ESTUDIOS ECONÓMICOS
ISSN: 0425-368X
estudioseconomicos@uns.edu.ar
Universidad Nacional del Sur
Argentina

UNA REFLEXION JUEGO -TEORETICA AMPLIA PARA EL DILEMA DEL VIAJERO

Maximiliano Gabriel Miranda Zanetti[*]

enviado: Septiembre 2011- *aceptado:* Noviembre 2011

Resumen

El presente artículo analiza el Dilema del Viajero desde una perspectiva juego-teorética. El Dilema del Viajero, introducido por Basu (1994, 2007) como una interacción ejemplificadora de las fallas de la teoría a la hora de explicar el comportamiento de los agentes, presenta para el investigador sugerentes motivaciones para ampliar la visión teórica. El citado juego puede caracterizarse como un dilema del prisionero extendido en el que las estrategias para cada agente están representadas por la elección de un entero dentro de un intervalo. Llamativamente, la experimentación con este dilema arroja hechos estilizados en fuerte contraste con los resultados esperados teóricamente. El objetivo es presentar, dentro de un contexto teórico, un marco lo más amplio posible que intente echar luz sobre comportamientos de los agentes compatibles con los resultados empíricos.

Clasificación JEL: C72, D83
Palabras clave: teoría de juegos - Dilema del Viajero - economía experimental

Abstract

This article analyzes the Traveller's Dilemma from a game-theoretic point of view. The Traveller's Dilemma, introduced by Basu (1994, 2007) as an example of the failures of theory to explain the behavior of agents, presents the researcher with suggesting opportunities to increase the theoretic toolkit. The mentioned game can be characterized as an extended version of a Prisoner's

---

[*] Universidad Nacional del Sur, Departamento de Economía, IIESS, (UNS-CONICET)

Dilemma, in which the strategies for each agent are represented by the choice of an integer from an interval. Interestingly, the stylized facts revealed by experiments carried out under this interaction contrast sharply with the expected results predicted by theory. The aim of this work is to present, from a theoretical point of view, the broadest possible framework that could shed light over agents' behaviors which could be deemed more compatible with the empirical results.

JEL *Classification*: C72, D83
Keywords: game theory - Traveller's Dillema- experimental economics

---

INTRODUCCION

El propósito del presente trabajo es efectuar un análisis en términos conceptuales amplios dentro de la Teoría de Juegos de un ejemplo de juego que en la práctica (y en sus implementaciones experimentales) revela la dificultad que posee la teoría para brindar explicaciones de actitudes o elecciones de los agentes en ciertos marcos.

I. OBJETIVOS

El objetivo es intentar investigar distintas respuestas dentro del ámbito de la Teoría de Juegos que puedan en principio echar luz sobre los resultados del juego citado, sin abandonar la concepción de un agente que efectúa sus decisiones con la información de la que pueda disponer siguiendo patrones racionales o al menos predecibles de razonamiento.

II. EL JUEGO

II.1. Motivación del dilema

En un artículo de 1994, Kaushik Basu presenta un juego con el objetivo de desafiar la noción de racionalidad presente en la Teoría de Juegos. En otras palabras, introduce el dilema del viajero (en adelante, *DV*) para generar una crítica a los principios normativos habituales de la teoría.

II.2. Descripción

Como planteo inicial, consideremos el siguiente marco de interacción, que incluye la historia motivadora inicial del planteo[1]:

Dos personas, que no se conocen, han viajado por la misma compañía aérea; ambos individuos compraron en el destino turístico del que vuelven el mismo artículo. Lamentablemente los dos artículos se deterioraron en vuelo, por lo que la compañía se ve obligada a indemnizar a los turistas por el monto de la pérdida. La empresa aérea no puede conocer con exactitud el monto exacto que cada turista ha pagado por el producto, aunque sabe que debe hallarse dentro de un rango de valores posibles.

Suponiendo que ambos turistas han adquirido el bien por el mismo valor, y este importe se halla entre, digamos, 180 y 300 unidades monetarias, la empresa presenta a ambos turistas, enfrentados por separado, la siguiente propuesta: cada uno de ellos informa a la empresa de un valor a recibir en concepto de indemnización: se le otorgará a cada turista la menor de las cifras manifestada, y además, en caso de diferir los montos, se le descontará en concepto de castigo un adicional de 10 unidades monetarias a quien ha hecho la mayor oferta, mientras que se le agregará a quien ha comunicado el menor valor un premio de 10 unidades monetarias.

II.3. Formulación alternativa

Resulta evidente que, desde un aspecto juego-teorético habitual, el verdadero valor pagado por el objeto es irrelevante a la hora de predecir el comportamiento de los agentes (aunque ciertamente pueda discutirse su validez como valor de anclaje o referencia).

Para clarificar ciertos aspectos teóricos que subyacen bajo el marco de interacción, independientemente de problemas de enmarque o *framing*, permitámonos reescribir el esquema del juego en forma más abstracta:

Dos agentes, 1 y 2, tienen la posibilidad de elegir una cifra entera entre $\underline{v}$ y $\overline{v}$ sin posibilidad de comunicación entre ellos. El agente *empresa*, en función

[1] Se mencionan en esta descripción los montos 180, 300 y 10, habitualmente empleados a partir de Rubinstein (2006) en implementaciones experimentales del dilema.

de las cifras s_1 y s_2 independientemente manifestadas por los dos individuos, paga a cada uno de ellos en función de la siguiente regla:

$$\pi_i = \begin{cases} s_i & s_1 = s_2 \\ \min\limits_{j} s_j + \alpha & s_i < s_{-i} \\ \min\limits_{j} s_j - \beta & s_i > s_{-i} \end{cases}$$

La versión encuadrada precedente equivale a la aquí formulada con $\underline{v} = 180, \overline{v} = 300$, y $\alpha = \beta = 10$.

II.4. El planteo de Basu

Basu ha expuesto este juego como una crítica hacia la concepción descriptiva clásica de la teoría de juegos, en los siguientes aspectos:

- La teoría prescribe que la única decisión racional para cualquier agente es la estrategia $s_i = \underline{v}$.
- Tal estrategia constituye conjuntamente el equilibrio de Nash.
- Además de ser equilibrio de Nash, la conjunción de estrategias $(\underline{v}, \underline{v})$ es el único equilibrio racionalizable. Es más, $(\underline{v}, \underline{v})$ constituye la única alternativa viable o predecible normativamente en término de varias generalizaciones posibles del concepto de equilibrio de Nash. (Esto último se halla en cierto grado vinculado con el hecho de que el equilibrio de Nash es único.)
- El equilibrio de Nash provee a los agentes con un resultado ciertamente subóptimo, y alejarse simultáneamente de él otorga importantes beneficios a los agentes.
- Todas las puntualizaciones precedentes son válidas para cualquier valor (posible) de $\underline{v}$, $\overline{v}$, α y β (con $\alpha > 0 < \beta$).

Basu aduce que los resultados teóricos esperados son altamente contraintuitivos y no esperables en la práctica. En otras palabras, resulta evidente que los agentes poseen fuertes incentivos para mencionar un número alto, y los posibles castigos por efectuar la puja mayor son más que compensados por la posibilidad de coordinar una valuación mayor para la indemnización. En palabras de Basu, “resulta muy improbable que cualquier par de individuos, no importa cuán racionales sean ni cuán seguros puedan estar sobre la racionalidad del respectivo rival, elijan $(\underline{v}, \underline{v})$”. Basu argumenta mediante este problema de elección la dificultad inherente de la teoría de juegos para captar la lógica de decisión en ciertos procesos.

Brandenburger ha criticado, entre otros, la posición de Basu, en el sentido de que es posible ampliar el herramental de conceptos teóricos para abarcar procesos de decisión que aparentemente una visión estrecha o demasiado rígida no permite concebir.

II.5. Primeras aproximaciones teóricas

En principio, podemos esperar, siguiendo una lógica habitual juego-teorética, que si cada agente es racional, y supone que el restante agente lo es, puede esperarse que $(\underline{v}, \underline{v})$ sea el resultado de la interacción.

Esta conclusión (que exploraremos más en detalle más abajo) puede fundamentarse con las siguientes observaciones:

1. Claramente, los agentes se beneficiarían con desviaciones en el sector “alto” de la matriz de estrategias de la forma normal (esto es, conviene desde el punto de vista conjunto elegir valores altos para s_i.
2. Sin embargo, los agentes pueden descartar la estrategia $s_i = \overline{v}$ como posible opción del rival.
3. Una vez que se ha descartado la mayor estrategia como opción, mediante un proceso similar puede recortarse el resto del conjunto de estrategias, quedando solamente $\underline{v}$ como opción viable.

II.6. Resultados experimentales

Las implementaciones experimentales [solo por nombrar algunas referencias, considérense Capra et al. (1999), Becker et al. (2005), y Basu et al. (2011)] han arrojado abundante evidencia sobre la dificultad de adscribir a los agentes como única respuesta posible el resultado $s_i = \underline{v}$. Podemos caracterizar los resultados genéricamente, especificando los siguientes *hechos estilizados* de los resultados experimentales:

i. Se observa un abundante rango de dispersión en las estrategias seguidas por los participantes

ii. Las estrategias $\underline{v}$ y $\overline{v}$ están habitualmente presentes, y su distribución parece depender en general de los valores de α y β.

iii. El patrón general parece indicar que la distribución de respuestas altas disminuye con el valor de premio y castigo, y para altos valores de α y β, las respuestas colapsan hacia $\underline{v}$.

iv. Los resultados pueden estar influenciados por problemáticas de información, racionalidad y comprensión del problema. Sin embargo, no podemos evitar mencionar el hecho de que la experimentación con grupos con efectivo conocimiento de conceptos de teoría de juegos [Becker et al. (op. cit.)] arroje resultados similares a los de otras poblaciones experimentales.

III. ANALISIS

III.1.Visiones tradicionales

III.1.1. Equilibrio de Nash

El concepto de equilibrio de Nash involucra, como es sabido, un punto fijo de mejores respuestas dentro del espacio de estrategias de (en este caso) los dos agentes.

La proposición de Basu referente a que la solución $(\underline{v}, \underline{v})$ es el único resultado esperado por la prescripción teórica debe cualificarse de la siguiente forma: en primer término, $(\underline{v}, \underline{v})$ es el único equilibrio de Nash. Además, es el único punto de interés que sobrevive una serie de criterios habituales que funcionan como extensión del concepto de equilibrio de Nash; ya nos referiremos a ellos en el siguiente apartado.

Procuremos profundizar el siguiente concepto: es cierto que la combinación $(\underline{v}, \underline{v})$ constituye el único equilibrio de Nash, pero bajo la condición de que los dos individuos valoren exclusivamente su propio bienestar a la hora de efectuar su elección. En otras palabras, la concepción de Basu hace referencia a un juego en que la función de utilidad de los agentes depende en forma creciente de su exclusivo pago (y no ingresen en ella factores como la justicia del resultado, aspectos de cooperación, externalidades referidas al bienestar del restante individuo, etcétera).

Resulta evidente, entonces, que un campo de exploración teórica (aunque quizá no tan fácilmente implementable en su concepción explicativa, la verdad sea dicha) está dado en la posibilidad de incorporar otras formas funcionales más abarcativas que permitan exponer la posible racionalidad del comportamiento de los agentes. Volveremos a este punto en el parágrafo III.3.2.

III.1.2. Racionalizaciones semejantes

Como se precisa en el apartado precedente, existen conceptos teóricos que buscan ampliar el conjunto posible de acciones de los agentes, abarcando la solución de equilibrio pero permitiendo que otras opciones comprensibles sean abarcadas dentro de la conceptualización racional.

Un problema en la exploración de este camino es que algunas de estas soluciones consisten en el refinamiento del concepto de equilibrio, cuestión que en nuestro caso se ve obstaculizada por el hecho de que existe (al menos bajo la concepción utilitarista descripta arriba) un único equilibrio de Nash.

Aun así, podemos explorar una serie de conceptos de mayor amplitud que el equilibrio que pueden echar luz sobre la dicotomía teoría/resultados empíricos.

Eliminación de estrategias dominadas

Luego de considerar la noción de equilibrio de Nash, el primer concepto que podremos explorar es el de la eliminación o descarte de estrategias dominadas.

Como es sabido, el juego reducido a partir del original por medio de eliminación iterada de estrategias (débilmente) dominadas contiene el equilibrio de Nash. Lamentablemente, en nuestro caso, al menos considerando la formulación de utilidad inicial, no hacemos otra cosa que reducir la matriz a la única celda $(\underline{v}, \underline{v})$. En otras palabras, la estrategia $s_i = \overline{v}$ se halla débilmente dominada por $s_i = \overline{v} - 1$, y una vez descartada la estrategia dominada en primer término, pueden ser descartadas una a una todas las estrategias alternativas hasta restar solo $s_i = \underline{v}$.

Eliminación de estrategias estrictamente dominadas y racionalidad

Un concepto en principio más amplio hace referencia a la eliminación de estrategias estrictamente dominadas; el conjunto mutuo de estrategias para ambos individuos que han sobrevivido la eliminación iterada de estrategias dominadas en el sentido fuerte constituye el grupo de posibles equilibrios racionalizables (bajo nuestro marco de interacción, con dos agentes, esto equivale el conjunto de alternativas que incorporan para cada individuo una posible mejor respuesta a una creencia racional de estrategias del rival).

Desafortunadamente, también aquí nos encontramos con un resultado conocido: si bien la estrategia $s_i = \overline{v}$ no es estrictamente dominada por, digamos, $\tilde{s}_i = \overline{v} - 1$, nos encontramos con el siguiente resultado:

Sea $s_0 = \overline{v}$, y $s_i = \overline{v} - i$. Entonces, puede verse que la estrategia s_0 se encuentra estrictamente dominada por la estrategia mixta con pesos $\left(\frac{\epsilon^i}{\sum_j^{\overline{v}-\underline{v}} \epsilon^j}\right)_{i=0}^{\overline{v}-\underline{v}}$ para un valor de ϵ convenientemente elegido.

De forma iterativa, podemos ir eliminando por el mismo procedimiento las estrategias s_1, s_2,… hasta reducir la matriz a $(\underline{v}, \underline{v})$.

III.2. Discusión de los hechos estilizados

Retomemos los conceptos referenciados en §II.6: es evidente que la formulación estándar, estudiada a través de los conceptos habituales de solución o racionabilidad, no nos permiten predecir el comportamiento de los individuos, o al menos algunos de sus rasgos centrales.

Resulta indiscutible que la opción solución $\overline{v}$ solo pareciera estar justificada si consideramos alternativas para la valoración de los resultados (alguna concepción diferente de la función de utilidad). Por otro lado, la presencia de la opción por $\underline{v}$ indica que parte de la población bajo análisis parece estar representada por el proceso racional habitual descripto en §III.1.

Finalmente, resta también explicar dos aspectos: la amplia diversidad de respuestas estratégicas (quizá explicable con la inclusión de diferencias intrínsecas, o variedad de tipologías de agentes) y cierta dispersión sobre las respuestas focales, especialmente en el rango $290 - 300$ [bajo los valores canónicos indicados supra], fenómeno quizá justificable bajo apreciaciones de acotamiento de racionalidad, o consideraciones epistémicas. Volveremos a estas cuestiones más adelante.

III.3. Algunas propuestas explicativas

Esbozaremos aquí tres ítems que permiten abarcar, sin abandonar del todo la condición de una cierta caracterización de racionalidad, algunos de los aspectos presentes en las respuestas elegidas por los sujetos experimentales.

III.3.1. Reacciones y creencias

Una primera conceptualización de los requisitos de racionalidad más exiguos posibles que nos permitan explicar resultados más allá del equilibrio de Nash es la siguiente:

Sea una función de probabilidad $f(\cdot)$ definida sobre el soporte $[\underline{v};\overline{v}]\backslash\mathbb{Z}$, la *creencia*, que indica la estimación del agente sobre la probabilidad que el asigna a cada posible estrategia del rival.

Sea una función de utilidad u que determina la valoración de los resultados dados por la función de pagos π bajo estrategias s_i, s_{-i} (propia y rival), $u(\pi(s_i, s_{-i}))$.

Sea una operación $E(f,\sigma)$ definida para cualquier creencia f y aleatorización σ elegida por el individuo, que determina la *expectativa* del resultado de σ. (La expectativa puede ser, como es el caso habitual, la *utilidad esperada*:

$$\sum_i \sum_j \sigma_i f(s_j)\, u\left(\pi(s_i, s_j)\right)$$

Proponemos el requisito más simple posible bajo este ámbito en términos de racionalidad: simplemente llamaremos a un agente *racional* si, dadas su creencia f, su utilidad u y su expectativa E, decide efectuar una aleatorización $\tilde{\sigma}_i$ que maximiza $E(f,\sigma)$ sobre todas las posibles loterías por optar.

Ejemplo 1

Sea un agente que estima que el restante individuo elegirá su estrategia totalmente al azar, es decir, $f(\cdot) = \frac{1}{\overline{v}-\underline{v}+1}$. Consideremos que el agente guía su proceder mediante las expectativas usuales, empleando la fórmula de utilidad esperada. Su función de utilidad que emplea para evaluar los pagos es la identidad, es decir, $u(\pi(s_i, s_{-i})) = \pi(s_i, s_{-i})$.

¿Qué estrategia empleará un agente "racional" bajo estas hipótesis?

Considerando los valores canónicos $\underline{v} = 180$, $\overline{v} = 300$, y $\alpha = \beta = 10$, el resultado a primera vista sorprendente es que un individuo racional, bajo los supuestos del caso, optará por la estrategia de anunciar un valor $s_i \in \{280; 281\}$.

La elección, quizá sorpresiva considerando los considerables montos del premio y el castigo, puede entenderse si se pondera el hecho de que disminuir la estrategia a partir de s_i en una unidad solo tiene efectos sobre el pago bajo la interacción contra una estrategia $s_{-i} = s_i - 1$, mientras que disminuye los pagos toda vez que se enfrente a una estrategia $s_{-i} > s_i$.

Este ejemplo nos permite concluir que no es indispensable más que que se pueda racionalmente suponer cierta probabilidad de respuestas altas para que un agente racional (bajo la concepción expresada arriba) y totalmente desinteresado pueda optar por una estrategia relativamente alta.

Ejemplo 2

Consideremos un agente 1, caracterizado por igual función de utilidad y expectativas que en el ejemplo precedente. Supongamos que las creencias del individuo siguen los siguientes lineamientos: el agente valúa dos posibles tipos de individuo 2 con los que puede interactuar: 1 valúa con una probabilidad γ que el agente 2 aleatorice uniformemente su elección; asimismo, estima con una probabilidad $1 - \gamma$ que el individuo 2 jugará $s_2 = 280$.

Resulta que para todo valor de γ contenido en $\left[0; \frac{1089}{1090}\right)$, la mejor respuesta para nuestro agente racional es elegir $s_1 = 279$.

Concretamente, este ejemplo nos muestra que es posible considerar variadas concepciones de creencias o niveles de conocimiento que nos permita describir comportamientos de dispersión en rededor de respuestas *focales*.

Un aspecto a destacar de la visión que hemos expuesto en esta sección es la diferenciación de esta postura con el análisis tradicional. En otras palabras, ¿en qué consiste la diferencia entre esta caracterización y las prescripciones normativas que habíamos expuesto en §III.1.

Como podrá adivinar el lector, el paso adicional que la visión tradicional espera del agente es la consistencia de sus creencias con el proceso de análisis racional en base con su conocimiento del rival (utilidad…), el conocimiento del conocimiento del rival, y demás encadenamientos inductivos bien tratados por la correspondiente subárea juego-teorética [véase el buen análisis de Osborne y Rubinstein (1994) sobre conocimiento en teoría de juegos].

Al respecto, y simplemente para efectuar una argumentación de por qué puede ser interesante relajar esta estructura de conocimiento e información, tomemos un ejemplo bien conocido en la literatura [Osborne y Rubinstein (op. cit.)]: El llamado juego del correo electrónico, que ilustra en forma sucinta la profundidad del el tópico de excesivas suposiciones sobre la racionalidad y conocimiento de los agentes.

Consideremos dos agentes, *Fila* y *Columna*, que deben interactuar en dos posibles situaciones, para cada una de las cuales puede asignarse una matriz de pagos que describa las utilidades de los agentes. Consideremos la siguiente matriz, que describe los pagos para los jugadores bajo el subjuego o situación *A*, de acuerdo con la estrategia que elijan de las dos posibilidades: *Entrar* o *Salir*.

Tabla 1 - Pagos bajo el subjuego *A*

$Fila \backslash Columna$	Salir	Entrar
Salir	$(7,7)$	$(1,-10)$
Entrar	$(-10,1)$	$0,0$

Como se puede observar, este subjuego presenta un único equilibrio dado por la conjunción (*Salir*, *Salir*).

Tabla 2 - Pagos bajo el subjuego *B*

$Fila \backslash Columna$	Salir	Entrar
Salir	$(0,0)$	$(1,-10)$
Entrar	$(-10,1)$	$(7,7)$

Puede observarse en este caso que el subjuego presenta dos equilibrios: (*Salir*,*Salir*) y (*Entrar*,*Entrar*), por lo que es válido describir esta subsección como un juego de coordinación; sin embargo nótese que se trata de una coordinación intricada, ya que el equilibrio dominador en el sentido de Pareto exige a ambos individuos coordinar la entrada, estrategia que de no ser correspondida presenta a los jugadores amplios costos.

Si los jugadores deben elegir entre salir y entrar, sin conocer en qué subjuego se hallan, resulta obvio que la estrategia salir puede resultar tentadora.

De hecho, aun si el subjuego que se ha de jugar es conocido por el jugador *Fila*, y este no puede comunicárselo al restante agente, el equilibrio bajo el juego en que se sortea el subjuego B con probabilidad γ tiene como único equilibrio (bayesiano) (*Salir*, *Salir*) para valores de γ menores a $1/2$.

Analicemos la siguiente complicación del juego: El jugador *Fila* puede decidir enviar un mensaje anunciando que el subjuego es el B, el que será procesado por un protocolo de mensajes. El protocolo tiene el siguiente funcionamiento: si *Fila* decide enviar el mensaje, se remite a una terminal a la que tiene acceso *Columna* el mensaje original; la terminal de *Columna* reenvía un mensaje de confirmación de recibo cada vez que le llega un envío de la terminal de *Fila*, y lo propio hace esta última terminal; cada mensaje y acuse de recibo tiene una probabilidad ϵ de no llegar a destino, pequeña pero no nula. Así es que todo proceso de comunicación termina ciertamente en una determinada cantidad de mensajes y acuses intercambiados finita. Supongamos que los agentes tienen acceso al número final y total de mensajes que han arribado a cada terminal.

Sea por ejemplo la situación en que el agente *Columna* conoce que su terminal ha recibido 16 mensajes (el original y 15 acuses adicionales). Resulta más que claro que ambos jugadores conocen que el subjuego que se está llevando a cabo es el B: Es más, cada uno sabe que el restante jugador conoce este hecho. Sin embargo, el único equilibrio bayesiano es la situación en la que cada jugador elige *Salir*[2].

El resultado precedente ilustra la importante restricción que el requerimiento de eliminación de posibles jugadas del rival con base en

[2] La demostración de este resultado [véase Osborne y Rubinstein (op. cit)] depende simplemente de un proceso iterativo de descarte basado en la inducción matemática.

exigencias del uso de la información y deducción provoca sobre las posibles elecciones estratégicas de un individuo.

Es cierto que, en nuestro caso, el enfoque racional amplio que hemos desarrollado en esta subsección permitiría explicar una gran parte de los resultados empíricos; sin embargo, no podemos dejar de notar que el elegir la estrategia $\overline{v}$ no puede ser justificado aún por esta vía. Accederemos en el parágrafo siguiente a la profundización de algunos aspectos que pueden echar luz sobre el particular.

III.3.2. Solidaridad, interdependencia y nociones de justicia

Como se habla más arriba, una posible noción subyacente a la elección de valores intrínsecamente altos puede ser la conceptualización de la decisión considerando, además del bienestar individual, otros aspectos, tales como nociones de justicia o equidad, de desprecio de la traición o ausencia de cooperación, efectos externos causados por la evaluación del bienestar ajeno, etc.

De hecho, pareciera que al menos alguna de las nociones del párrafo previo es imprescindible para explicar los resultados experimentales obtenidos: como se dijo arriba, podemos concebir fácilmente una explicación de por qué un agente individualista elija números elevados, si espera que el restante agente se comporte de forma parecida; empero, no podemos brindar una explicación convincente de por qué algún agente elige $s_i = \overline{v}$ si no consideramos alternativas a la función de utilidad $u\big(\pi(s_i, s_{-i})\big) = \pi(s_i, s_{-i})$.

Gintis et al. (2003) constituye un muy interesante aporte, relacionando aspectos de retribución e interdependencia en ámbitos de interacción semejantes al descripto en el *DV*: por ejemplo, el conocido dilema del prisionero.

Existen varias formas posibles de incluir estos aspectos en una forma funcional útil que podamos aplicar a nuestro problema y permita la justificación (o mejor dicho, la explicación) de la elección de, por ejemplo, $s_i = \overline{v}$.

Un problema asociado a este enfoque es la capacidad predictiva del mismo, toda vez que la estipulación de (los parámetros de) las formas funcionales extendidas en los sentidos señalados debe previamente calibrarse con datos reales, o definirse de acuerdo con algún criterio previamente justificable, de ser posible.

Estudiemos, aunque más no sea brevemente, algunas formulaciones simples que puedan permitir aventurarnos a conjeturar la posibilidad de respuestas coordinadas:

Fórmulas lineales

1. $u_1(\pi_1, \pi_2) = \pi_1 + \eta \cdot \pi_2, \qquad \eta > 0.$
2. $u_1(\pi_1, \pi_2) = \pi_1 + \zeta \cdot |\pi_1 - \pi_2|, \quad \zeta > 0.$

Fórmulas no lineales

3. $u_1(\pi_1, \pi_2) = \pi_1 \cdot \pi_2^{\phi}, \qquad \phi > 0.$
4. $u_1(\pi_1, \pi_2) = \pi_1 \cdot |\pi_1 - \pi_2|^{\xi}, \qquad \xi > 0.$

Fórmulas como las precedentes nos permiten modelizar comportamientos cooperativos que involucren resultados de cooperación. Por ejemplo, de la fórmula dada en el punto **3** se generan los siguientes resultados de interés:

- $\hat{\phi}$, El valor de ϕ a partir del que se genera la solución cooperativa $s_i = \overline{v}$, puede computarse fácilmente en $\hat{\phi} = \frac{\ln(\overline{v}+\alpha-1)-\ln\overline{v}}{\ln\overline{v}-\ln(\overline{v}-\beta-1)}$.
- $\hat{\phi}$ presenta interesantes propiedades: por ejemplo, $\frac{\partial\hat{\phi}}{\partial\alpha} > 0, \frac{\partial\hat{\phi}}{\partial\beta} > 0.$
- Putativamente, también $\frac{\partial\hat{\phi}}{\partial\overline{v}} > 0$, lo que concuerda en gran medida con los resultados experimentales.

Estas consideraciones son sumamente interesantes, puesto que coinciden con los hechos estilizados de la experimentación.

III.3.3. Heterogeneidad

Otro aspecto que parece deducirse de los hechos estilizados comentados previamente es la posible justificación de la consideración de diversos tipos o caracterizaciones de agentes: en otras palabras, la consideración de distintas tipologías de agentes (por otro lado, supuesto de trabajo considerablemente razonable, dada la amplia dispersión de las respuestas halladas).

III.4. Propuestas epistémicas adicionales

III.4.1. Descripción general

Los presupuestos de la teoría de juegos epistémica [Brandenburger (2007a, 2007b)] involucran efectivamente consideraciones generales de problemáticas de conocimiento relacionadas con algunos tópicos esbozados en el trabajo.

Esta teoría se aboca centralmente a modelizar explícitamente cualificaciones o tipificaciones sobre los distintos posibles grados de conocimiento, entendimiento o información que puedan poseer los agentes.

No se hará aquí un desarrollo de las temáticas de la teoría epistémica; describiremos empero dos caminos que pueden resultar interesantes ampliaciones para nuestra modelización.

III.4.2. Enfoques de números limitados de pasos de razonamiento

El tópico que describiremos aquí hace referencia a un grado acotado de razonamiento. Podemos concentrarnos, por ejemplo, en la modelización de agentes con k grados de capacidad cognitiva deductiva.

El agente económico racional es supuesto con una capacidad cognitiva $k = \infty$. Por el contrario, un agente con una capacidad $k = 1$ podrá deducir el carácter de dominado de una estrategia como $\overline{v}$, pero no necesariamente el de $\overline{v} - 1$.

Esta concepción permite explicar o abordar el comportamiento de agentes para los cuales el juego relevante (esto es, la matriz de la que extraerán su estrategia) es comparativamente reducido, en rededor de alguna respuesta focal o natural.

III.4.3. Modelos categóricos

Se explicita aquí un tópico de la concepción epistémica relacionado con el precedente, y que puede permitir modelizar categorías o dispersiones de respuestas alrededor de puntos focales.

El concurso de belleza [véase el enfoque epistémico de Camerer et al. (2004)] constituye un juego con complicaciones semejantes al *DV*, en términos de discrepancias entre las respuestas halladas y el equilibrio de Nash derivado de las suposiciones habituales de conocimiento, capacidad de razonamiento y racionalidad.

Se puede pensar en categorías de agentes con las siguientes características:

Existen agentes de categoría o inteligencia cero, τ_0, que eligen su estrategia de alguna manera trivial, no estrictamente inferida mediante razonamiento; por ejemplo, pueden aleatorizar su respuesta, o concentrarse en una estrategia focal.

Luego, se hallan presentes agentes de tipo de inteligencia creciente, τ_1, τ_2,… cada uno de los cuales supone la existencia de los tipos precedentes.

La distribución de tipos τ_i puede asumirse gobernada por una función de probabilidad $f(i)$ sobre $\{0; 1; 2 \dots\}$.

Para $i > 0$, el agente de tipo τ_i presupone que se enfrentará con un agente de tipo τ_{i-1} o inferior, e inferirá que la distribución de posibles agentes rivales τ_j estará dada por la probabilidad condicionada $g(k) = \frac{f(k)}{F(i-1)} = \Pr(\tau_j = \tau_k | j < i)$. Una modelización habitual emplea para $f(\cdot)$ la distribución de Poisson de parámetro λ: $f_\lambda(i) = \frac{e^{-\lambda}\lambda^i}{i!}$.

En el presente caso, la implementación más sencilla del modelo permitiría predecir cierto comportamiento dado por el hecho de que un agente de tipo τ_i elegirá una estrategia muy cercana (pero separada al menos en una unidad) de la estrategia del tipo precedente.

IV. Discusión

Se ha desarrollado en el presente trabajo todo un conjunto de propuestas que se contraponen a la proposición de que la única solución viable indicada por la teoría es $(\underline{v}, \underline{v})$.

Resulta necesario intentar incorporar al herramental teórico conceptualizaciones que nos permitan vislumbran intencionalidades presentes en el razonamiento de los individuos experimentales; este proceder nos brindará la posibilidad de ampliar el poder predictivo de la modelización, sin necesariamente abandonar el análisis positivo del comportamiento racional humano, y quizá avanzando en el necesario camino de explorar limitaciones plausibles del proceso inductivo de los jugadores.

Se ha presentado varias vías de incorporación de conceptos al análisis que permitirían extender el campo predictivo de la teoría.

Por ejemplo, se puede comenzar por considerar un núcleo del procedimiento racional dado por elección basada en procesos de optimización con base en ciertas creencias, teniendo en consideración que el desarrollo de esas creencias puede seguir variados procedimientos, no siendo exclusivamente la inferencia con capacidad ilimitada la única posibilidad.

Se explora esta visión racional “primigenia” en la sección III.3.1. Resulta interesante comparar esta visión con la correcta predicción que efectúan los agentes (o al menos un buen número de ellos) sobre el comportamiento del rival [véase, por ejemplo, Becker et al. (2005)]; esta predicción pareciera indicar que este modelo sencillo de racionalidad basado en la creencia de comportamiento del rival parece explicar una parte importante del razonamiento de los agentes.

Es posible acoplar además al análisis la posibilidad de heterogeneidad en las motivaciones de los agentes. Eso permitiría tal vez explicar un gran conjunto de los hechos estilizados indicados en §II.6.

A su vez, parte de esta heterogeneidad puede involucrar la interacción de agentes relativamente individualistas con un cierto número de agentes con visiones distintas, como las modelizadas por medio de formalizaciones de utilidad similares a las indicadas en §III.3.2.

Se finaliza el desarrollo incorporando concepciones epistémicas tales como las explayadas en la sección III.4. Especialmente interesante parece ser el proceso explicado en §III.4.3, que permite explayar la concepción de racionalidad comentada en la sección III.3.1 con una especificación interesante, aunque algo ingenua.

En respuesta a la crítica de Basu, y en consonancia con la defensa llevada a cabo por Brandenburger y otros, puede concluirse que la Teoría de Juegos posee herramental y margen de desarrollo suficiente para abordar e intentar explicar fenómenos empíricos que contradicen las posiciones más tradicionales y en cierta forma simplistas de la teoría estándar.

Basu, K.., (1994), “The Traveler's Dilemma: Paradoxes of Rationality in Game Theory”, *American Economic Review* 84.

Basu, K., (2007), “The Traveler's Dilemma”, *Scientific American* 296.

Basu, K., Becchetti, L. y Stanca, L., (2011), “Experimens with the traveler's dilemma: Welfare, strategic choice and implicit collusion”, *Social Choice and Welfare* (en prensa).

Becker, T., Carter, M. y Naeve, J., (2005), “Experts Playing the Traveler’s Dilemma”, Manuscrito.

Brandenburger, A., (2007a), “The power of paradox: some recent developments in interactive epistemology”, *International Journal of Game Theory* 35.

Brandenburger, A., (2007b), “Epistemic Game Theory”, En: *The New Palgrave Dictionary of Economics*, segunda edición, ed. por Durlauf, S. N. y Blume. Palgrave Macmillan. Norton, Nueva York.

Camerer, C., Ho, T. y Chong, K., (2004), “A Cognitive Hierarchy Model of One-Shot Games”, *Quarterly Journal of Economics* 119.

Capra, C., Goeree, J., Gomez, R. y Holt, C., (1999), Anomalous Behavior in a Traveler’s Dilemma? *American Economic Review* 89.

Fudenberg, D. and Tirole, J., (1993), *Game Theory*. MIT, Cambridge, Massachusetts.

Gintis, H., Bowles, S., Boyd, R. y Fehr, E., (2003), “Explaining altruistic behavior in humans”, *Evolution and Human Behavior* 24.

Mas-Colell, A., Whinston, M. D. and Green, J. R., (1995), *Microeconomic Theory*, Oxford University Press, New York.

Osborne, M. J. and Rubinstein, A., (1994), *A Course in Game Theory*. Second Impression (1995), MIT, Cambridge, Massachusetts.

Rubinstein, A., (2006), “Dilemmas of an Economic Theorist”, *Econometrica* 74.