas variables que el teórico necesita conocer de antemano, ya sea para predecir el curso de acción que emprenderá un sujeto, o para recomendarle el más racional. (Starmer, 2000; Weisman 2009) El contenido empírico de EUT –su capacidad de formular predicciones referidas a la decisión- está fundado en tal conocimiento.

El advenimiento de *Prospect Theory* ha modificado el supuesto relativo al objeto de preferencia de los sujetos. Los seres humanos no poseen preferencias definidas sobre bienes y servicios posibles. Existe una tercer variable causal operando junto a las anteriores (Cap. IV) Según la nueva propuesta, las decisiones individuales están motivadas por probabilidades, resultados y una descripción vinculada al punto de referencia (Kahneman y Tversky, 1979). Esto da lugar al curioso fenómeno de *framing effects*, la reversión de preferencias motivada por cambios no significativos en la descripción de una situación dada.

La tercera variable causal posee, sin embargo, una naturaleza muy distinta respecto a las convencionales. Probabilidades y resultados son propias del problema de decisión que enfrenta el sujeto, y el grado de control que puede ejercer el sujeto sobre ellas es nulo, exactamente igual que el del teórico. Esto coloca a decisores y observadores (teóricos) en pie de igualdad frente a una situación de incertidumbre. La situación es muy diferente en el marco de *Prospect Theory*. Como la descripción es algo pasible de caer bajo el control directo del teórico, se abre una brecha de *asimetría*: los sujetos reaccionan pasivamente frente a la descripción propuesta por el teórico, fenómeno que está en la base de los *framing effects*. Tal asimetría genera la posibilidad de manipular

las decisiones individuales (Tversky, 2002) Las consecuencias de este hecho sobre el *concepto* de preferencia convencional son profundas, y sus efectos se extienden hasta disciplinas muy alejadas en la estructura reticular del conocimiento. En este capítulo argumentaremos a favor de las siguientes tesis:

7.1-Las preferencias individuales que intentan sacarse a la luz mediante las encuestas bien podrían estar fatalmente viciadas por la descripción de las preguntas formuladas; el instrumento básico de las ciencias sociales, la encuesta, es puesto en cuestión por la existencia de “framing effects”

7.2-Las preferencias políticas de los ciudadanos pueden no cumplir con los requisitos mínimos de ciudadanía competente, desafiando así el concepto mismo de democracia.

7.3-Las preferencias individuales pueden mejorarse mediante la intervención de un experto que conozca cómo manipular las decisiones, si el experto tuviera además el conocimiento de aquello que más le conviene al sujeto. El paternalismo gana nuevas e inesperadas armas en contra de la visión liberal ortodoxa.

Las dos primeras tesis son consecuencias relativamente directas de *framing effects*; la tercera requerirá algo más de argumentación. Empezaremos por el primer punto.

2-El problema de las Encuestas⁸⁹

La idea de que las encuestas no suelen brindar información confiable está bastante extendida, no sólo entre quienes las diseñan y utilizan con regularidad sino entre el público en general. Si bien las metodologías con base estadística

⁸⁹ En esta sección nos restringimos únicamente a aquellas encuestas que intentan elucidar *preferencias* mediante preguntas del tipo “¿Qué prefiere, A o B?”, o bien “¿Está de acuerdo con la implementación del programa X?”. Las encuestas del tipo “boca de urna”, donde se busca indagar no preferencias, sino acciones cercanas, pasadas o futuras, o las encuestas numéricas no serán consideradas aquí.

intentan minimizar los sesgos, y se confía en la probidad de los científicos sociales para descontar manipulaciones, los resultados muchas veces son desalentadoramente paradójicos.

Más allá de la honestidad científica de quienes diseñan/aplican las encuestas, o de la implementación de mejoras estadísticas, hay una razón teórica que explica –y predice– la falla de las encuestas, al menos en un dominio relevante. El culpable de la falla en las encuestas puede ser el mismo supuesto básico que guía su uso: la existencia de *preferencias libres y autónomas*.

¿Puede repararse este inconveniente? Como veremos, hay un argumento que impide mostrarse demasiado optimista al respecto. La *existencia* de una propiedad comparativa (cómo “más preferido que”, o “más pesado que”) exige la invariancia procedural: dicho de manera escueta, la escala (comparativa) construida a partir de dicha propiedad debe ser *única*, e *independiente* del procedimiento de construcción. En la medida en que esto no ocurra, y que encuestas equivalentes den lugar a escalas de preferencias distintas, es difícil sostener la existencia de esas entidades.

En términos simples, los supuestos detrás de la utilización de las encuestas son los siguientes:

1-Existe una *escala de preferencias* que es independiente de, y previa a la encuesta. Llamaremos a esta característica el supuesto de *preferencias libres*

2-La decisión entre las opciones ofrecidas está motivada únicamente por características *internas* de esas opciones. Llamaremos a este segundo supuesto el de *preferencias autónomas*, porque salvo excepciones las decisiones aparecen motivadas por la relación de un sujeto con el mundo, sin posibilidad de que un tercero interfiera de manera sistemática en esta relación.

Las encuestas pretenden idealmente sacar a la luz preferencias libres y autónomas. Hemos visto que, si *Prospect Theory* puede *predecir* la conducta

individual, es porque ha identificado una variable causal relevante en las decisiones humanas: además de los resultados y las probabilidades, la *descripción* de la situación determina las decisiones de los agentes. Como las encuestas no pueden sino describir una situación, existe un serio peligro de que interfieran en aquello que pretenden elucidar. Si es así, las precondiciones para la utilización de las encuestas como *instrumento* para sacar a la luz las preferencias parecen no estar dadas.

2.1-El mito de las preferencias libres y autónomas

Es hora de evaluar una objeción frecuente relativa a la naturaleza última de las preferencias. La objeción comienza reconociendo que la manipulación puede afectar a algunos sujetos, sea por inadvertencia o por falta de educación. Sólo en tales condiciones es posible engañar *momentáneamente* a la gente con respecto a qué es lo que verdaderamente prefiere. Pero en la medida en que se la *educación* o el *aprendizaje* tengan lugar, concluye, sus *verdaderas* preferencias (libres y autónomas) saldrán a la luz.

Es necesario reconocer que la objeción no impacta como una idea extraña. La idea iluminista que vincula el aprendizaje con la liberación individual subyace a ella y permea la cultura liberal. Sin embargo, tomada en su justo valor, la afirmación no deja de ser un argumento empírico, y debe ser evaluado como tal. Hasta qué punto y bajo qué condiciones los individuos aprenden reclama investigaciones de campo o de laboratorio. Y una gran cantidad de estudios⁹⁰ señalan que el argumento del aprendizaje se halla infundado en gran medida. Al parecer el aprendizaje no es una tarea fácil, y sólo tiene lugar bajo condiciones sumamente excepcionales:

“The problem with many economic models of learning is that they seem to apply to a very static environment. In fact, such models seem to be directly applicable

⁹⁰ Entre ellos, Einhorn y Hogarth, 1978, y Tversky y Kahneman, 1986.

only to the situation in which Bill Murray finds himself in the movie Groundhog Day. In the movie, Bill Murray is a TV weatherman sent to report on whether the groundhog sees his shadow on Feb. 2. Murray's character ends up reliving the same day over and over again. Although he is a slow learner, the opportunity to rerun the same day repeatedly, and to learn from the consequences of his actions each time, creates a controlled experiment in which he is able to learn many things eventually, from how to prevent accidents to how to play the piano. Alas, life is not like Groundhog Day. In life, each day is different, and the most important of life's decisions, such as choosing a career or spouse, offer only a few chances for learning!" (Thaler, 1986, pp 135-136) . Véase también Tversky y Kahneman (1986, pp. S274 y S275).

Es difícil sostener que pueda haber aprendizaje cuando los resultados de las acciones no son inmediatos, y no pueden atribuirse fácilmente a una acción en particular. Además, en un gran número de circunstancias no es posible saber qué hubiera ocurrido o cuál hubiera sido el resultado obtenido si se hubieran emprendido acciones alternativas. Más aun, en aquellas ocasiones donde la elección es única –las decisiones vitales, por ejemplo- no hay posibilidades de aprender. En contextos complejos no parecen darse las condiciones necesarias para el aprendizaje.

La idea de preferencias libres y autónomas remite a los oscuros orígenes históricos en los que fue forjado el concepto de un hombre autónomo y racional – un sujeto iluminista, portador de preferencias que se revelaban en, pero que eran previas a sus elecciones. Está implícita tanto en la justificación conceptual de la democracia como la de los sistemas de libre mercado (Samuelson, 1970, p. 946; Hayek, 1945; Friedman y Friedman, 1980). Si la elección ha de ser libre (la elección democrática, o la elección de régimen de vida, o la elección de bienes y servicios) es porque detrás de ella hay preferencias autónomas, no manipuladas por terceros. Justamente era la clase de preferencias libres presupuestas por la

teoría de la decisión tradicional (EUT), pero que se pierden irremisiblemente con *Prospect Theory*.

En relación a los supuestos implícitos en el uso de las encuestas, resulta notorio que:

1-Las preferencias *no son libres*: el primer supuesto no se cumple, porque las encuestas incorporan necesariamente una determinada *descripción* de la situación, y el ranking de preferencias depende de tal descripción;

2-Las preferencias *no son autónomas*, porque la descripción es susceptible de ser *manipulada*; más aun, si pueden predecirse las elecciones es porque pueden ser manipuladas en un sentido definido.

Podemos ahora responder a la pregunta de por qué fallan las (al menos, algunas) encuestas. En la medida en que las encuestas pretenden extraer información referente a preferencias libres y autónomas, están tratando de elucidar una entidad mítica. De manera semejante a los químicos anteriores a Lavoisier, los encuestadores persiguen una quimera con instrumentos particularmente inadecuados para ello.

Consideremos el siguiente ejemplo, el cual bien podría formar parte de una encuesta posible:

“La aplicación del paquete de medidas A está asociada a un 90% de empleo. ¿Estaría dispuesto a apoyarlo (A) o no (B)?”

Prospect Theory permite predecir que la mayoría de la gente encuestada optaría por la alternativa A. Formulando la *misma* pregunta en términos de pérdidas:

“La aplicación del paquete de medidas X está asociada a un 10% de desempleo. ¿Estaría dispuesto a apoyarlo (A) o no (B)?”

Una mayoría de los encuestados seleccionarán ahora la opción B. La literatura relevante (Lacy, 1997, Bartels, 1998, Zaller, 1992) registra casos semejantes⁹¹, en los cuales alteraciones en descripciones *equivalentes* conducen a preferencias distintas, dando cuenta de la presencia disruptiva del *homo intensionalis*. En este contexto cabe preguntarse ¿Qué prefiere la gente realmente? ¿Apoya realmente el programa o no lo hace? Esta clase de preguntas son difíciles de contestar, porque supone preferencias libres y autónomas, ejercidas sobre estados finales.

Volvamos ahora a la afirmación relativa al carácter mítico de las preferencias libres y autónomas. La afirmación se fundamenta en que una condición esencial para aceptar la existencia de una entidad es lo que llamamos en el Capítulo IV “invariancia procedural”. La invariancia afirma que el orden comparativo de preferencias es *independiente* del procedimiento utilizado para elucidar dicho orden. Tversky lo define y señala su importancia de la siguiente manera:

“Procedure invariance plays an essential role in measurement theories. For example, the ordering of objects with respect to mass can be established either by placing each object separately on a scale, or by placing both objects on the two sides of a pan balance; the two procedures yield the same ordering, within the limit of measurement errors. Analogously, the classical theory of preference assumes that each individual has a well defined preference order (or a utility function) that can be elicited either by offering a choice between options, or by observing their reservation price. Procedure invariance provides a test for the existence of a measurable attribute. It would have been difficult to attribute mass to objects if the orderings of these objects with respect to mass were dependent on the measuring device. Similarly, it is difficult to defend the proposition that a person has a well-defined preference order (or equivalently a utility function) if

⁹¹ En otro caso relativamente reciente “A majority of citizens expressed a preference “to use military force” at the start of the Gulf War, while the same majority also expressed a preference not “to engage in combat” (Mueller, 1994, p. 30. cit en Druckman, p.230)

different methods of elicitation give rise to different choices" (Tversky, 1999, p. 189).

Contar con un orden único, con una escala que sea independiente del método de medición es una condición necesaria para poder afirmar que una propiedad *existe*. Tomemos como ejemplo una propiedad comparativa cualquiera, digamos el volumen, y apliquémosla dentro de un dominio restringido de tres elementos, A, B y C. La medición del volumen por métodos equivalentes arroja un orden que es *único*. La escala ordinal del volumen de un elefante, un ruiñeñor y una hormiga es *la misma*, ya sea que se la construya mediante métodos gravimétricos o geométricos. Esta circunstancia nos permite afirmar que el volumen es una propiedad independiente de los métodos que se usen para determinarla, lo cual es una forma trabajosa de decir que cualquier método adecuado arrojará el mismo orden comparativo. *Mutatis mutandis*, asumiendo que existen dos métodos equivalentes para cuantificar el flogisto, el ordenar objetos por la cantidad de flogisto dará lugar probablemente a órdenes inconsistentes. Esto puede ser percibido al principio como una falla en la medición, pero si el panorama no mejora en el tiempo quizás daría lugar a que se sospeche de la *existencia* de la entidad misma.

3-Preferencias Democráticas e Incompetencia de la Ciudadanía

El fenómeno de *Framing Effects* pone en cuestión la existencia misma de una estructura de preferencias subjetivas sobre cual pueda montarse la elección individual *económica*. ¿Qué ocurre con las preferencias subjetivas que subyacen a la decisión o elección *política*?

Existe una multitud de estudios que trabajan en los *framing effects* como evidencia de *incompetencia de la ciudadanía*, un fenómeno según el cual los individuos basan sus preferencias en información arbitraria y/o están sujetos a la manipulación por parte de las elites (Zaller, 1992; Bartels, 1998; Edelman, 1993).

La teoría política tradicional le otorga a la noción de preferencia convencional, libre y autónoma, un rol central en la justificación del sistema democrático.

El liberalismo propugnado por teóricos como Milton Friedman encuentra una marca de origen común entre la democracia liberal y el sistema liberal de mercado: tanto la política como la economía liberales están fundamentadas en la expresión libre de preferencias individuales, y ambos requieren del *homo extensionalis*. Los electores en un sistema democrático *revelan* sus preferencias por medio del voto; alternativamente, las elites en los sistemas democráticos representan la “voluntad” popular, en el sentido de que satisfacen el orden de preferencias de los electores. Aun esta reconstrucción mínima reclama la presencia de una *ciudadanía competente*, capaz de evaluar de manera óptima las alternativas de programas políticos que enfrenta.

Hemos visto que los *framing effects* abren la puerta para la *manipulación* de las decisiones. A diferencia de las variables convencionales, la descripción está directamente bajo el control del teórico. La posibilidad de manipular las decisiones descansa en este hecho. Retomemos el ejemplo de *Asian Disease* para ilustrar la posibilidad de manipulación, sus vínculos con la predicción y de qué manera se relacionan con la incompetencia ciudadana. Allí el hecho de que una mayoría de individuos reviertan sus preferencias depende de la manera en que se describa el problema. Es claro ahora que si alguien tuviera fuertes incentivos, por motivos de conveniencia política o económica, para que se eligiera democráticamente la opción *A* simplemente debería presentar el problema bajo la primera descripción. De manera semejante, quienes estén a cargo podrían presentar la situación bajo el marco de pérdidas, para que la votación democrática diera paso al programa opuesto.

La aceptación pasiva del punto de referencia, que constituye la esencia de los *framing effects* (Kahneman y Tversky, 1979) no es un fenómeno aislado, ni está restringido al laboratorio. Consideremos el siguiente ejemplo:

*“Thomas C. Schelling presented a compellingly realistic example of the dilemmas raised by framing. Schelling reports asking his students to evaluate a tax policy that would allow a larger child exemption to the rich than to the poor. Not surprisingly, his students found this proposal outrageous. Schelling then pointed out that the default case in the standard tax table is a childless family, with special adjustments for families with children, and led his class to agree that the existing tax schedule could be rewritten with a family with two children as the default case. In this formulation, childless families would pay a surcharge. Should this surcharge be as large for the poor as for the rich? Of course not. The two versions of the question about how to treat the rich and the poor both trigger an intuitive preference for protecting the poor, but these preferences are incoherent.”*⁹²

Implícito en el funcionamiento de cualquier democracia se encuentra el concepto de *competencia ciudadana*. En palabras de Dahl (1971) *“a key characteristic of a democracy is the continuing responsiveness of the government to the preferences of its citizens”* Puede concebirse, siguiendo esta línea, al democrático como un sistema formado por dos esferas en interjuego permanente: una elite política, cuya función es responder de manera consciente a las preferencias de los individuos, y un conjunto orgánico de individuos capaces de formar y expresar sus preferencias. La competencia ciudadana, entonces, no es otra cosa que la capacidad del conjunto de formar/expresar preferencias.

No toda preferencia es legítima en teoría política. El significado de la expresión “capacidad de formar preferencias” es un tema largamente debatido. Una parte importante de la discusión pasa por encontrar *límites* a las preferencias, de manera de excluir de antemano la satisfacción de aquellas preferencias basadas en datos erróneos o mal ponderados. Algunos requisitos han sido propuestos, como restringir las preferencias legítimas a aquellas basadas en gran cantidad de

⁹² Kahneman, 2002, pág. 1457/8

información (Somin, 1998), o bien en heurísticas efectivas (Lupia y McCubbins, 1998). Todas ellas pueden ser entendidas como intentos de demarcar entre *buenas* y *malas* preferencias (*i.e.*, preferencias admisibles como propias de una ciudadanía competente y aquéllas incompatibles con la democracia) Existe un problema previo, sin embargo, que vinculado a las condiciones de posibilidad que subyacen a la formación de *cualquier* preferencia posible, ya sean “buenas” o “malas”.

El concepto de preferencia supone una relación entre objetos. Brevemente mencionado, dados dos programas alternativos, Programa A y Programa B, un sujeto o bien prefiere el Programa A sobre el B, o bien prefiere el Programa B sobre el A, o bien es indiferente entre ambos. Esta definición mínima excluye cualquier situación no contemplada en aquella triada, y en particular descarta relaciones contradictorias (del tipo “X prefiere A sobre B y prefiere B sobre A”) El primer requisito mínimo, y sin duda poco controversial, es que las preferencias de una ciudadanía competente no deben estar basadas en información arbitraria (Druckman, p. 232; Zaller, 1992), en particular no deben ser afectadas por *la manera en que se describen las opciones* (Arrow, 1982, p.6; Bartels, 1998, p. 7) Es lo que hemos llamado más arriba, en el contexto de las encuestas, “preferencias libres” (Weisman, 2009) y surge del requisito de Invariancia (procedural y/o descriptiva). Su fuerza normativa deriva del hecho de que pocos dirían que el ciudadano que prefiere el Programa A al B cuando se le pregunta si “*Prefiere usted el programa A al B?*”, y a continuación se inclina por el Programa B en lugar del A cuando se le pregunta si “*Favorecería usted el Programa A o B?*” es un ciudadano que ha formado sus preferencias de manera competente. Resulta difícil además imaginar de qué manera las elites pueden satisfacer preferencias tan lábiles como la descripta.

El segundo requisito exige que las preferencias de ciudadanos competentes no deben ser producto de la *manipulación* por parte de las elites, y es el sucedáneo del requisito de autonomía en las preferencias (Weisman, 2009) La manipulación

ocurre por ejemplo cuando las elites influncian sobre las preferencias de los ciudadanos ocultando o brindando información limitada. Las preferencias construidas sobre una información de tal naturaleza difieren de aquellas que se hubieran manifestado sobre la base de la mejor información disponible (Zaller, 1992, p.313), e incluso pueden operar *en contra* de los legítimos intereses de la ciudadanía (Lupia y McCubbins, 1998, p.70).

3.1-Del elector económico al elector político

Robert Dahl ha establecido las condiciones definatorias de una democracia ideal. En primer lugar figura el requisito de *participación efectiva*, según el cual

“Citizens must have adequate and equal opportunities to form their preference and place questions on the public agenda and express reasons for one outcome over the other”; también se requiere *comprensión iluminista*: *“Citizens must enjoy ample and equal opportunities for discovering and affirming what choice would best serve their interests.”* (1989, p. 113 y ss.)

Sin duda, el funcionamiento del sistema democrático representativo requiere de un ciudadano competente, capaz de evaluar información relevante y de expresar preferencias informadas respecto a las políticas formuladas por las elites. La expresión de *preferencias* por parte del electorado *competente* constituye un engranaje central en la maquinaria democrática liberal.

La pregunta acerca de si las condiciones necesarias para su existencia se cumplen o no es una cuestión empírica. De manera notable, el proyecto de *Prospect Theory* implica que ninguna de esas condiciones se cumple:

1-Las preferencias individuales están sujetas a *framing effects*, y los individuos violan la invariancia procedural (Lichtenstein y Slovic, 1971) y descriptiva (Kahneman y Tversky, 1979) en un amplio rango de circunstancias.

2-Las preferencias son manipulables en un sentido preciso, y vinculado con lo anterior: la violación de invariancia descriptiva es causada porque la descripción juega un rol causal en las decisiones, lo que permite manipular éstas por medio del control de las primeras.

En cierto sentido, *Framing Effects* sugiere que no existen preferencias subjetivas a expresar, o alternativamente que las preferencias individuales se modifican con las descripciones asociadas a las opciones. Las eventuales consecuencias son ominosas:

“framing effect render people’s preferences uninterpretable. For example, when people prefer an economic program described as resulting in 95% unemployment but then oppose the same program when told that it will result in 5% unemployment, it is impossible to determine if they support or oppose the program” (Druckman, 2001),

lo cual levanta interrogantes sobre la posibilidad de que *exista* después de todo una opinión pública a ser satisfecha por representantes democráticos (Entman, 1993)

Si este análisis es correcto, se puede afirmar que:

- El fenómeno de *Framing Effects* introduce un quiebre conceptual en la teoría de la decisión tanto económica como política.
- Los fundamentos ideológicos de la democracia liberal son incompatibles con *Framing Effects*.
- Específicamente, la idea de un sujeto autónomo y portador de preferencias individuales es desafiada, lo que introduce –como veremos– la posibilidad de explorar opciones paternalistas en el diseño de políticas económicas y sociales.

La teoría económica tradicional había pensado que los agentes de mercado eran entidades separadas, autónomas y racionales. Dentro de los supuestos básicos convencionales figuraba el de que los individuos podían maximizar su utilidad. Después de todo, ¿quién mejor que el propio agente para hacer lo que le conviene, o elegir lo que más le gusta?

4-Paternalismo libertario

Hasta ahora hemos visto que *Prospect Theory* abre la puerta para la manipulación de las decisiones. En algunas ocasiones pareciera que el carácter de esta manipulación es *necesaria*, porque no puede evitarse dar una descripción de la situación. Pero que sea *necesaria* no quiere decir que sea *beneficiosa* para los individuos. Lo que separa a la manipulación del *paternalismo* es el vínculo con el mayor interés o beneficio para los individuos sobre cuyas decisiones se está influyendo. No toda manipulación es paternalista, pero toda intervención paternalista supone manipulación.

Es interesante notar que estos problemas no tienen ocasión de darse en la teoría de la decisión convencional. El supuesto de conocimiento perfecto de parte de los agentes (*unbounded rationality*) imbricado en EUT implica que nadie puede mejorar el resultado que éstos logran decidiendo por sí mismos (Davidson, 2003). En estas condiciones, el intervencionismo no sólo no es necesario, sino que es *perjudicial* para los intereses de los individuos: genera un resultado sub-óptimo. Modelar a los agentes como siendo *acotadamente* racionales (como lo hace la corriente de *Bounded Rationality*) parece ser *necesario* para que alguna forma de intervencionismo pueda ser defendida. Pero no es *suficiente*, porque los interventores pueden padecer de idénticas (o peores) limitaciones. Si se considera que los interventores poseen también racionalidad acotada, el intervencionismo debe ser defendido con argumentos adicionales. Una manera de defenderlo es asumir alguna forma de *asimetría en el conocimiento a favor del interventor*. Tal como fue entendido tradicionalmente, el intervencionismo asumía que los

expertos conocían mejor que los agentes el modelo de la economía, o sus preferencias y necesidades, lo que autorizaba a eliminar o reducir el papel del tomador de decisiones.

Prospect Theory, y muchos de los modelos de *Behavioral Economics* fundados en ella, permiten la defensa de una forma particular de intervencionismo que se conoce como *paternalismo libertario*. El paternalismo libertario supone una *asimetría cognoscitiva*: asume que las elecciones que los agentes realizan entre opciones están sujetas a *framing effects*, que no son perceptibles para los propios tomadores de decisiones, pero sí para los teóricos (científicos o políticos).

Uno de los últimos golpes de ariete a la fortaleza del sujeto cartesiano lo constituye el fenómeno de la manipulación. Se ha sugerido que la ausencia de un orden de preferencias a elucidar en las elecciones genera problemas conceptuales para los instrumentos y saberes convencionales que de algún modo lo suponen (encuestas, ciudadanía competente). Las elecciones individuales *pueden ser manipuladas*, por medio de las descripciones. Este es el mensaje de *framing effect*. Pero además de ser manipuladas, el resultado de esta manipulación puede ser beneficioso para el individuo. No parece ser una empresa sencilla, porque requiere demostrar, en contra del saber convencional, que las elecciones realizadas por los agentes con perspectivas de maximizar la utilidad fallan sistemáticamente en lograr su objetivo. De hacerlo, habría lugar para tomar decisiones por el individuo y en su propio beneficio. Esto es un requisito presente en la definición de paternalismo:

*Paternalismo: interference with a person's liberty of action justified by reasons referring exclusively to the welfare, good, happiness, interests or values of the person*⁹³.

⁹³ Dworkin (1972) p. 65. Citado por Muramatsu y Fonseca (2009)

El concepto de libertad de acción juega un papel central en la definición de paternalismo. Más adelante volveremos sobre esa cuestión. Por el momento, notemos que el paternalismo reclama una nota normativa con el fin de ser ejercido legítimamente. Para ello se requiere una *segunda* asimetría cognoscitiva, diferente a la primera: es necesario que los interventores conozcan mejor que los propios agentes *lo que es mejor para éstos*. *Prospect Theory* brinda el fundamento para la conformación de una ciencia de la manipulación, sin embargo, para que el paternalismo pueda ser ejercido, se requiere *otra* forma de conocimiento experto que establezca cómo influenciar las decisiones individuales *en bien de los propios individuos*.

5-La Utilidad Experimentada

Cuando alguien emprende una decisión de consumo, selecciona una alternativa en el momento presente T0 con la idea de disfrutar del bien en el momento T1. Podría decirse que en el instante de toma de decisión realiza una conjetura o hipótesis referida al placer o disfrute que experimentará en T1. Llamaremos a la hipótesis realizada en T0 *previsión hedónica*. Esquemáticamente:

T0-----T1

El hiato temporal entre ambos momentos implica que pueden producirse errores en la previsión hedónica. En otros términos, la conjetura realizada por el individuo en T0 bien podría estar equivocada. Supongamos que nuestro agente debe elegir entre A y B en el momento inicial. ¿Qué determina su decisión, desde el punto de vista teleológico? Parece poco controversial pensar que la causa final es alcanzar el placer o rehuir del displacer originado por la realización de su elección. El individuo anticipa cada uno de ellos, y sobre esta base construye un orden de preferencias, del tipo “A es preferido a B”, o “B es preferido a A” o “A y B son indiferentes”. A continuación, realiza la decisión.

En este simple análisis están involucradas dos nociones diferentes de utilidad. La primera es aquella que el agente anticipa o conjetura en el momento T0, y a la cual toma como base para tomar su decisión. Siguiendo a Kahneman (1999) podemos denominar a este primer sentido de utilidad “*decision utility*” (DU). El segundo es aquel que el individuo experimenta efectivamente en T1, una vez que su conjetura se ha realizado y nuestro hombre se reúne con el resultado de su decisión. Llamaremos a este segundo concepto de utilidad “*experienced utility*” (EU)

Es útil notar que EU, la utilidad efectivamente experimentada, constituye muy posiblemente aquello que los individuos tratan de maximizar. Por otra parte, es difícil concebir otro objeto para la rama de la economía conocida como *Welfare Economics*, o Economía del Bienestar. En los últimos años ha venido tomando impulso una disciplina conocida como *Happiness Economics*, la cual intenta introducir marcadores de felicidad de los ciudadanos dentro de la medición del desarrollo de los países, un ranking que tradicionalmente se construye considerando el Producto Interno Bruto o PIB. No es sólo el crecimiento monetario lo único que importa, afirman los propulsores de *Happiness Economics*, sino primordialmente la felicidad de los individuos.

Curiosamente el concepto de utilidad experimentada no es tomado en cuenta por la teoría de la decisión convencional. Esto es así porque en la medida en que la utilidad conjeturada en T0 es la única que el sujeto considera para tomar su decisión, una teoría interesada en el éxito descriptivo en tal ámbito sólo necesita incorporar la utilidad que dispara la decisión. Por otra parte, no existen razones obvias para pensar que la conjetura o hipótesis realizada en el momento inicial se fuera a ver sistemáticamente defraudada en T1. La confianza en el carácter generalmente adecuado de la previsión hedónica implica el orden de preferencias en T0 se verá reproducido exactamente, al menos en la mayoría de los casos, en T1. Este optimismo es el que posibilita dejar fuera de consideración la utilidad experimentada y concentrarse en la utilidad que interviene en la decisión, un

movimiento estratégico que permitió soltar amarras con la psicología, y exornar al término utilidad de connotaciones a estados mentales internos.

6-Cambio en el concepto de Utilidad

Históricamente el término *utilidad* tiene obvias connotaciones psicologistas, asociadas a su origen en la obra de J. Bentham. Para el fundador del Utilitarismo, la utilidad es una característica del objeto vinculada con su capacidad de generar reacciones hedónicas en el sujeto, reacciones (psicológicas) de placer o displacer.

“ By utility is meant that property in any object, whereby it tends to produce benefit, advantage, pleasure, good, or happiness, (all this in the present case comes to the same thing) or (what comes again to the same thing) to prevent the happening of mischief, pain, evil, or unhappiness to the party whose interest is considered: if that party be the community in general, then the happiness of the community: if a particular individual, then the happiness of that individual.”
(Bentham, 1907 p.13)

Durante sus comienzos, la economía albergó a la utilidad con un sentido más o menos benthamita, como un estado interno de placer o displacer. Sin embargo, tal como hemos visto en capítulos precedentes, el operacionalismo⁹⁴ gestado al calor del positivismo lógico en la década de 1930 hizo que cualquier apelación a “entidades ocultas” fuera vista con aprensión, y el concepto de utilidad (benthamita o psicológica) entró en proceso de revisión. El proceso arrojó una definición mínima, según la cual la utilidad no tiene otra entidad fuera de ser una marca en la escala jerárquica de preferencias.

Como hemos visto, si el orden jerárquico se derrumba, la utilidad pierde todo referente, y es una palabra que –como “flogisto” para la química posterior a

⁹⁴ Brevemente, se puede entender el operacionalismo como sosteniendo que el significado de un concepto científico es el conjunto de operaciones realizadas para medirlo. Ver Bridgman (1927)

Lavoisier- no designa nada. Por esta razón tener preferencias bien ordenadas es un requisito que difícilmente pueda salir de la teoría sin lastimarla profundamente.

Sin embargo, muy lejos de estos agoreros pronósticos, algunos de los representantes más conspicuos de la profesión⁹⁵ veían a comienzos de la década de los '50 un avance promisorio en la disciplina, por cuanto se había logrado acuñar un concepto de utilidad metodológicamente intachable y que –acercando la economía al resto de las ciencias duras- incluso se podía medir. La utilidad de la que hablan muchos modelos económicos dista mucho de ser *experienced utility*, y en general no tiene contenido hedónico alguno. Ahora se puede apreciar la ruptura propuesta por Kahneman cuando reintroduce factores psicológicos en la utilidad, volviendo a la tradición de Bentham.

Hemos dicho que existía cierto optimismo en la profesión respecto a que el orden realizado *ex ante*, en T0 coincidiría con el orden verificado en T1. Si esto fuera así, la vuelta a Bentham sería trivial. Sin embargo, numerosos experimentos muestran que, por un lado, la *previsión hedónica* falla, porque los sujetos se equivocan sistemáticamente al predecir las utilidades que les prestarán bienes futuros, y por el otro, tienen un recuerdo distorsionado de las utilidades obtenidas de los bienes en el pasado (*bounded memories*). Ambas capacidades parecen ser necesarias para que las estimaciones *ex ante* de utilidad resulten coincidentes

⁹⁵ Friedman se congratula de poder contar con una teoría unificada de la decisión. Hasta entonces, las decisiones en contextos de certidumbre eran analizadas por la Teoría Ordinal de la Utilidad, y las de riesgo o incertidumbre por teorías ad hoc, las cuales por lo general partían de principios distintos a la anterior. Ver Friedman y Savage (1952, p.463)

con las estimaciones *ex post*⁹⁶. En la medida en que ambas limitaciones subsistan, la posibilidad del paternalismo no puede ser desechada⁹⁷.

Por supuesto, la gente no se equivoca con respecto a todas las situaciones donde una previsión hedónica debe ser hecha. En situaciones que son repetitivas, familiares, y poco espaciadas en el tiempo, la conjetura avanzada en T0 suele ser validada exitosamente por la experiencia. Por ejemplo, cuando se prefiere un helado de chocolate a uno de vainilla sobre la base del recuerdo del placer experimentado, es poco probable que distorsiones cognitivas subyacentes terminen por defraudar nuestra previsión: la utilidad experimentada coincide en estos casos con la utilidad de la decisión, y puede afirmarse entonces que el orden de preferencias *ex ante* coincide con el orden hedónico *ex post*. Sin embargo, cuando las decisiones son únicas, o son decisiones vitales –decisiones relativas a comprar una casa, elegir el lugar geográfico donde vivir, o a la persona con quien se va a compartir la vida- las limitaciones adquieren una presencia real, y entonces los órdenes *ex ante* y *ex post* pueden diferir (Kahneman y Thaler, 2006). La libertad de elección, un valor central para el liberalismo y la teoría económica liberal, podría conducir en estos casos a resultados subóptimos, si el fracaso de la previsión hedónica es sistemático o previsible.

7-Fallas en la previsión hedónica

La literatura especializada abunda en casos donde la previsión hedónica falla sistemáticamente. En aquellos casos donde las circunstancias emocionales entre T0 y T1 son diferentes, cuando el estado en que se toma la decisión está teñido por el hambre, el deseo sexual, o la violencia, los resultados de la decisión en términos de la utilidad experimentada distan mucho de ser maximizadores. El

⁹⁶ “Retrospective evaluations of the experienced utility of past episodes are undoubtedly the most important source of predictions of the hedonic quality of future outcomes” (Kahneman, 1999, p. 211).

⁹⁷ “Is it likely that an objective observer (or a government) could make more accurate predictions than individuals would make on their own behalf?” (Kahneman, 1999, pp. 205 – 206).

caso del “comprador hambriento” que compra alimentos como si esperara permanecer hambriento indefinidamente, ilustra esta posibilidad (Nisbett y Kanouse, 1968). Previsiblemente, cuando se le suministra un *snack* antes de entrar al supermercado, el comprador hambriento se restringe a los ítems planeados (Gilbert, Gill y Wilson, 1998. Citado en Kahneman, 2006, p. 224). El error consistente en proyectar el estado actual al futuro de manera indiscriminada también se advierte en los consumidores de heroína (Badger et. al, 2004), en las compras por catálogos (O’Donoghue y Vogelsang, 2004), en las inscripciones a gimnasios, etc. El *contexto* también puede modificarse entre T0 y T1, generando fallas de previsión y decisiones subóptimas (Hsee, 2000) Al evaluar un electrodoméstico se lo suele hacer por comparación con otros modelos similares, lo que hace resaltar características que motivan la compra pero que resultan ser imperceptibles una vez que el contexto ha cambiado y el bien deja de estar en exposición junto a otros para estar aislado en el living de una casa.

El aprendizaje a partir de experiencias pasadas también presenta problemas en ciertos casos. En una serie de experimentos (Kahneman, Fredrickson, Schreiber y Redelmeir, 1993) un grupo de participantes se involucraron en una serie de dos pruebas conocidas como “cold pressor”, en las cuales se sumerge la mano en agua dolorosamente fría durante un cierto tiempo. El segundo ensayo era más largo que el primero, pero durante los segundos finales la temperatura del agua se aumentaba en un grado centígrado. A pesar de que un principio básico de evaluación racional indica que incrementar la duración de un episodio doloroso no puede incrementar su utilidad, los participantes evaluaron el experimento largo como menos displacentero que el corto, y expresaron preferencias por repetir la experiencia temporalmente más larga. Los recuerdos de una situación hedónica pueden estar distorsionados, lo que los torna poco confiables como base para tomar decisiones.

El fenómeno de *adaptación* también suele ser pasado por alto al evaluar el futuro. La adaptación es el fenómeno según el cual los estímulos continuados

dejan de afectar al sistema perceptivo, y por ende dejan de ser percibidos. Desde el punto de vista evolutivo, la adaptación puede ser vista como una forma ingeniosa de maximizar el flujo de nueva información a procesar. Trasladado el sistema de valuación, el fenómeno implica que las ganancias o pérdidas dejan eventualmente de proporcionar placer o displacer, y una vez que el ser humano se ha adaptado, simplemente percibe su estado como neutral, ni bueno ni malo. Gráficamente puede verse considerando la Función Valor de *Prospect Theory* (Cfr. Cap. IV): cualquier valor monetario a la derecha o a la izquierda termina por acoplarse con el punto de referencia, y deja de ser percibido.

La euforia de ganar la lotería, o la disforia de quedar parapléjico se disuelve en un tiempo sorprendentemente corto:

“The newly paraplegic [or] lottery winner (...) is almost continually aware of that state. But as the new state loses its novelty it ceases to be the exclusive focus of attention (...) Research indicates that paraplegics are in a fairly good mood more than half the time as soon as one month after their crippling accident”.
(Kahneman, 2006. p.230)

Sin embargo, cuando se evalúa el estado de ánimo de un ganador de lotería o de un parapléjico una vez que ha transcurrido un año desde el cambio, se le suelen asignar puntajes invariablemente desplazados. No es difícil ver aquí la sombra del Error de Bernoulli: los portadores de utilidad no son los estados finales (“ser” parapléjico, o “ser” millonario), sino los cambios con respecto a un punto de referencia: es el cambio de gozar de buena salud a tener una discapacidad lo que genera malestar a nivel psicológico, o el pasaje hasta convertirse en rico, pero una vez logrado el fenómeno de adaptación entra en escena y los estados se incorporan al punto de referencia. Sólo se perciben cambios, y sólo los cambios tienen contrapartidas hedónicas. Dejar de lado este hecho psicológico conduce a tomar decisiones cuyos resultados serán desalentadores en términos de utilidad experimentada. Esto parece sugerir que, abandonados a su propia libertad, los

individuos son incapaces de maximizar su utilidad. No sorprende que las limitaciones de la previsión hedónica abren la puerta para la propuesta paternalista. Kahneman termina por reconocer el hecho, aunque, conciente de sus implicaciones, antes lanza un proviso:

“The fact that people sometimes fail to maximize utility does not imply that someone else (spouses, parents, employers, governments) should usurp the right to choose. Spouses, parents, employers and bureaucrats also make errors, and the best parents are those who let their children make some of their own mistakes. But in some cases, those who are “in charge” can guide and influence choices without restricting anyone’s freedom to choose, adopting what Sunstein and Thaler (2004) call ‘libertarian paternalism’” (Kahneman, 2006, p.232)

8-Libertad y Manipulación.

La economía tradicional tiene un fuerte vínculo con la libertad de acción y elección. El intervencionismo está fuera de consideración fundamentalmente por dos razones principales:

- 1- se asume que la libertad dará lugar a decisiones óptimas, y
- 2- aun cuando esto no fuera así no existe ninguna figura capaz de decidir mejor que los propios individuos. No hay asimetría cognoscitiva.

Las investigaciones empíricas de Kahneman y otros ponen en cuestión *ambas* aserciones. La libertad de decisión no conduce a la maximización de la utilidad, en el sentido relevante para el propio individuo. Por otra parte, al identificar un conjunto de mecanismos que gobiernan la percepción y la evaluación de situaciones de elección, surge la asimetría cognoscitiva necesaria para la puesta en marcha del paternalismo. Es conveniente no perder de vista que las investigaciones psicológicas no se limitan a establecer desvíos respecto a la conducta maximizadora, sino que postulan *mecanismos* suficientemente precisos como para habilitar el manejo de los resultados. En tales condiciones es posible

todavía levantar las banderas de la libertad como valor último, pero ciertamente sus vínculos con la eficiencia ya no le dan sustento.

En la sección siguiente diremos algunas palabras más sobre el paternalismo libertario, con el objeto de arrojar luz sobre los vínculos entre la nueva forma de hacer economía, basada en la psicología cognitiva, y la economía convencional de cuño neoclásico, que abjura y se mantiene alejada de cualquier intervención psicologista.

9-De la Teoría de la Decisión a la Teoría Económica: *Behavioral Economics*

Behavioral Economics constituye una serie de modelos basados en las investigaciones psicológicas llevados a cabo fundamentalmente por Kahneman y Tversky. No intentan reemplazar a la teoría económica de cuño neoclásico, aunque tampoco adscriben a ella (Camerer, 2002). Su intención es más modesta: allí donde surgen anomalías inexplicables para los modelos tradicionales, avanzan explicaciones que toman en cuenta las limitaciones cognitivas del individuo. Su estrategia puede ser vista como la de ir avanzando progresivamente sobre las explicaciones convencionales, manteniendo la teoría aceptada con la expectativa de poder erigir otra sobre nuevas bases, posiblemente más *realistas* en el sentido de incorporar mecanismos descubiertos por la psicología cognitiva (Thefts y Weisman, 2010)

En varios sentidos, *Behavioral Economics* vuelve a anudar dos líneas disciplinares que habían sido separadas a comienzos del Siglo XX (Cfr. Cap. V): la psicología y la economía. Pero como hemos visto, lo hace sobre nuevas bases: la psicología utilizada no es la psicología *folk* de Adam Smith, Edgeworth, o Bentham, sino la psicología cognitiva que ha sido renovada a partir de los años 60' (Camerer, *loc.cit*). El estudio del cerebro como una máquina de procesos (lo que corrientemente se denomina “metáfora de la computadora”) permitió a partir de esos años el estudio de fenómenos como la memoria, el procesamiento de información, y finalmente la toma de decisión. Es entonces cuando Tversky y Kahneman golpean la puerta de la economía con sus aportes fundamentales,

mostrando la posibilidad de reemplazar la ficción del individuo perfectamente racional de los modelos económicos neoclásicos por agentes provistos de racionalidad acotada, obteniendo –a diferencia de Simon, 1955 – predicciones testables.

El hecho de contar con predicciones testables es una característica central del proyecto. Más allá del reciente desarrollo de ciertas corrientes epistemológicas sofisticadas dentro del campo de la filosofía, una parte importante de la disciplina económica aun adscribe a la visión positivista según la cual la predicción es el signo de la “buena ciencia”, y los propulsores de *Behavioral Economics* no son la excepción⁹⁸ Las características centrales de sus modelos son las siguientes:

1- Retienen el método neoclásico de maximización. En nuestros términos, conservan la Hipótesis *Conductual* de Maximización

2- Reemplazan el proceso de toma de decisiones por un mecanismo basado en la psicología cognitiva, y que cuenta por tanto con *apoyo empírico independiente* (Weisman y Thefs, 2010; Marqués, 2011).

El segundo punto contribuye a echar luz sobre el sentido en el cual *Behavioral Economics* reclama para sí el título de realista. No es el realismo exigido por Simon (1955), porque los agentes siguen siendo maximizadores, y entendida de manera sustantiva tal actividad continua reclamando capacidades olímpicas que los sujetos reales no poseen. Antes bien, el realismo pasa por incorporar *heurísticas* y *propensiones*, del tipo de *loss aversion*, falta de fuerza de voluntad, cuentas mentales, y otras que han entrado en la caja de herramientas de la psicología cognitiva en los últimos tiempos. Es un realismo delimitado por la *naturaleza de las variables*, y no por el proceso de maximización, que sigue utilizándose a la manera del *como si*. Los agentes, que son racionalmente acotados (pero ya no en el sentido de Simon) se comportan *como si* maximizaran

⁹⁸ “We share the positivistic view that the ultimate test of a theory is the accuracy of its predictions. But we also believe that, *ceteris paribus*, better predictions are likely to result from theories with more *realistic* assumptions” (Camerer, 2002, p. 2)

su utilidad; el proceso que siguen es una interrelación de variables constitutivas de un *mecanismo* sobre el cual es posible operar.

La manipulación introduce otra brecha importante con respecto a la economía neoclásica. Hemos mencionado que allí la intervención no es posible ni deseable. En contraste, los nuevos modelos behavioristas cuentan con un mecanismo de toma de decisiones *realista* que permite *intervenir* en las decisiones, y el descubrimiento de que los individuos en ciertas ocasiones no son capaces de maximizar su utilidad brinda la justificación *normativa* para la intervención paternalista. Veamos el asunto más de cerca.

En primer lugar, ¿de qué *utilidad* se está hablando cuando se afirma que los seres humanos “fallan en maximizar su utilidad”? Hemos visto hace unos momentos que existen al menos dos sentidos del término, la utilidad implícita en la decisión o *decision utility* y la utilidad experimentada al alcanzar el fruto de la decisión, o *experienced utility*.

Es interesante notar que cada concepto de utilidad da pie a una forma particular de paternalismo (Marqués, 2010): se puede intervenir para ayudar al individuo a que maximice la utilidad convencional o a que maximice su utilidad hedónica o experimentada. La presencia de dos sentidos de utilidad conduce a una serie de enredos teóricos, porque no es sencillo determinar a cuál se refieren los modelos económicos, e incluso en algunas ocasiones se opera con ambos sentidos de utilidad al mismo tiempo (Marqués y Weisman, 2009), lo cual dificulta aún más el análisis.

Una forma de clarificar la cuestión es dividiendo el universo de agentes en dos clases: consumidores, cuyo objetivo es maximizar la *utilidad experimentada* derivada del consumo de bienes y servicios, y productores o empresarios, que actúan persiguiendo una *utilidad “convencional”*, traducible como beneficio monetario. No debe entenderseles necesariamente como agentes físicamente distintos, sino más bien como funciones que pueden ser ocupadas alternativamente por la misma persona. La economía tradicional suele utilizar

una ficción semejante en sus modelos. Una diferencia importante es que aquí ambos agentes están expuestos a limitaciones cognitivas, y por lo tanto su racionalidad es acotada en el sentido de *Prospect Theory*. En los modelos convencionales no se discute la capacidad de los agentes para arribar a sus objetivos hedónicos o monetarios, y tampoco es necesario, dados los convenientes supuestos de partida. Sin embargo, el conocimiento empírico brindado por la psicología cognitiva permite adelantar que surgirán distintos tipos de fallas o errores de cálculo que desviarán sistemáticamente a los agentes de sus fines, volviéndolos incapaces en muchas ocasiones de maximizar la utilidad en cualquiera de los dos sentidos. Esta situación torna necesaria la intervención del teórico, en aquellas situaciones donde es esperable que se presenten fallas en la maximización, para “ayudar” al *homo intensionalis* a maximizar sea la utilidad convencional, sea su utilidad hedónica. Ambos conceptos de utilidad son necesarios.

9.1- La necesidad *normativa* del concepto convencional de utilidad

El punto merece ser destacado. Los modelos económicos basados en *Prospect Theory* no pueden abandonar el sentido convencional de utilidad. Hacerlo conduciría a paradojas, dado que el concepto de utilidad hedónica no permite asignar utilidad alguna a niveles de riqueza. Como consecuencia la meta de hacer crecer los niveles de riqueza de un país –uno de los objetivos últimos de la disciplina económica, el crecimiento económico- se vería súbitamente despojada de sustento normativo. *Prospect Theory* implica que los países más ricos no poseen necesariamente ciudadanos más felices, porque el fenómeno de adaptación borrar eventualmente la satisfacción derivada de las mejoras. Los estados no cuentan para el aparato cognitivo, sólo los cambios.

Una confirmación indirecta del argumento según el cual no es posible defender normativamente el *desideratum* de crecimiento económico sin el concepto convencional de utilidad lo constituye la conocida *Paradoja de Easterlin*⁹⁹, según

⁹⁹ Ormerod (2008)

la cual no existe relación entre los *niveles* de ingreso de un país y la felicidad reportada por los individuos. Incluso sucede con mediciones realizadas en un mismo país para períodos diferentes: Japón duplicó varias veces su nivel de riqueza desde la Segunda Guerra Mundial; sin embargo los niveles de felicidad reportados por sus ciudadanos son comparativamente similares. Esto constituye una *paradoja* para la teoría tradicional, según la cual la utilidad está asociada con *niveles* de riqueza, por lo que es esperable que quienes sean más ricos sean consistentemente más felices. Sin embargo, no constituye un fenómeno inesperado para *Prospect Theory*, porque la utilidad se predica de cambios de estado de riqueza, no de niveles de riqueza, y eventualmente desaparece cuando sobreviene el insoslayable fenómeno de adaptación. La utilidad hedónica o felicidad no es idéntica con la utilidad convencional, ni conceptual ni empíricamente. Sin embargo, retomar el Error de Bernoulli es imprescindible para hacer economía, por cuestiones normativas. Los modelos de *Behavioral Economics* son muestra clara de ello.

Sunstein y Thaler (2004) utilizan el paternalismo vinculado a la utilidad neoclásica, que presupone la existencia de asesores expertos y asesorados inexpertos. El experto domina una “ciencia” (*i.e.*, posee el monopolio de una clase de conocimiento del que se encuentra privado el lego). Y es esta la base que le permite intervenir en provecho del lego. Por ejemplo, un experto puede aconsejar la elección del portafolio A y no del B, sobre la base de la estimación de que A arrojará mayores rendimientos. Nada se presume aquí sobre la felicidad que estos rendimientos proporcionarán a los asesorados. Que A debe ser preferido a B porque arrojará mayor utilidad significa simplemente que arrojará mayores dividendos. Estamos en el dominio de la utilidad convencional.

9.2-Paternalismo basado en *Experienced Utility*

La segunda clase de paternalismo exhibe notas comunes con la primera. Nuevamente, hay aquí expertos y legos, pero en este caso el tipo de conocimiento experto es diferente al anterior: es un conocimiento que no lo proporciona la

ciencia económica, sino la psicología cognitiva, especialmente la referida a los procesos de decisión. También contiene un fuerte componente normativo. Sin embargo no recomienda pautas de acción a los agentes usuales (tomadores de decisiones), sino a otros “agentes”, de segundo orden, como empresas y autoridades. Ofrece a estos “agentes” una teoría paternalista que les permitiría adoptar medidas que favorecen a los individuos a su cargo. Podríamos designar “instituciones” o “entidades” a estos agentes de segundo orden. Así podríamos decir que el paternalismo libertario es una herramienta para dichas instituciones, como empresas, asesores financieros y autoridades gubernamentales y económicas embarcadas en la generación de políticas. Las prescripciones adoptan en cada caso rasgos diferentes.

A las empresas se les puede aconsejar una política de compensaciones a sus empleados que maximice el bienestar (felicidad) de éstos. Lo que la gente trata de “maximizar” es su utilidad experimentada. Las políticas de bienestar procuran contribuir a que los sujetos consigan este objetivo. Experiencias psicológicas muestran que un incremento en los ingresos aumenta la satisfacción de quienes lo perciben. Pero éstos se *adaptan* rápidamente al nuevo nivel y al percibir el ingreso incrementado como normal, el plus de satisfacción obtenido desaparece. En la sección anterior veíamos que la satisfacción de un ganador de lotería o la disforia de una persona devenida en parapléjica termina por desaparecer eventualmente, debido al efecto de la adaptación. Un objetivo de las políticas de remuneración debería ser *vencer* o al menos *retardar* la adaptación, lo que permitiría obtener *más útiles por unidad de pago*. Usando de diversas maneras la misma suma de dinero es posible producir diferentes cantidades de satisfacción (utilidad). Un problema importante, descuidado por la teoría económica estándar, es qué distribución de pagos facilitará el objetivo de los individuos de “maximizar” su utilidad *experimentada*. En rigor, la teoría estándar ni siquiera cuenta con las herramientas conceptuales como para plantear el problema, porque la utilidad de una suma cualquiera es, por definición, función del monto total, y cambios insustanciales en la forma de entregarlo son incapaces de alterarla.

Kahneman y Thaler (1991) abordan este problema y concluyen que los individuos obtienen mayor satisfacción de una paga anual consistente en el salario normal mensual más dos bonos adicionales, uno a mitad de año y otro al final del año (es decir, nuestro aguinaldo), que de una paga uniforme en que los bonos se reparten equitativamente a lo largo del año y por ende forman parte del salario normal incrementado por dichos aportes¹⁰⁰.

Esta política de compensación parece dar lugar a un panorama donde todos ganas. Los empresarios no han aumentado las retribuciones, pero al repartirlas de otra manera generan más *experienced utility* en los trabajadores. Por otra parte, la mayoría de las personas ven como algo injusto que las empresas bajen salarios en épocas de bajas ventas, pero consideran *normal* que los trabajadores dejen de percibir ciertos bonos extraordinarios. No es lo mismo anunciar una baja de salarios que un recorte de beneficios, aunque la cantidad de dinero sea exactamente la misma. De esa manera, es posible también establecer políticas de ajuste minimizando la resistencia.

9.3-Libertad en el Paternalismo Libertario

¿En qué consiste la libertad en el paternalismo libertario? No es una respuesta sencilla, porque en la medida que los individuos posean libertad de elección, nadie puede decidir por ellos de manera paternalista. Parece haber aquí una contradicción en los términos. Sunstein y Thaler recogen el desafío, y en *Libertarian Paternalism is not an oxymoron* ensayan una defensa. El error de sus detractores, aducen allí, es pensar que el paternalismo es evitable. Sin embargo hay una multitud de casos donde esto no sucede. Consideremos el siguiente hecho psicológico: cuando existen opciones predeterminadas y se le ofrece a los

¹⁰⁰ Las teorías estándar del ahorro (la hipótesis de Friedman del ingreso permanente y la teoría de Modigliani del ciclo vital) recomiendan un patrón de retribución diferente al aconsejado por Kahneman y Thaler, en que lo más conveniente para los individuos es distribuir *uniformemente* su nivel de consumo a lo largo de sus vidas.

individuos una por *default*, la gente se apega a ella, a pesar del costo nulo de cambiarse de opción. Supongamos que se ofrece a los empleados de una empresa asociarse a un plan de ahorro; para esto, se lo puede asociar de manera automática, y solicitarle que se dé de baja si está en desacuerdo, o bien no asociarlo, y pedirle que se asocie si así lo prefiere. En el primer caso, la opción por default es “asociado”, y en el segundo “no asociado”. Multitud de investigaciones empíricas muestran que las reglas impuestas por default difícilmente se modifican, a pesar de que los individuos son libres de hacerlo. (Camerer, 2000; Sunstein y Thaler, 2003; Madrian y Shea, 2001)

El ejemplo precedente muestra de qué manera es posible *influir* en las decisiones de los agentes sin comprometer su libertad de elección. Conociendo que las personas ahorran de manera insuficiente, es posible ofrecerles la posibilidad de asociarlos en un plan de ahorros dejando la opción “asociado” por default. Los individuos ahorrarán más, lo cual es beneficioso para ellos. Por otra parte, se respeta su libertad de elección, en la medida que siempre pueden optar por salir del plan. Por último, cualquier opción que se tome por predefinida va a influir en la decisión. En tal sentido el paternalismo es inevitable:

“Our central empirical claim has been that in many domains, people’s preferences are labile and ill-formed, and hence starting points and default rules are likely to be quite sticky. In these circumstances, the goal should be to avoid random, inadvertent, arbitrary, or harmful effects and to produce a situation that is likely to promote people’s welfare, suitably defined. Indeed, many current social outcomes are, we believe, both random or inadvertent, in the sense that they are a product of default rules whose behavior-shaping effects have never been a product of serious reflection” (Sunstein y Thaler, 2003, p.42)

La optimalidad de los resultados es justamente lo que le da sustento normativo a la práctica paternalista. La legitimación de tal práctica ha sido intentada colocándola cerca de la tradición liberal clásica. Marumatsu y Fonseca (2009) citan a Mill:

“If either a public officer or anyone else saw a person attempting to cross a bridge which has been ascertained to be unsafe, and there were no time to warn him of his danger, they might seize him and turn him back, without any real infringement of his liberty; for liberty consists in doing what one desires, and he does not desire to fall into the river (...) Nevertheless... no one but the person himself can judge of the sufficiency of the motive which may prompt him to incur the risk: in this case, therefore... he ought, I conceive, to be only warned of the danger; not forcibly prevented from exposing himself to it (On Liberty, chapter 5, paragraph 5).”

Sin embargo, es necesario notar que las palabras de Mill se dirigen al Sistema II. El *homo intensionalis* no carece de información relevante, simplemente está atravesado por el Sistema I, y sus decisiones reflejan este hecho. Al poner una opción por *default* no se le está dando nueva información al sujeto: se está modificando el entorno para que su sistema evaluativo tome una decisión que más adelante lo va a beneficiar, aunque en el momento presente él no lo sepa. La asimetría del conocimiento es distinta en uno y otro caso: en el de Mill, el comportamiento erróneo se produce porque hay desconocimiento por parte del individuo de datos relevantes. En *Behavioral Economics*, hay desconocimiento de cómo funciona el proceso de los datos, y aunque existiera tal conocimiento, sería fútil. La presencia de dos tradiciones irreconciliables a nivel ontológico – aquella del *homo extensionalis*, o en términos de la psicología cognitiva el Sistema II, y el *homo intensionalis* o Sistema I- genera inconvenientes y equívocos cuando se trata de defender posiciones. Líneas arriba Sunstein y Thaler defienden el paternalismo sobre la base de que el ser humano es *intensional*, y defienden el carácter libertario sobre la base de que el ser humano es *libre* de decidir, es decir, es *extensional*. Este tipo de enredos muestran que las rupturas entre la teoría de la decisión convencional y los aportes de *Prospect Theory* son más profundas de las que pueden sospecharse a primera vista.

Comentarios finales

Tesis defendidas y nuevas líneas de investigación

La distinción aristotélica entre el campo de la prueba demostrativa (aquellas cosas de las que podemos tener *certeza*, como la matemática o la lógica) y el campo donde el conocimiento es meramente *probable* ha ejercido una influencia profunda y prolongada en el pensamiento occidental. Durante mucho tiempo se identificó el campo de la ciencia y la racionalidad con el del conocimiento exacto. Hubo de esperarse hasta el siglo XVII para que surgiera una concepción más débil de racionalidad, afincada en lo irreducible incertidumbre del conocimiento humano¹⁰¹. La *Teoría de la Probabilidad* emergió en la Europa iluminista como el nuevo modelo que recortaba el campo de lo racional. Aplicado a los juegos de azar, se estimó que el comportamiento razonable al elegir entre juegos alternativos pasaba por maximizar el *valor esperado*, definido a partir de los resultados y las probabilidades de obtenerlos.

La definición de la conducta aceptable en términos del valor esperado pronto enfrentó conflictos: colisionaba con la intuición de individuos razonables. La *Paradoja de San Petersburgo*, consistente en un juego de valor esperado infinito pero valuado por personas educadas en una cantidad finita (y modesta), impulsó a Daniel Bernoulli a modificar el criterio, reemplazándolo por el de la *utilidad esperada*. Los resultados de un juego no poseen un valor uniforme para todos los individuos; la riqueza, por ejemplo, condiciona la manera en la que se evalúa una cierta cantidad. La estrategia consistente en remediar un contraejemplo modificando una teoría normativa mediante la introducción de consideraciones psicológicas, dio origen a un patrón que se repetirá a lo largo de casi 300 años, con las Paradojas de Allais, Ellsberg y finalmente *preference reversals*.

El hecho de que una teoría de la conducta *racional* sea modificada a partir de la conducta *observada* de los individuos puede granjear suspicacias en el campo filosófico, porque desde Hume suele hacerse una distinción tajante entre los

¹⁰¹ Gigerenzer, 2001, p.2.

planos normativo y descriptivo, los cuales, se argumenta, discurren por niveles separados. La naturaleza del campo de las decisiones humanas, sin embargo, obliga a repensar la relación entre ambas esferas.

La relación entre lo normativo y lo descriptivo se hace explícita en *Expected Utility Theory*, surgida en 1944. John von Neumann y Oskar Morgenstern proponen en *Teoría de los Juegos y Comportamiento Económico* una teoría axiomática de la elección bajo riesgo. A partir de una serie de axiomas y supuestos normativamente defendibles se deriva la hipótesis de utilidad esperada de Bernoulli. Si las preferencias sobre loterías satisfacen dichos axiomas y supuestos, se dice que están bien formadas y son representables mediante una función de utilidad que posee la propiedad de la utilidad esperada y es única hasta una transformación afín positiva. La elección racional bajo riesgo e incertidumbre consiste en maximizar dicha función.

Por varias razones EUT ejerció un fuerte atractivo para los economistas: proponía un concepto de utilidad medible, intachable desde el punto de vista formal, despojado de conexiones incómodas con estados internos psicológicos, y enmarcado en una teoría normativamente adecuada y descriptivamente promisorio. Esta doble característica constituía el *Golden Dream* que por varias décadas pareció situar a la economía en una posición de privilegio, en virtud de la promesa de contar con una teoría definitoria de la conducta racional y al mismo tiempo poseedora de contenido empírico. Convenientemente desarrollada, EUT podría dar lugar a un programa de investigación creciente, impulsado por predicciones precisas.

Esta es una característica central de EUT: permite predicciones precisas, a diferencia de otros análisis basados en la utilidad, que son más bien generales y por lo tanto difíciles de testear. La posibilidad de testeo, por otra parte, convirtió a EUT en un objetivo claro para los críticos de la visión estándar y rápidamente comenzaron a surgir y extenderse sus dificultades empíricas, conocidas como anomalías o paradojas. Diversos intentos de disolverlas mediante los argumentos

tradicionales –falta de incentivos, o mal diseño experimental- no hicieron sino revelar la robustez de las dificultades.

De manera independiente a sus problemas empíricos, había quienes mostraban una actitud recelosa en relación a las posibilidades descriptivas de EUT por una cuestión de principio, vinculada a su falta de realismo. Los seres humanos reales, argumenta Herbert Simon hacia 1955, carecen de las capacidades computacionales hiperbólicas y el acceso a la información gratuito e inmediato que les adjudica a sus agentes EUT. La dirección en la cual debe ser reparado lo que se percibe como un error de fundamento conducía hacia la incorporación de consideraciones psicológicamente “realistas” en la construcción de alternativas para los modelos basados en agentes con racionalidad perfecta. Nace así la corriente conocida como *Bounded Rationality*, la cual dio lugar a una multitud de modelos de elección y continúa ejerciendo una enorme influencia hasta nuestros días. Entre los modelos a los que ha dado lugar se destaca la propuesta del mismo Simon, conocida como *Behavioral Model*. Hemos analizado el modelo, y encontrado que, a pesar de su enorme poder sugestivo y hermenéutico, *carece de contenido empírico*, lo que lo vuelve inadecuado como alternativa *descriptiva* a EUT.

La aparición de una clase de contraejemplo particular, conocido como *preference reversals*, significó uno de los obstáculos más importantes en la aceptabilidad de EUT como teoría capaz de describir la conducta observable. Lo que es peor, si la reversión de preferencias se mantiene como un fenómeno robusto, conceptos centrales para la economía convencional –como el de utilidad- no pueden subsistir sin cambios. Se ha argumentado que, en rigor, el concepto de utilidad convencional *desaparece* con la reversión de preferencias, algo que no parece haber sido anticipado por la profesión. Atribuido a la violación del Principio de Extensionalidad por parte de los individuos “reales”, este contraejemplo implica un final abrupto para el *Golden Dream*, porque la extensionalidad es un requisito necesario para cualquier teoría *normativa* de la decisión.

Prospect Theory, una teoría perteneciente a la corriente de *Bounded Rationality* propuesta en 1979 por dos psicólogos, David Kahneman y Amos Tversky, predice exitosamente la reversión de preferencias. Lo hace renunciando a cualquier aspiración normativa; sin embargo, se ha sostenido aquí que de hecho existe una normatividad específica en ella, vinculada al concepto de maximización, el cual suele tomarse dentro de la tradición económica como un sinónimo laxo de racionalidad. *Prospect Theory* resulta psicológicamente realista y descriptivamente exitosa en un rango significativamente amplio de circunstancias, lo cual implica que, a diferencia del modelo propuesto por Simon en 1955, posee contenido empírico.

El pasaje de EUT a *Prospect Theory* ha sido enmarcado en una visión acumulativa del conocimiento, de raíz positivista. Contra esta posición, se ha argumentado que el patrón de cambio teórico es mejor comprendido como revolucionario en el sentido de Thomas Kuhn. Cambios fundamentales en la semántica de términos como utilidad sugieren dicha tesis. En EUT la utilidad es medible, refiere a estados finales, y es un mero índice de preferencias; en *Prospect Theory* no existe método de medición, el término se predica de cambios respecto a un punto de referencia, y posee contenido hedónico. La estrategia de considerar EUT como un caso especial de *Prospect Theory* pasa por alto este hecho, y deja abierta la puerta para eventuales dificultades en la comunicación intra disciplinaria a causa del fenómeno de inconmensurabilidad.

Hemos defendido además que en el pasaje de una a otra teoría se verifica un cambio ontológico, o alternativamente un cambio en la lógica subyacente. Aquí hemos tomado ventaja del holismo de Quine para echar luz sobre este inusual fenómeno, así como sobre otros estrechamente relacionados. Fenómenos como la manipulación, el paternalismo, han sido rastreados hasta la asimetría que en el terreno lógico se abre entre el *homo extensionalis* característico de EUT y el *homo intensionalis* modelado por *Prospect Theory*. También se ha sugerido la existencia de un sentido de predicción particular, que podría estar operando al

interior de *Prospect Theory*, basado en la posibilidad de manipular la conducta humana. Por último, hemos planteado dudas respecto a la compatibilidad entre el *homo intensionalis* y la práctica eficiente de la democracia.

Restan indicar algunas líneas de investigación abiertas por el presente trabajo, las cuales sería conveniente explorar a futuro. En primer lugar, como se ha mencionado hace unos momentos, la dicotomía habitual en la literatura filosófica tradicional, cuyos orígenes se remontan a la célebre “falacia naturalista” de David Hume, distingue entre enunciados *normativos* y enunciados *descriptivos*. Para quienes hemos sido formados en esa tradición, ambas esferas constituyen categorías excluyentes y exhaustivas, en las cuales es posible distribuir de manera decidible cualquier enunciado. Contra esta suposición, sin embargo, parece haber razones para pensar que la dicotomía es *insuficiente* en Teoría de la Decisión, y ha sido reemplazada por una *tricotomía* (ver *ut. supra*, nota # 24) que diferencia entre enunciados normativos, descriptivos y *prescriptivos*¹⁰². La elaboración filosófica de esta posición no ha sido emprendida, y constituye una línea de investigación que podría ser de importancia para clarificar las distintas interpretaciones relativas a las teorías de la decisión que subyacen a los modelos económicos.

El uso *idealizante* de EUT también genera perplejidades (ver pág. 36) . De acuerdo con el mismo, la optimización como base de la decisión de los agentes en contextos donde rigen condiciones de competencia, asegura que los actores prevaecientes al final del día de mercado serán aquellos que implícitamente se hayan comportado *como si* hubieran tomado sus decisiones siguiendo los

¹⁰² La tricotomía se viene utilizando de manera efectiva desde hace un tiempo: “For our purposes, we shall augment the usual dichotomy that distinguishes between the normative and descriptive sides (the “ought” and the “is”) of decision making, by adding a third component: the prescriptive side. We do this because much of our concern in this paper addresses the question: “How can real people—as opposed to imaginary, idealized, super-rational people without psyches—make better choices in a way that does not do violence to their deep cognitive concerns?” And we find that much of what we have to say does not fit conveniently into the usual normative or descriptive niches. Loosely speaking, prescriptive analysis exploit some of the logical consequences of normative theories and the empirical findings of descriptive studies bur, in addition, something else has to be added to that is far from the spirit of normative or descriptive analyses. (Bell, Raifa y Tversky, 1988, p.9)

dictámenes de EUT. Esta es una característica particular de la economía, que comparte quizás con otras disciplinas como la biología: la presencia de *modelos de estado final* (Weisman, 2010) El propulsor de esta visión, asociada con la Escuela de Chicago ha sido Milton Friedman, en su célebre ensayo de 1953. Mäki ha estado trabajando sobre el tópico del *como si* y el realismo de Friedman en los últimos años (v.gr. 2003, 2005, 2010), desde una perspectiva puramente filosófica, vinculada a la noción de una verdad *posible* de los modelos económicos, pero los vínculos con la experiencia han estado ausentes de su núcleo de interés teórico. Es un *realismo* meramente *posible*, lo cual puede reducir su relevancia para el campo de la disciplina económica. Postular y defender un realismo *empírico* parece ser una línea de investigación sugerente en este sentido¹⁰³

Conectado con lo anterior, el *realismo* que introduce *Prospect Theory*, y que ha sido legado a *Behavioral Economics* parece ser de naturaleza muy distinta al de Simon (1955). Cabe conjeturar si la presencia de *mecanismos* bien definidos y descubiertos en laboratorio que caracteriza a *Behavioral Economics* está en la base de la diferencia, y ayuda a echar luz sobre la clase de realismo que es compatible con la capacidad predictiva de los modelos.

Esto entronca con la siguiente cuestión. Tradicionalmente, se concibió a los *modelos* como meras representaciones del mundo. El problema consistía entonces en determinar si resultaban o no adecuados (falsos, verdaderos o aproximadamente verdaderos respecto de la realidad representada). Tal problema está presupuesto en la discusión tradicional entre Friedman (1953) y los realistas de su tiempo, así como en posturas actuales como las de Mäki (2009a, 2009b), o Sugden (2000). Refleja el modo típico que tienen las ciencias naturales de relacionarse con la realidad. Una visión más “tecnológica” de los modelos económicos, como herramientas para localizar puntos de intervención efectiva

¹⁰³ El Proyecto CEREM (*Credible Empirical Realism for Economic Models*), de la mano de G. Marqués, avanza en esta dirección.

sobre la economía, es la vigente en macroeconomía (especialmente en econometría). Recientemente, se han hecho intentos por pensar epistemológica y filosóficamente este cambio de perspectiva, avanzando la idea de que la *explicación* a través de modelos se realiza sustituyendo el empleo de *leyes* por el de *mecanismos*. Elster (1998b) y Bunge (1998) tratan la noción de mecanismo como un primitivo y se concentran en examinar la forma que adoptan típicamente los argumentos que incorporan este tipo de recursos explicativos. Woodward (2000, 2002), Glennan (1996, 2002b), Machamer (2002) y Machamer *et.al.*, (2000), han abierto la “caja negra” y han procurado definir los mecanismos de una manera que involucre la noción de *regularidad*, pero no la de ley. Este tipo particular de regularidad permitiría la intervención (manipulación) a los efectos de activar el mecanismo y de obtener ciertos resultados previstos (Woodward, 2002; Leuridan, 2007). Las conexiones con la manipulación y la capacidad predictiva de *Prospect Theory* resultan sugestivas, y merecen ser ahondadas.

La extensa literatura sobre mecanismos refiere principalmente a las ciencias naturales. Sin embargo, en el ámbito de la economía, se puede conjeturar que muchos modelos no convencionales podrían ser examinados y comprendidos desde esta óptica. En particular, el análisis de los mecanismos ayudaría a clarificar la capacidad de ciertos modelos para intervenir y contribuir a conformar la realidad de un modo esperable (y muchas veces deseable). Un caso paradigmático de ello son los modelos de ahorro e inversión de *Behavioral Economics* (Benartzi and Thaler, 2001, 2007). En tanto que los modelos convencionales han favorecido principalmente la adopción de una actitud no intervencionista, la concepción “mecanicista” de los modelos podría favorecer la intervención en su forma clásica (Keynesianismo) o en la forma particular que alienta *Behavioral Economics*: el paternalismo libertario (Sunstein and Thaler, 2003), vínculos que hasta el momento no han recibido la atención que su relevancia parece exigir.

IX- Fuentes utilizadas

Akerlof, G., (1982), "Labor Contracts as Partial Gift Exchange", *Quarterly Journal of Economics* 97: 749-75.

Arrow, K., Colomatto, E., Perlman, M., and Schmidt, M., (eds), *The rational Foundations of Economic Behaviour*, Great Britain, Macmillan Press, 1999.

Baumol, William J., (1951), "The Neumann – Morgenstern Utility Index –An Ordinalist view", *Journal of Political Economy*, LIX, pp. 61 – 66.

Bernoulli, Daniel (1834) "Exposition of a New Theory on the Measurement of Risk", *Econometrica*, Vol. 22, No. 1. (Jan., 1954)

Bell, Raifa y Tversky (1988)–eds-: *Decision Making: Descriptive, Normative and Prescriptive Interactions*. Cambridge University Press.

Benartzi, Sh. and Thaler, R., (1995), "Myopic Loss Aversion and the Equity Premium Puzzle", *The Quarterly Journal of Economics*, Vol. 10, N° 1, pp. 73 – 92.

Benartzi, Sh. and Thaler, R., Naive Diversification Strategies in Defined Contribution Saving Plans, *The American Economic Review*, Vol. 91, No. 1. (Mar., 2001), pp. 79-98.

Benartzi, Sh. and Thaler, R., "Heuristics and Biases in Retirement Savings Behavior", *Journal of Economic Perspectives*, Vol. 21, n°3, Summer 2007, pp. 81-104.

Boland, Lawrence (1981), "On the Futility of Criticizing the Neoclassical Maximization Hypothesis", *The American economic Review*, Vol. 71, N° 5, pp. 1031 – 1036.

Bunge, Mario (1982), *Economía y filosofía*, Madrid, Tecnos.

(1998) *Social Science under Debate: A Philosophical Perspective*, Toronto, University of Toronto Press.

Cartwright, Nancy (1989), *Nature's Capacities and Their Measurement*. Oxford: Clarendon Press.

Cartwright, Nancy (1999) *The Dappled World*. Cambridge: Cambridge University Press.

Caldwell, Bruce, (1983), "The Neoclassical Maximization Hypothesis: Reply", *The American Economic Review*, Vol. 73, N° 4, pp. 824 – 827.

Camerer, Loewenstein and Rabin, (2003), *Advances in Behavioral Economics*, Princeton University Press.

Camerer, C., and Loewenstein, G., 2002, "Behavioral Economics: Past, Present and Future (<http://www.hss.caltech.edu/~camerer/ribe239.pdf>)

Crespo, R., y García-Huidobro, J., 2010, "Una revisión crítica sobre el carácter interesado de todas las acciones humanas", *mimeo*.

Darden, Lindley (2002), "Strategies for Discovering Mechanisms". *Philosophy of Science* 69: S354—S365.

Dahl, Robert A: *Democracy and Its Critics*, Yale University Press, 1989

Davidson, P. (2003), "The terminology of uncertainty in economics and the philosophy of an active role for government policies", en Runde, J. y Mizuhara S., ed., 2003.

Elster, J. (1996), "Rationality and Emotions". *The Economic Journal*, 106, pp.1386-97.

Friedman, Milton and Savage, L. J., (1948), "The Utility Analysis of Choices Involving Risk", *Journal of Political Economy*, LVI, pp. 279 – 304.

Friedman, Milton and Savage, L. J., (1952), "The Expected utility Hypothesis and the Measurability of utility", *The Journal of Political Economy*, Vol. LX, N° 6.

Friedman, M., (1953), *La metodología de la economía positiva*, en Ensayos sobre economía positiva, Gredos, Madrid, 1967, 14.

Giere, R., (2009), "Why Scientific Models Should not be Regarded as Works of Fiction" en In *Fictions in Science: Philosophical Essays on Modeling and Idealization*. ed. Mauricio Suárez, 248-258. Routledge,

Gigerenzer, G., 1996, "On Narrow Norms and Vague Heuristics: A Reply to Kahneman and Tversky", *Psychological Review*, vol. 103, n° 3, 592-596.

Gigerenzer, G., and Selten, R., eds., (2001), *Bounded Rationality –The Adaptive Toolbox*, Massachusetts, The MIT Press.

Glennan, Stuart S. (2002b), "Rethinking Mechanistic Explanation". *Philosophy of Science* 69: 342—S353.

Hausman, D., (1992), *The Inexact and Separate Science of Economics*, Cambridge University Press, Cambridge.

Ho, T.H., Lim, N. and Camerer, C., (2005), "Modeling the Psychology of Consumer and Firm Behavior with Behavioral Economics"

(<http://www.hss.caltech.edu/~camerer/JMRFinal.pdf>)

Hutchison, T.W: (1988), "The Case for Falsificationism", in de Marchi, N. (Eds), *The Popperian Legacy in Economics*, Cambridge: Cambridge University Press.

Koperski, Jeffrey (1998), Models, Confirmation, and Chaos, *Philosophy of Science*, Vol. 65, No. 4, pp. 624-648.

(2006), "Models", *Internet Encyclopedia of Philosophy*.

Kahneman, D, (2003a) "Maps of Bounded Rationality: Psychology for Behavioral Economics", *AER*, Vol. 93, N° 5, pp. 1449 – 1475.

Kahneman, D., (2003b), "A psychological Perspective on Economics", AER, Vol. 93, N° 2, pp. 162 - 168.

Kahneman, D. and Tversky, A., (1979), "Prospect Theory: An Analysis of Decision Under Risk", *Econometrica* 47, pp. 263-291.

Kahneman, D., (1999), "New Challenges to the Rationality Assumption", en Arrow, Colombatto, Perlman and Schmidt, (eds), *op. cit.*, pp. 203 - 219.

Kahneman, D, and Sugden, R., (2005) "Experienced Utility as a Standard of Policy Evaluation", *Environmental & Resource Economics*, 32, pp. 161 – 181.

Kahneman. D. and Riepe, M., (1998), "Aspects of Investor Psychology – Beliefs, preferences, and biases investment advisors should know about", *Journal of Portfolio Management*, Vol. 24, n° 4.

Kahneman, D., and Thaler, R., (1991), "Economic Analysis and the Psychology of Utility: Applications to Compensation Policy", *The American Economic Review*, Vol. 81, n° 2, pp. 341 – 346.

Kahneman y Thaler, "Utility Maximization and Experienced Utility", *Journal of Economic Perspectives*, Vol. 20, N°1, 2006

Kahneman, D.; Wakker,P.; Sarin, R., "Back to Bentham? Explorations of Experienced Utility", *The Quarterly Journal of Economics*, Vol. 112, No. 2, (May, 1997), pp. 375-405.

Landro, A.,(1999) Acerca de la Probabilidad, Economizarte

Lerner, S., Marqués, G. y Weisman, D., (2010), *Some remarks about scientific change – logic and philosophical aspects* (seleccionado para ser presentado en INEM 2010, November 12 – 14, Birmingham)

Lipman, B. (2002), Bounded Rationality: The Adaptive Toolbox by Gerd Gigerenzer; Reinhard Selten. (Review.) *Journal of Economic Literature*, Vol. 40, No. 3., pp. 926-927.

Mäki, U., (2008), "Scientific realism and ontology", en The New Palgrave Dictionary of Economics, Edited by Steven N. Durlauf and Lawrence E. Blume, Second Edition, Palgrave Macmillan.

(2009b), "Models and the locus of their truth", en Synthese, ed. Springer Netherlands, 13.

McEwan, M., (2006), The Semantic View of Theories: Models and Misconceptions. <http://artsweb.uwaterloo.ca/~mmcewan/LSE Presentation.pdf>

Machamer, Peter (2002), "Activities and Causation: the Metaphysics and Epistemology of Mechanisms", unpublished manuscript, presented at the 2002 *Philosophy of Science Association* Conference.

Marqués, G. *Recientes experimentos psicológicos y su impacto sobre la teoría de la utilidad esperada*, XIV Jornadas de Epistemología de las Ciencias Económicas, 2 y 3 de Octubre de 2008, FCE – UBA, Buenos Aires.

Marqués, G. *Maximización de la utilidad en la teoría estándar de la decisión: el impacto de los "framing effects"*, ECON 2008, 4 de junio de 2008, Facultad de Ciencias Económicas, UBA.

Marqués, G. "De qué sirve maximizar la utilidad esperada?", en: *Racionalidad, economía e interdisciplinariedad*, Gustavo Marqués (comp.), CIECE, Facultad de Ciencias Económicas, UBA, ISBN: 978-950-29-1182-3, Noviembre de 2009

Marqués, G. y Weisman, D., (2008), *La futilidad de criticar el modelo de Simon de elección racional*, *Revista Economía*, Universidad de Los Andes, Mérida, Venezuela, XXXIII, N° 26 (julio – diciembre, 2008), pp. 43 – 61. ISSN 1315 – 2467

Marqués, G. y Weisman, D., (2009), *Intensional Decision-Makers –An outline of their impact on Decision Theory and Economics* (mimeo)

Marqués, G. y Weisman, D., (2010), *Situational Logic as Meta Model: a Reassessment of its Theoretical Role* (enviado al Cambridge Journal of Economics)

Marqués, G. (2010), *Paternalismo Libertario: ¿escollo o último refugio para el liberalismo anti-intervencionista?* mimeo

Morgenstern, Oskar y John von Neumann (1947): The Theory of Games and Economic Behavior. Princeton University Press

Muramatsu, R. and Fonseca, P. (2009), Freedom of Choice in a World of Boundedly Rational Agents –Remarks about the light paternalistic policy implications of behavioral economics. Leído en EAEP. November 2009. (http://www.smu.ca/iarepsabe09/documents/muramatsu_fonseca-p.pdf)

Peterson, M., (2009) “An Introduction to Decision Theory”, Cambridge UP

Popper, K. (1996), *The Myth of the Framework*, Edited by M. A. Notturmo, London and New York, Routledge.

Sargent. T. (1993), *Bounded Rationality in Macroeconomics*, Oxford Univ. Press.

Scarano, E. (2008) *La Teoría Estándar de la Racionalidad: perspectivas metodológicas*, Revista: Economía, v. XXXIII, jul.-sep., 2008, pp.63-92; ISSN1315-2467 Indexada en Latindex y EconLit

Schoemaker , P. (1982) “The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations”, *Journal of Economic Literature*, Vol. 20, No. 2. (Jun., 1982), pp. 529-563.

Shefrin, H. and Thaler, R. (1988), "The Behavioral Life – Cycle Hypothesis", *Economic Inquiry*, Vol. XXVI, pp. 609-643.

Solow, R. (1991) "Sustainability: An Economist's Perspective" (Dorfman, R. y Dorfman, N.S. (eds.), *Economics of the Environment*, 3. Ed., Nueva York.)

Simon, H., (1955), "A Behavioral Model of rational Choice", *The Quarterly Journal of Economics*, LXIX, February 1955.

Simon, H. (1986), "Rationality in Psychology and Economics", en Hogarth, R., M., y

Reder, M., W., (eds), *Rational Choice -The Contrast between Economics and Psychology*, The University of Chicago Press, Chicago.

Starmer, Chris, (2000), "Developments in Non-Expected Utility Theory: The Hunt for a Descriptive Theory of Choice under Risk", *Journal of Economic Literature*, Vol., XXXVIII, pp. 332-382.

Strotz, Robert H. (1953), "Recent Developments in Mathematical Economics and Econometrics: An Expository Session – Cardinal Utility", *The American Economic Review*, Vol. 43, Nº 2, May 1953, pp. 384 – 397.

Sugden, R., (2000), "Credible worlds: the status of theoretical models in economics", en *Journal of Economic Methodology*, Taylor and Francis Ltd

Sunstein, Cass and Thaler Richard, (2003), "Libertarian Paternalism Is Not an Oxymoron", *The University of Chicago Law Review*, Vol. 70, nº 4.

Thaler, R., (1985), *Mental Accounting and Consumer Choice*, *Marketing Science*, Vol. 4, No. 3 (Summer, 1985), pp. 199-214.

(1999), *Mental Accounting Matters*, *Journal of Behavioral Decision Making*, 12: 183~206.

Thaler, Richard, 2000, "From Homo Economicus to Homo Sapiens", *Journal of Economic Perspectives*, Volume 14, Number 1, Winter 2000, pp. 133-141).

Tversky, A., (1999), "Rational Theory and Constructive Choice", in Arrow, K., Colombatto, E.,

Tversky, A. and Kahneman, D, (1981), "The Framing of Decisions and the Psychology of Choice", *Science, New Series*, Vol. 211, N° 4481, pp. 453 – 458.

Tversky, A. and Kahneman, D. (1986), "Rational Choice and the Framing of Decisions", *The Journal of Business*, Vol. 59, N° 4, Part 2, pp. S251 – S278.

Tversky, Amos and Kahneman, Daniel, (1992), *Advances in Prospect Theory: Cumulative Representation of Uncertainty*, *Journal of Risk and Uncertainty*, 5: 297-323.

Tversky, Amos, (1974), "Assessing Uncertainty", *Journal of the Royal Statistical Society, Series B (Methodological)*, Vol. 36, n° 2, pp. 148 – 159.

Tversky, Amos and Kahneman, Daniel, (1974), "Judgment under Uncertainty: Heuristics and Biases", *Science, New series*, Vol. 185, n° 4157, pp. 1124 – 1131.

Weisman, D.: Neoliberalismo y crisis del Estado en la Argentina de los noventa. Algunas consecuencias de Framing Effects para el liberalismo político y económico, (pp.57-66) FCE-UBA, ISBN 978-950-29-1161-8, Ed. Llairó-Díaz. (Capítulo de Libro)

Weisman, D.: *Del Sujeto Iluminista al Estado Paternalista*. En Ensayos sobre Economía y Racionalidad. Imprenta de la FCE, UBA. ISSN: 1851-0922. Junio 2008

Weisman, D.: *Simon (1955): los límites de la racionalidad limitada*. En Ensayos sobre Economía y Racionalidad. Imprenta de la FCE, UBA. ISSN: 1851-0922. Junio de 2008

Weisman, D.: *La teoría estándar de la decisión racional bajo riesgo*, (con G. Marqués) En Ensayos sobre Economía y Racionalidad. Imprenta de la FCE, UBA. ISSN: 1851-0922. Junio de 2008.

Weisman, D.: *La Incierta Decisión*, en Ensayos de filosofía y Metodología de las Ciencias Económicas. FCE, UBA. ISBN: 978-950-29-1060-4. Marzo de 2008

Weisman, D.: *La inutilidad de las encuestas*, en Racionalidad, Economía e Interdisciplinariedad FCE,UBA, ISBN: 978-950-29-1182-3. Mayo 2009.

Woodward, Jim (2000), “Explanation and Invariance in the Special Sciences”, *British Journal for the Philosophy of Science* 51: 197—254.

(2002), “What Is a Mechanism? A Counterfactual Account”, *Philosophy of Science* 69: S366—S377.