

Una introducción a la metodología experimental en economía

Enrique Fatás*

José M. Roig**

LINEEX y Universidad de Valencia

RESUMEN

El objetivo de este trabajo es analizar la metodología experimental en economía, identificando los elementos esenciales de todo experimento económico. Los tres objetivos esenciales de la economía experimental son analizados en profundidad a partir de algunas de las aportaciones esenciales de esta técnica. Por último, se presentan las principales críticas que la utilización de laboratorios experimentales ha recibido, así como las respuestas con que los experimentalistas han intentado rebatirlas.

1. UNA BREVE INTRODUCCIÓN HISTÓRICA

La economía, tal y como es actualmente enseñada y aprendida en las escuelas de negocios y universidades, es más intensiva en teoría y menos intensiva en observaciones que quizás ninguna otra ciencia. Probablemente nuestro entrenamiento como economistas nos condiciona a pensar en la economía como una ciencia apriorística y

¹ LINEEX y Universidad de Valencia, *fatás@uv.es*, Facultad de Economía, Campus Tarongers 46022 Valencia (España).

² LINEEX y Universidad de Valencia, *roigjm@uv.es*, Facultad de Economía, Campus Tarongers 46022 Valencia (España).

no observacional en la que la interrelación entre teoría y observaciones es poco efectiva. La conclusión es que acabamos pensando que los problemas económicos pueden ser interpretados básicamente pensando en ellos. Con el suficiente rigor científico, la adecuada coherencia interna y el conveniente consenso académico, el pensamiento de los economistas parece ser una herramienta lo suficientemente potente como para enfrentarse al mundo real.

El uso del enfoque deductivo no es por supuesto un enfoque exclusivo de la economía. La teoría siempre organiza nuestro conocimiento y nos ayuda a predecir el comportamiento en situaciones nuevas. En particular, la teoría nos dice qué datos merece la pena recoger y nos sugiere cómo analizar las nuevas observaciones. A medida que la teoría progresa, nos guía en el proceso de refinar los datos obtenidos y en seleccionar las preguntas que deberíamos hacernos. Por su parte, la recolección de los datos y el análisis de los mismos a menudo descubren regularidades que no son explicadas por la teoría existente en esos momentos. Esta interacción entre teoría y el trabajo empírico, a menudo mediante ajustes menores y a veces por medio de revoluciones, es el verdadero motor de progreso científico en cualquier disciplina científica.

Sin embargo, en la aproximación tradicional a la investigación económica, la observación de fenómenos económicos que suceden de forma natural ha sido la única fuente de datos a la que se recurría para estimular la revisión de la teoría. Si los datos relevantes a una proposición económica no eran directamente observables en un ambiente económico natural, entonces la proposición simplemente se quedaba sin los beneficios del contraste empírico. En los años recientes, sin embargo, los métodos experimentales han dado a los economistas acceso a nuevas fuentes de datos y han ampliado el conjunto de proposiciones económicas en los que los datos pueden ser consultados.

Roth (1995) establece como antecedente más remoto de los experimentos en ciencias sociales el trabajo desarrollado por los hermanos Bernoulli sobre la llamada Paradoja de San Petersburgo cuando recurrieron a encuestar a distinguidos colegas acerca de la resolución de ese problema. Probablemente la prehistoria de la economía experimental pueda situarse en los años 30 cuando Thurstone (1931) analizó experimentalmente el problema de determinar las curvas de indiferencia individuales haciendo elegir a sus sujetos experimentales entre cestas de bienes hipotéticas de manera sistemática. Pero dos acontecimientos en los años 40 fueron probablemente los que desencadenaron el verdadero pistoletazo de partida de la experimentación en economía. El primero, la publicación de la piedra angular de la teoría de juegos, la obra de von Neumann y Morgenstern (*Teoría de juegos y comportamiento económico*) en 1944 que dota al contexto estratégico, propio de todo experimento, de carta de

plena naturaleza. El segundo, la realización de los primeros experimentos de mercado propiamente dichos en la Universidad de Harvard por Edward H. Chamberlin, aunque sus consecuencias tuvieran un efecto retardado.

La publicación de la obra de von Neumann y Morgenstern provoca el interés en la realización de experimentos individuales en una serie de académicos que desemboca en el trabajo de Allais (1953) sobre la violación de la teoría de la utilidad esperada en decisiones bajo riesgo que terminó conociéndose como la Paradoja de Allais. De manera independiente Dresher y Flood realizaron en 1950 una serie de experimentos en la Rand Corporation con la colaboración de John Nash que iban a tener una enorme influencia posterior, al realizar el primer test empírico de lo que más adelante se conocería como el dilema del prisionero. En 1952 tuvo lugar la primera conferencia académica sobre *El diseño de Experimentos en procesos de decisión*, financiada por la Fundación Ford y la Universidad de Michigan, que se celebró en Santa Mónica. Por primera vez se presentan resultados experimentales sobre juegos de negociación extraídos de entornos económicos donde los pagos de los sujetos eran determinados directamente por sus acciones y las acciones de los demás.

Los experimentos de Chamberlin (1948) tuvieron una doble virtud. En primer lugar, suponen el primer experimento de mercado registrado, tal y como el autor establece en su introducción al afirmar que *el propósito de este artículo es el de (...) describir un experimento real con un mercado bajo condiciones de laboratorio y establecer algunas de las conclusiones en él obtenidas* (página 95). Chamberlin creó un mercado experimental informando a cada comprador y vendedor de su precio de reserva de una mercancía (esto es, para cada comprador el precio por debajo del cual pueden comprar con beneficio y para cada vendedor el precio por encima del cual pueden vender con beneficio), en el que compradores y vendedores podían interactuar libremente para llevar a cabo transacciones de manera totalmente descentralizada. Dados los precios individuales de reserva de compradores y vendedores, se podían construir las curvas de demanda y oferta del mercado y, por lo tanto, elaborar de manera inequívoca una predicción sobre el equilibrio resultante.

La segunda virtud de los experimentos de Chamberlin tardó años en ponerse de manifiesto, y tiene que ver con el entusiasmo que despertó en los participantes, uno de los cuales habría de convertirse, primero en la Universidad de Purdue y luego en la de Arizona, en el padre de la economía experimental, y principal responsable de la siguiente generación de experimentalistas: Vernon Smith.

Si bien ya desde los primeros experimentos de Chamberlin y Allais, la economía experimental fue creciendo de manera exponencial, fue precisamente esa nueva generación de experimentalistas (en EEUU: Charles Poltt en Caltech, Charles Holt en

Virginia, Alvin Roth en Pittsburgh; en Europa, los «discípulos» del más tarde Nobel Selten: John Hey en York, Werner Guth en Berlin, por citar sólo algunos ejemplos) la que alumbró la siguiente fase de expansión. Esta expansión tiene que ver con la creación a mediados de los 80 de los primeros laboratorios experimentales plenamente informatizados, que permitieron a la investigación experimental resolver algunos de los problemas técnicos ligados a la validez interna de los experimentos por el nivel de sofisticación del diseño experimental. Esta sofisticación permitió emplear esta metodología en la práctica totalidad de los campos de investigación económica.

A pesar de esa expansión de la economía experimental por la mayoría de los centros de investigación en economía del mundo, y de la carta de plena naturaleza que supone la concesión en 2002 del premio Nobel de Economía a Vernon Smith (y al también experimentalista y psicólogo Danny Kahneman), esta metodología todavía resulta novedosa y lejana a buena parte de los economistas. El objetivo del presente trabajo no es el de realizar un recorrido por los principales resultados obtenidos por la economía experimental en las últimas dos décadas³, sino realizar una aproximación a esta metodología que cubra ese vacío. Para ello, recorreremos el tipo de evidencia experimental y los principios esenciales de esta metodología en el punto 2, analizaremos sus objetivos esenciales en el 3 y responderemos a las críticas más habituales que despierta este método en el punto 4. La sección 5 concluye.

2. LA METODOLOGÍA EXPERIMENTAL

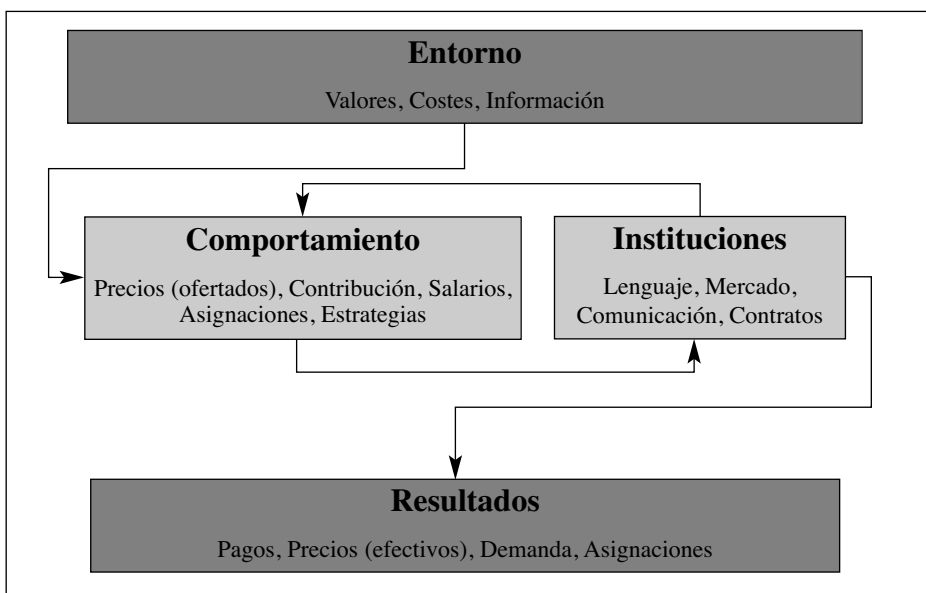
2.1. *Los principios de la Economía Experimental*

Podemos definir los componentes básicos de la economía experimental recurriendo a Smith (1994) cuando identifica tres ingredientes básicos en todo experimento económico: el entorno, las instituciones y el comportamiento. El entorno especifica los recursos iniciales, las preferencias, los costes que pueden motivar el intercambio entre los agentes en el experimento; las instituciones definen la manera en la que los sujetos pueden interactuar, el método por el que se intercambia la información y las normas bajo las que estos mensajes tienen capacidad vinculante, todo

³ Tarea imposible de cubrir en el espacio de este trabajo: en 1995, cuando todavía la economía experimental estaba despegando, la excelente revisión de Kagel y Roth (1995) ocupaba 720 páginas. Un trabajo equivalente hoy ocuparía varios volúmenes.

ello definido por las instrucciones que reciben los sujetos y que describen los procedimientos que rigen en el experimento; el último componente es el comportamiento de los agentes que participan en ellos, que siempre debe ser interpretable como reacción al entorno y a las instituciones. A esta clasificación de Smith hay que añadir un componente fundamental más para poder completar el esquema gráfico siguiente: los resultados experimentales.

Cuadro 1. Un experimento económico



Así, y siguiendo el Cuadro 1, un experimento tiene lugar en un entorno económico controlado, en el que una serie de agentes económicos interactúan mediante una institución estrictamente definida. Lo que caracteriza al entorno económico del laboratorio es el grado de control que se puede ejercer sobre las decisiones de los agentes. Este nivel de control sobre los flujos de información que relacionan entorno, instituciones y comportamiento normalmente se lleva a cabo mediante la utilización de redes de ordenadores.

De esta manera, una de las principales características de estos entornos controlados es la generación de ambientes virtuales en los que discurren de manera totalmente controlada los flujos de información entre los participantes. Por ejemplo, los agentes participantes en un experimento pueden ser compradores o vendedores, que

interactúan a través de una institución que puede ser un tipo particular de mercado. Sus decisiones (comportamiento) fluyen a través del entorno informatizado de manera que cada participante tiene información acerca de sus decisiones, de las decisiones de todos o de algunos de los otros participantes (por ejemplo sus precios ofertados de venta o sus propuestas o pujas de compra), de los resultados agregados de la interacción (por ejemplo, los precios vigentes en el mercado en el que participan, o de todos los mercados que están en funcionamiento en ese momento) o de cualquier otra variable relevante para el estudio (por ejemplo, las cantidades demandadas, las vendidas, los beneficios de los compradores, los de los vendedores, o incluso estimaciones económicas de la función de demanda o de oferta).

En este entorno controlado, los agentes se definen por sus características económicamente relevantes: preferencias, tecnología, dotación de recursos o información. Aunque los sujetos experimentales tengan sus propias características en el momento de entrar en el laboratorio, se pretende analizar el comportamiento en el marco de teorías que postulan que algunas variables controladas pueden afectar al resultado del experimento, lo que como veremos asegura la teoría del valor inducido.

Una institución económica especifica las acciones disponibles a los sujetos y los resultados que se derivan de cada combinación de acciones de los sujetos. Incluso para los experimentalistas, el primer instinto al diseñar un experimento es el de tratar de trasladar al laboratorio de la manera más fidedigna posible algún entorno real de interés económico sustancial. De manera análoga, y especialmente entre los teóricos, un instinto alternativo recomienda diseñar un experimento de tal manera que replique tan precisamente como sea posible los supuestos en los que se basa el modelo económico objeto de estudio.

¿Qué aproximación es la correcta? ¿imitar a la realidad o a los modelos formales? La respuesta correcta para Friedman y Sunder (1994) es que ninguno. El objetivo debería de ser encontrar un diseño que ofrezca la mejor oportunidad para aprender algo útil y para responder las preguntas que han motivado la investigación. Es fútil tratar de replicar en el laboratorio las complejidades de un entorno de campo, dado que la realidad tiene infinitos detalles y es su mejor modelo. No importa cuánto se detenga uno en trasladar los detalles de la realidad al laboratorio, un infinito número de detalles se quedarán sin capturar, con la dificultad práctica añadida de que el presupuesto nunca dejará avanzar en esa dirección lo suficiente y muy probablemente el entorno se habrá hecho tan complejo que será difícil separar las causas de los efectos. Como en cualquier otra disciplina, la simplicidad mejora el control. Tratar de encontrar los entornos más sencillos en el laboratorio que al mismo tiempo incorporen aspectos esenciales del entorno económico real es una buena norma de aproximación.

Es igualmente inútil intentar replicar en el laboratorio los supuestos precisos de un modelo formal. Una dificultad práctica añadida es que la mayoría de los modelos formales dejan fuera detalles vitales desde el punto de vista del diseño experimental, y la elección de qué detalles dejar fuera es por definición arbitraria en términos teóricos pero importante en términos de comportamiento. Incluso si se triunfa en la creación de una economía de laboratorio que reproduzca fielmente los supuestos del modelo formal, es muy posible que no se aprenda mucho con ello.

Si el comportamiento observado en la economía es consistente con las implicaciones del modelo formal, sólo se obtendrá una evidencia débil acerca del poder explicativo del modelo. La evidencia hubiera sido más fuerte si se hubiera observado el mismo comportamiento en una economía de laboratorio en la que se hubieran relajado los supuestos más restrictivos. En el caso extremo de una completa y perfecta recreación la consistencia entre los datos y nuestra predicción sólo nos diría que no hay ninguna inconsistencia lógica interna dentro del modelo. Una pírrica victoria, dado que los experimentos de laboratorio siempre tienen menos poder que las simulaciones o el análisis matemático.

Por otro lado, si el comportamiento observado que se encuentra no es consistente con el modelo formal, te enfrentas al criticismo sobre la inadecuación del diseño o sobre los procedimientos empleados para que los sujetos comprendieran las instrucciones (o más probablemente sobre los dos aspectos). A no ser que el propósito sea demostrar la estrechez del modelo o su nulo campo de aplicación, se puede aprender muy poco de ese ejercicio⁴.

⁴ Una analogía de Friedman y Sunder (1994) puede ser clarificadora al respecto. Si un artista desea expresar un acontecimiento humano, digamos la muerte de un hermano, dado que es poco partidario por motivos estéticos, morales o prácticos de replicarla, elige un medio de expresión, como el lienzo o la piedra. La cualidad de su pintura será juzgada por lo apropiadamente que simplifique la realidad para capturar y comunicar la esencia de su pérdida. La escultura será juzgada también por su impacto en el público, no por su fidelidad extrema a la realidad. De manera análoga, un experimento será juzgado, o debería de serlo al menos, por su impacto en nuestro entendimiento y no por su fidelidad a la realidad o al modelo.

2.2. *La evidencia empírica experimental*

Friedman y Sunder (1994) clasifica las fuentes de datos del trabajo empírico en cualquier tipo de ciencia distinguiendo entre datos experimentales, deliberadamente creados por el científico en un entorno controlado, y datos no experimentales, no generados por el científico en un entorno controlado, con un origen por tanto no totalmente dirigido por el científico. Una distinción complementaria a la anterior puede realizarse entre los datos de laboratorio, reunidos en un entorno artificial diseñado por el científico, y datos de campo, que son recogidos de un entorno natural, sean estos de carácter experimental o no.

El esquema resultante de cruzar ambos tipos de clasificaciones sería el recogido en el Cuadro 2.

Cuadro 2. Tipos de evidencia empírica

	Entornos no experimentales (N)	Entornos experimentales (E)
Datos de campo (C)	Tasa de inflación	Experimentos naturales
Datos de laboratorio (L)	Penicilina	Mercados experimentales

Obviamente, las cuatro combinaciones son posibles. Por ejemplo, un experimentalista interviene en un proceso que ocurre naturalmente y toma nota de los resultados para obtener el tipo de datos CE correspondiente a *Experimentos de campo*. Un ejemplo económico de este tipo de observaciones serían el *experimento natural* de mantenimiento de los ingresos en Denver o el *experimento natural* de evasión fiscal de Minnessota⁵. Tradicionalmente, sin embargo, la mayoría del trabajo empírico en economía ha sido del tipo CN: observaciones de campo no experimentales y por tanto no totalmente controladas por el economista, como la tasa de inflación de una eco-

⁵ Véase Pencavel (1986) para el primer caso y Coleman (1997) para el segundo.

nomía o cualquier otro dato macroeconómico agregado. El ejemplo arquetípico de LN es la penicilina, descubierta como subproducto generado por un accidente en un experimento⁶.

Por supuesto, la economía experimental proporciona datos tipo LE, fruto de experimentos realizados en un entorno controlado de un laboratorio, cuyo valor añadido específico es el de resultar especialmente valiosos para los propósitos científicos por ser relativamente sencillos de interpretar. Si el resultado A, digamos un nivel eficiente de asignación, se asocia siempre con la institución B, digamos un cierto tipo de mercado, estando el resto de variables institucionales y no institucionales bajo control, entonces podemos fácilmente concluir que B causa en alguna medida A. Los datos provenientes de fuentes no controladas no pueden soportar este tipo de asociaciones, ya que dada la ausencia de control sobre el proceso de generación de los mismos, la observación A puede estar causada tanto por la variable B, como por otra variable desconocida y oculta que las genera a las dos al mismo tiempo.

Los entornos de laboratorio, flexibles y controlables, son ciertamente más costosos de crear, mantener, y operar, dado que cada experimento requiere costes adicionales como los pagos a los sujetos, en el caso de la economía experimental. Así, tanto los costes fijos y/o hundidos como los costes variables y marginales pueden ser significativos en el caso de la economía experimental, y por supuesto más elevados que los derivados de la observación de campo, aunque este tipo de datos también tengan un coste.

Sin embargo, este coste debe compararse con el incremento de la validez o relevancia de los datos así obtenidos, aspecto crucial de toda fuente de datos. Cuando los entornos de campo tienen un interés en sí mismos, los datos de tipo CN y CE son automáticamente relevantes. Sin embargo, los datos CN son normalmente recogidos por agencias privadas o gubernamentales por razones no científicas, y mientras los astrónomos son directamente responsables de la credibilidad científica de sus observaciones la situación es bien distinta en el caso de los economistas que generan este tipo de datos. Así, en economía, cuando los datos no tienen la apariencia esperada, lo primero que se cuestiona es la calidad de los mismos (Smith, 1987).

⁶ Aunque este tipo de observaciones son realmente minoritarias dado que el uso del laboratorio ya implica el objetivo intencional de controlar el entorno en el que se generan las observaciones, en economía también existen han existido numerosas observaciones de este tipo, como por ejemplo las ligadas a los mecanismos de provisión de los bienes públicos, investigación dirigida en numerosas ocasiones por la emergencia de *anomalías* teóricas observadas en el laboratorio.

Los datos de laboratorio tienen problemas de validez distintos. Los problemas de validez interna están relacionados con el uso de las metodologías adecuadas y del rigor suficiente como para poder hacer una inferencia causal correcta. Dado que la validez interna es una cuestión directamente relacionada con los adecuados controles experimentales, el diseño experimental apropiado y el análisis riguroso de los datos, este requisito es esencialmente análogo al exigible a los datos experimentales de cualquier otra ciencia. En segundo lugar está la más interesante cuestión de la validez externa, que cuestiona la posibilidad de generalizar nuestras inferencias del laboratorio a la vida real y que como problema metodológico central analizamos en el punto siguiente.

2.3. *El requisito central de todo experimento: el valor inducido*

La idea central de este requisito es que el uso apropiado de una recompensa como medio permite al experimentalista inducir características pre-especificadas en los sujetos experimentales, y que las características innatas de los sujetos acaben siendo irrelevantes. Tres condiciones deben satisfacerse para inducir las características de los sujetos para Smith (1982): monotonicidad, prominencia⁷ y dominancia.

Por el requisito de *monotonicidad*, los sujetos deben preferir más recompensa a menos y no deben quedar saciados. Formalmente, si $V(m, z)$ representa las preferencias no observables de los sujetos sobre la recompensa monetaria y z representa todo

lo demás, la condición de monotonicidad exige que $\frac{\partial V}{\partial m}$ exista y que sea posi-

va para cualquier combinación posible de (m, z) . Esta condición parece fácil de satisfacer usando pagos monetarios en moneda local frente a la que los sujetos experimentales siempre preferirán más a menos y no quedarán saciados para un rango de recompensa y costes de oportunidad apropiados.

Por la condición de *prominencia* se deriva que la recompensa recibida por el sujeto depende de sus acciones y de las acciones de los demás agentes de la manera que definen las normas institucionales de manera comprensible. Esto es, la relación entre las acciones y la recompensa formaliza la institución experimental, y los sujetos entienden esa relación. Por ejemplo, un pago fijo de 5 Euros a los sujetos por parti-

⁷ Es realmente difícil traducir el original inglés *saliency*; hemos optado por prominencia aun a sabiendas de que no es un traducción perfecta.

cipar no es prominente por que el pago no depende de las acciones de los sujetos en el laboratorio una vez ya se ha presentado. Por otra parte, un pago de un centavo de Euro por cada punto de beneficio ganado en el mercado lo es debido a que el pago depende de las acciones de los sujetos.

Por el requisito de *dominancia*, cambios en la utilidad de los sujetos debido al experimento proceden predominantemente del medio de recompensa y el resto de influencias son despreciables. Esta condición es la más problemática de las tres, dado que tanto las preferencias V como la variable z , pueden no ser observables por el experimentador. La dominancia es más plausible si las recompensas m se incrementan y si los más obvios componentes de z permanecen constantes. Por ejemplo, los sujetos a menudo se preocupan de las recompensas ganadas por los otros sujetos. Si los procedimientos experimentales hacen imposible conocer o estimar las recompensas del resto de los sujetos (mediante el recurso a la privacidad estricta en el proceso de los pagos) entonces un componente importante de z está neutralizado. Los efectos demanda, provenientes de los deseos de los sujetos de ayudar (u obstaculizar) al experimentador, deben controlarse evitando la revelación de los objetivos del experimento.

Si las tres condiciones son satisfechas, el experimentalista consigue controlar las características de los sujetos. Para ilustrar este hecho, supongamos que se quieren inducir unas preferencias determinadas (por ejemplo, una función Cobb Douglas) representadas por la función de utilidad que depende de dos variables $U(x, y)$ y que produce unas recompensas monetarias al sujeto en la forma especificada por la institución vigente en el experimento. Las preferencias inducidas $W(x, y) = V(m_0 + U(x, y), z_0 + \Delta z)$ donde (m_0, z_0) son los recursos iniciales inobservables del sujeto (tanto en términos de dinero como en términos de cualquier otra cosa), y Δz representa la recompensa no monetaria que el sujeto obtiene participando en el experimento. Por el lemma de Hicks podemos concluir que las dos funciones de utilidad (la utilidad absoluta del sujeto y la que extrae de la recompensa monetaria) representan las mismas preferencias si sus tasas marginales de sustitución coinciden⁸.

La intuición es que el experimentalista puede elegir libremente cualquier relación entre las variables y la recompensa monetaria siempre que sea capaz de explicar cla-

⁸ Algo más formalmente: $MRS_w = \frac{W_x}{W_y} = \frac{V_m U_x + V_z \Delta Z_x}{V_m U_y + V_z \Delta Z_y} + \frac{V_m U_x}{V_m U_y} = \frac{U_x}{U_y} = MRS_u$. Donde por la

definición de una relación marginal de sustitución se cumple la primera igualdad y la última, por la propiedad de la preponderancia, la segunda, por la de dominancia la tercera y por la monotonicidad, la cuarta.

ramente la relación a los sujetos (cumpliendo la condición de prominencia) y que los sujetos estén motivados por el medio de pago (por el requisito de monotonicidad) y no por otras influencias (cumpliendo con la cláusula de dominancia), ya que entonces el experimentador puede controlar las características de los sujetos para analizar la relación elegida en el laboratorio.

El concepto de prominencia sirve para diferenciar las encuestas de los experimentos económicos controlados. Los datos procedente de encuestas (respuestas humanas a preguntas hipotéticas) son empíricos, pero dado que esas respuestas no están motivadas económicamente por ningún incentivo económico su fiabilidad es cuestionable. Una encuesta típica pregunta a los sujetos para que informen acerca de algunos aspectos de sus características personales, acciones históricas, o acontecimientos determinados y a veces para pedirles que elijan en situaciones hipotéticas. La experimentación económica controlada no debe confundirse con esta última clase de encuestas; dado que no se ofrecen recompensas sobresalientes en estas encuestas los sujetos no están efectuando elecciones económicas bajo condiciones controladas por el experimentalista..

Dicho de otra manera: un procedimiento de laboratorio que incluya pagos idénticos a los sujetos para responder a preguntas hipotéticas es una encuesta y no un experimentos económico⁹.

Para terminar esta sección, y siguiendo a Hey (1996), podemos identificar 7 requisitos prácticos que todo experimento debe cumplir:

1. Los sujetos participantes deben ser pagados en efectivo para poder controlar el entorno económico. La mayor parte de los pagos deben estar ligados a las acciones que tomen los sujetos en el experimento en el que participen¹⁰ y el pago medio debería de superar el coste de oportunidad medio. Una buena aproximación al coste de oportunidad medio es el salario-hora del tipo de sujetos que participan en el experimento.

⁹ De manera similar, a menudo se confunden los datos procedentes de las simulaciones con datos económicos experimentales. Las simulaciones en el ordenador de algún modelo teórico son una extensión del modelo teórico debido a que ninguna decisión humana participa en el proceso excepto en el momento de escribir el programa. Si tradicionalmente los investigadores usaban la lógica deductiva y las matemáticas para descubrir las implicaciones de un modelo teórico, de manera creciente se recurre a la simulación bien debido a lo intratable del modelo bien debido a que la capacidad computacional de los ordenadores es cada vez mayor y el coste menor.

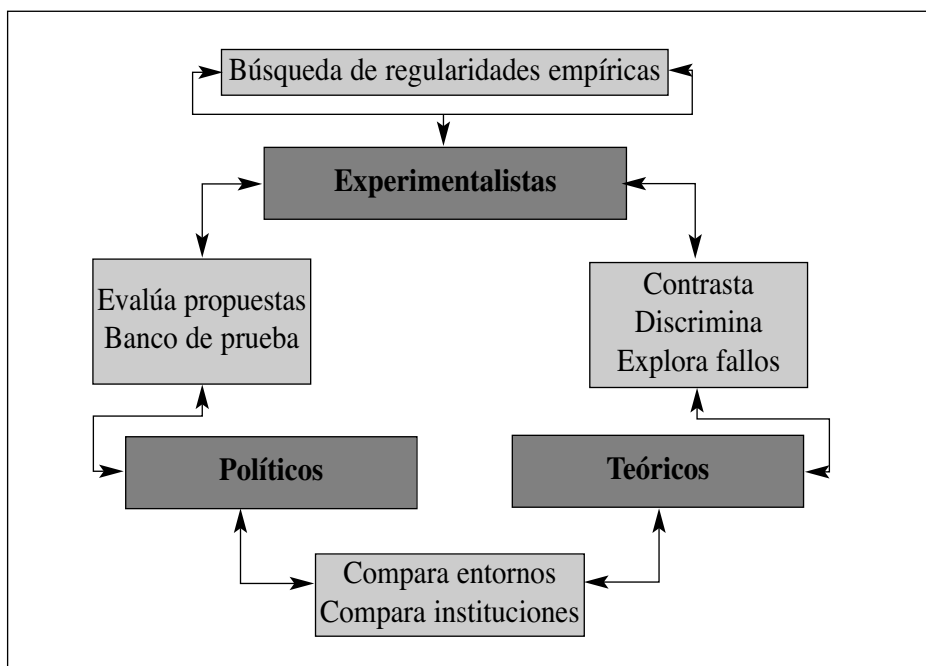
¹⁰ Es práctica habitual el efectuar un pago de segundo orden simplemente por presentarse, especialmente si se lleva a cabo un procedimiento de selección de la muestra por cuestionario.

2. Los sujetos participantes que mejor permitan cumplir las condiciones de dominancia y prominencia serán aquellos sujetos con bajos costes de oportunidad y curvas de aprendizaje con mucha pendiente. Los estudiantes universitarios cumplen perfectamente estas dos condiciones.
3. El entorno económico debe ser el más sencillo posible condicionado a la puesta en marcha de la institución. La sencillez promueve la prominencia y reduce la ambigüedad en interpretar los resultados. Las instrucciones deben estar escritas con la mayor claridad y detalle y es conveniente verificar el nivel de entendimiento con cuestionarios previos.
4. Para promover la dominancia, han de evitarse el empleo de palabras cargadas de sentido para los sujetos. En un dilema del prisionero, conviene etiquetar a las estrategias A y B, por ejemplo, en vez de Cooperar y Delatar. De la misma manera los roles de los sujetos deben ser denominados de la manera más neutra posible (compradores y vendedores, A y B, nunca rivales u oponentes).
5. Si la dominancia acaba siendo cuestionable y el presupuesto lo permite, se deben conducir algunas sesiones con un incremento proporcional en las recompensas. Un cambio sistemático en el comportamiento observado implicaría que la dominancia no ha sido conseguida para los niveles inferiores de pagos.
6. La privacidad de los sujetos y de los objetivos reales de la investigación debe mantenerse a resguardo de la curiosidad de los participantes siempre que no afecte críticamente al objeto de la investigación. Las preferencias previas de los sujetos pueden tener componentes malévolos o benévolos con respecto al objeto de estudio o al puesto relativo que ocupen en términos de ganancias, por lo que la dominancia puede verse comprometida si no se mantiene.
7. No debe engañarse a los sujetos o mentirles. Aunque los psicólogos sociales han efectuado interesantes experimentos acerca de este tipo de procedimientos, los experimentos económicos requieren de una credibilidad completa debido a que la prominencia y la dominancia se pierden si los sujetos dudan acerca de la anunciada relación entre acciones y recompensas o si los sujetos están apercibidos de la posible existencia de trampas o trucos. Los costes de este tipo de engaños, además, suelen proyectarse en el tiempo, por lo que futuros experimentalistas pueden encontrarse con problemas como consecuencia de ello.

3. LOS OBJETIVOS DE LA ECONOMÍA EXPERIMENTAL

Los experimentos económicos pueden tener muchos propósitos. El método más apropiado para diseñar y conducir un experimento depende de esos propósitos. Una buena revisión de los posibles propósitos de los experimentos está en Plott (1982), Plott (1987) o Smith (1994). A partir de una más tarde famosa clasificación de Roth (1995), podemos establecer tres tipos de propósitos generales en la economía experimental cuya interacción queda recogida en el Cuadro 3:

Cuadro 1. Un experimento económico



2.1. *Búsqueda de regularidades empíricas (diálogo entre experimentalistas)*

El primero de los cometidos o propósitos tiene que ver con la búsqueda mediante experimentos de variables acerca de las cuales la teoría tiene poco que decir. A menudo estos experimentos están motivados por experimentos previos y están diseñados para aislar las causas de alguna regularidad observada, variando detalles del diseño experimental previo que las descubrió. Estos experimentos formarían parte de un diálogo entre los propios experimentalistas, ya que a medida que la evidencia confirma acumulativamente alguna relación, la búsqueda de significado se hace posible, y nuevas teorías del comportamiento observado pueden ser propuestas y comprobadas, favoreciendo a su vez el diálogo entre teóricos y experimentalistas.

Uno de los primeros objetivos de la experimentación es conseguir regularidades empíricas. McCabe, Rassenti y Smith (1994) comparan diferentes instituciones de mercado alternativas cuyas propiedades teóricas no están completamente entendidas. En otros casos, diversas teorías compiten entre sí ofreciendo predicciones alternativas, por lo que los experimentos pueden dibujar el mapa de aplicabilidad de las distintas teorías. Fiorina y Plott (1991) estudian las decisiones políticas para descubrir que sólo algunos de los 16 modelos planteados son consistentes con los datos. Finalmente, hay áreas en las que un solo modelo es aplicable y entonces el trabajo de laboratorio consiste en demostrar si bajo determinadas condiciones la teoría produce el tipo de datos que predice.

Una segunda opción sería la de establecer regularidades empíricas como base de una nueva teoría. La medición de determinadas características individuales humanas, como el deseo de pagar por bienes ambientales o la aversión al riesgo, ver Cummings, Harrison y Rutstrom (1992), pueden estar en la base de nuevos modelos. Teorías bien formuladas en la mayoría de las ciencias tienden a ser precedidas por muchas observaciones, que a su vez estimulan la curiosidad de saber qué explica esas regularidades documentadas.

La teoría microeconómica tiende a construirse sobre la base de supuestos simplificados, y a desechar muchas de las instituciones de intercambio tal y como las conocemos. Pero en el laboratorio, especialmente si el laboratorio está informatizado, las instituciones complejas son tan sencillas de estudiar como las subastas de un único bien. Esto hace posible ir más allá de los actuales confines de las teorías actuales para establecer regularidades que pueden permitir a los teóricos anticipar cuáles son las dificultades sobre las que merece la pena trabajar. Un ejemplo mencionado por Smith

(1994) sería la subasta doble continua, usada en todo el mundo. En ella, los compradores anuncian precios de compra (pujas), mientras los vendedores anuncian precios de venta (ofertas). Cualquier nueva puja debe realizarse a un precio inferior (superior) que la puja (oferta) actual; esto es, el rango de pujas efectuadas debe estrecharse. Un contrato vinculante ocurre cuando un comprador acepta una petición de un vendedor o cuando un vendedor acepta una puja de un comprador. Esta institución, debido en parte a la robustez de sus propiedades ha sido estudiada intensivamente en el laboratorio (por ejemplo Smith, 1991) antes de que se formalizaran teóricamente sus propiedades (véase Freidman y Rust, 1992). Un estudio reciente de este tipo de instituciones en el campo de la economía pública es Broseta, Fatás y Neugebauer (2003).

Un tercer cometido de los experimentos puede ser el de comparar entornos usando la misma institución, lo que permite a los investigadores analizar la robustez de una institución y de paso constituirse en puente de diálogo entre académicos teóricos y políticos. El objetivo último es someter a tensión a los modelos teóricos bajo entornos extremos en los que las propiedades establecidas de la teoría pueden desmoronarse. Por ejemplo, en subastas de valor común (donde el producto que se subasta tiene el mismo valor para todos los compradores una vez la subasta se completa) el modelo de Nash se comporta mejor para un número de compradores bajo (ver Kagel y Levin, 1986). De manera similar, la predicción basada en el equilibrio de Nash se comporta bastante bien en un entorno de negociación al estilo de Fouraker y Siegel (1963) pero se hunde en el entorno del juego del ultimátum.

El último de los cometidos dentro de este primer bloque sería el de comparar instituciones, de nuevo participando en el diálogo entre gestores políticos y teóricos. De esta manera, usando instituciones idénticas, pero modificando las normas de intercambio se pueden establecer las propiedades comparativas de las instituciones. Ejemplos de este propósito pueden ser la comparación de distintos tipos de subastas (subastas inglesas, holandesas, selladas de primer o segundo precio, subastas múltiples de precio uniforme y de precio discriminatorio) en las que la institución es la misma, pero cambia el entorno.

2.2. *Dialogando con la teoría*

La segunda de las categorías incluye los experimentos diseñados para comprobar las predicciones de teorías formalmente bien articuladas, y para observar regularidades no predichas por la teoría, en un entorno controlado que permiten que estas observaciones sean interpretadas de manera no ambigua en relación con la teoría. Estos

experimentos en ese sentido forman parte del diálogo entre teóricos y experimentalistas.

El primer motivo dentro de esta categoría sería el de contrastar teorías o discriminar entre ellas. Esta motivación proviene claramente de la literatura de teoría de juegos y de la teoría económica. Se contrasta una teoría comparando sus predicciones o sus implicaciones sobre los resultados con las observaciones experimentales. Cuanto mayor sea la frecuencia con que las observaciones se corresponden con las predicciones, en el contexto de un diseño en el que los aciertos difícilmente se pueden producir por casualidad, mejor para la teoría.

Un segundo motivo sería el de explorar las causas del fallo de una teoría. Dado que las teorías siempre están sujetas a mejoras, esta segunda motivación puede colaborar en el proceso de refinamiento de una teoría. Si las observaciones de un experimento no son conformes a las implicaciones de la teoría, lo primero que hay que hacer siguiendo a Smith (1994) es revisar el diseño, para asegurarse de que el fallo en la predicción es responsabilidad de la teoría. Si las teorías bien articuladas modelizan formalmente el entorno y las normas de intercambio, los experimentalistas deben reproducir estas características en el laboratorio. En el proceso de contraste, si el diseño experimental continúa pareciendo apropiado y la teoría todavía continúa fallando, un examen experimental más detenido puede servir para descubrir la causa. Llegar a conocer las razones últimas de un fallo teórico es esencial para cualquier programa de investigación preocupado en modificar la teoría. Buenos ejemplos de ello serían la literatura sobre negociación (Roth, 1987 o Hoffman y Spitzer, 1985, o Bolton, 1991). A menudo teorías que se comportan no demasiado bien inicialmente mejoran si los sujetos tienen más experiencia, o si los sujetos reciben mayores pagos, pero a veces estas medidas fallan en arrojar mejores resultados que mejoren los niveles de acierto teóricos.

En tercer lugar, los experimentos pueden constituirse en test formal e irrefutable de la teoría. Este último propósito está abierto a un intenso debate. Algunos experimentalistas han dudado en los últimos años de si el propósito de un experimento es el de comprobar la validez de una teoría, ya que desde un punto de vista formal una teoría consiste en una serie de axiomas, supuestos de comportamiento y definiciones que dan origen siguiendo las leyes de la lógica deductiva a unas conclusiones. Una teoría es formalmente válida si es internamente consistente, esto es, si no lleva a afirmaciones que se contradigan unas a otras, y si las conclusiones se pueden probar a partir de los supuestos. ¿Qué podemos aprender entonces de los experimentos? Algunos experimentalistas, y muchos psicólogos, piensan en los experimentos como un medio de contrastar la validez descriptiva de los supuestos acerca del comporta-

miento humano que sustentan a los modelos teóricos. Otros, incluyendo a muchos economistas, piensan que los supuestos de comportamiento no necesitan cumplir ningún requisito o criterio de validez descriptiva al estilo psicológico. En cambio creen que una teoría tiene interés práctico en tanto en cuanto sus conclusiones proporcionen buenas aproximaciones (comparativamente con teorías alternativas) del comportamiento real incluso si los supuestos no se satisfacen.

2.3. *Diálogo con los políticos*

Este tercer tipo de propósitos representa el resultado del diálogo con los gestores de la política económica. Su característica básica es que el entorno experimental está diseñado para parecerse lo más posible en ciertos aspectos al entorno natural en el que ocurren los fenómenos de interés para los gestores de las políticas. Esta categoría de experimentos no ha dado lugar a una serie muy grande de experimentos repetidos por un gran número de investigadores, pero sin embargo ofrece la posibilidad de utilizar métodos científicos para llevar a cabo una de las más tradicionales actividades de los economistas: el formular recomendaciones en cuestiones de política económica cuyas respuestas suelen estar más allá del conocimiento científico de la profesión.

Una primera aproximación dentro de este tercer grupo sería el de evaluar las propuestas de políticas. La propuesta original de Friedman de que el Tesoro subastara sus títulos en subastas de un único precio ha llevado a que se efectúen comparaciones con normas discriminatorias (Smith 1991), recientemente analizadas para el caso español por Brandts y Abbink (2001). En la pasada década, la industria privada y los gobiernos han financiado estudios de los incentivos para mercados continuos de subasta doble, subastas de premios de emisión, mecanismos para adjudicar espacios a las aeronaves en los aeropuertos, o mecanismos de asignación de recursos para las naves espaciales, o mercados inteligentes para recursos compartidos en grandes empresas (como Hewlett Packard), o para desregular los mercados de energía en USA.

De esta forma, algunos experimentos se conducen para generar datos que puede influir en una decisión específica. Por ejemplo, Grether y Plott (1984) realizaron un experimento para asesorar en un caso específico de defensa de la competencia; Hong y Plott (1982) para asesorar a la comisión interestatal de comercio; Rassenti, Reynolds y Smith (1988) para asesorar a la Comisión Federal de Regulación de Energía; van Winden (2001) para analizar con financiación del gobierno holandés el denominado Plan van Elswijk de reforma del sistema de aportaciones a la seguridad social para cubrir los subsidios de desempleo, desde un sistema basado en los salarios a uno

basado en el valor añadido; Fatás y Pallardó (2004) un análisis de los distintos procedimientos de tomas de decisiones en los comités de política monetaria del Banco Central Europeo.

Adviértase que influenciar a las autoridades no es el único propósito persuasivo de los experimentos. Los consumidores o los votantes también han sido el objetivo de la experimentación económica como arma de persuasión, ver por ejemplo Cohen (1992) para un recorrido de los experimentos que se han empleado para influenciar a consumidores.

Un último propósito declarado de los experimentos en economía sería el servir como banco de pruebas antes de pasar a los experimentos a escala mayor. McCabe (1991) analiza mercados inteligentes para compartir recursos de informática, Friedman (1994) analiza un experimento piloto sobre redes eléctricas compartidas, Smith (1994) menciona el caso del mercado de valores de Arizona, derivado directamente de los experimentos desarrollados en el laboratorio de Tucson¹¹.

4. EL LIBRO DE RESPUESTAS DEL EXPERIMENTALISTA

Muchos economistas cuestionan la validez externa de los datos de laboratorio y sienten que ese tipo de datos no son representativos del comportamiento en el mundo real. Quizás sea interesante recordar que históricamente, los experimentalistas en otras disciplinas han encontrado ese mismo tipo de críticas. Los críticos de Galileo no creían que el movimiento de péndulos o pelotas en planos inclinados tuvieran ninguna relación con el movimiento planetario en la esfera celeste. En términos más generales, el problema de la validez de los datos obtenidos en el laboratorio se encuentra en cualquier observación recogida en el laboratorio y en buena medida este problema tiene que ver con el papel que se adjudica a las observaciones empíricas en el progreso de la ciencia.

Por mucho que observaciones efectuadas en el laboratorio (con todo tipo de controles para garantizar su validez interna) demuestren los efectos perniciosos de

¹¹ Un último propósito no analizado aquí por motivos de espacio y por no estar directamente relacionado con la investigación sería el pedagógico. En este sentido, Wells (1991) y Williams y Walker (1993) representan un buen ejemplo de cómo este empeño ha sido llevado a sus últimas consecuencias.

determinadas sustancias tóxicas en ratones, nadie cuestiona a la metodología experimental porque esas sustancias afectan a los seres humanos de una manera diferencial, siendo sus organismos más complejos y expuestos a dosis diferentes un período de tiempo distinto.

Obsérvese que la lógica deductiva no proporciona la base para rechazar este escepticismo. Por el simple hecho de que hayas observado salir el sol cada mañana durante 20 años no puedes deducir con seguridad la proposición de que saldrá mañana. El principio empleado es el principio general de inducción, mediante el que a partir de regularidades de comportamiento observadas sistemáticamente se deduce que éstas persistirán en nuevas situaciones siempre que las condiciones relevantes subyacentes permanezcan sustancialmente idénticas. Aunque una teoría general nos sugiera lo que es relevante y lo que es un cambio sustancial, el principio de la inducción es en sí mismo un supuesto (o un axioma) y no una proposición deducible.

Smith (1982) identifica el principio de la inducción en el contexto de la Economía Experimental con el precepto del paralelismo cuando afirma que *proposiciones acerca del comportamiento de los sujetos y la actuación de las instituciones que han sido comprobados en el laboratorio microeconómico se aplican también a entornos microeconómicos fuera del laboratorio siempre que similares condiciones ceteris paribus se sostengan*¹².

El cumplimiento de este precepto de paralelismo es una condición suficiente para deducir que los resultados encontrados en el laboratorio pueden tener validez en el mundo real. En cierto sentido, un observador escéptico de la metodología experimental debería de cargar con el peso de la prueba a la hora de identificar aquellas condiciones que cambian en el mundo real que son las causantes de la modificación de los resultados observados en el laboratorio y por tanto cuestionan la validez externa del método. En base a esta identificación, nuevos experimentos podrían ser diseñados y conducidos para comprobar la validez de la afirmación del escéptico.

Un ejemplo puede ilustrar este punto. Los mercados de valores experimentales usualmente son cuestionados bajo la crítica de artificialidad, pero cuando se pregunta a los escépticos acerca de las características que los convierten en artificiales suelen afirmar que las variables institucionales que provocan esa condición son las relativas al escaso número de compradores presentes en el mercado, en la cuantía de las recompensas monetarias empleadas para incentivar a los sujetos y en que los sujetos

¹² Pág. 936.

experimentales son distintos que los profesionales que suelen participar en esos mercados. En base a nuestro análisis previo de cada uno de estos requerimientos concretos, adviértase que en nuevos experimentos se puede pagar más, ampliar el número de sujetos por mercado y reclutar a profesionales para participar en esos experimentos y comprobar si de facto esas variables son las causantes de la ausencia de validez externa.

Tal y como dice Plott (1982), *el arte de plantear cuestiones descansa en la capacidad de convertir el estudio de sencillos casos especiales relevante para la comprensión de la complejidad. Los modelos y teorías generales por definición son aplicables a todos los casos especiales. Por lo tanto, deberían de funcionar en los casos especiales del laboratorio. Si los modelos fracasan en capturar los que se observa en estos casos especiales, pueden ser modificados o rechazados a la luz de la experiencia*¹³.

En un no demasiado conocido trabajo, Stigler (1977) proponía con humor que las conferencias en economía funcionarían de manera mucho más eficiente si los participantes en ellas pudieran utilizar una lista estándar de comentarios que pudieran ser referenciados numéricamente, produciendo un considerable ahorro de tiempo. En esa misma línea, Thaler (1991) propone un manual de Conferencias en Economía y Psicología que puede ser utilizado como punto de partida para elaborar una lista de preguntas.

Estas preguntas ponen de manifiesto en cierta medida las limitaciones de la metodología experimental aplicada al campo de las ciencias sociales en general y de la economía en particular, y permiten dar un recorrido por sus principales dificultades. Por motivos de espacio, nos vamos a centrar en tres críticas que consideramos principales. La primera afirmación más frecuentemente escuchada es la siguiente:

4.1. *Crítica 1: Selección de muestra*

Crítica 1: En el mundo real, los agentes no son estudiantes jóvenes. Los agentes económicos verdaderos se comportan de manera distinta, por lo que los resultados observados en el laboratorio no son extrapolables.

Los críticos de la economía experimental utilizan recurrentemente como objeción a la validez externa del método experimental el que se recurra a estudiantes en lugar

¹³ Pág. 1520.

de los profesionales que toman decisiones en el mundo económico real. El argumento básico tiene que ver en las diferencias que se pueden esperar en el comportamiento de los sujetos como consecuencia de la falta de experiencia de los sujetos experimentales en ese tipo de ámbitos. Dado que los profesionales (agentes económicos reales) acaban familiarizándose de manera profunda con las instituciones económicas con las que trabajan, internalizan las consecuencias de sus decisiones económicas bajo las condiciones económicas reales. De esta manera, sus decisiones reflejan el comportamiento económico real de una manera que los sujetos experimentales no pueden reproducir.

Independientemente del interés específico que puede tener la investigación de cómo se comportan futuros agentes económicos reales, este comentario plantea una cuestión eminentemente empírica, tal y como planteábamos al principio de esta sección. Si los *agentes económicos reales* hubieran tenido un comportamiento distinto de la muestra de sujetos presentes en el estudio experimental es una cuestión a la que se puede responder empíricamente.

De esta manera, esta crítica ha sido contestada de manera sistemática mediante la utilización de este tipo de expertos en lugar de sujetos sin experiencia. Schurr (1987) no encuentra diferencias en un juego de negociación en condiciones de riesgo en el que emplea además de estudiantes de MBA a un grupo de control de intermediarios financieros profesionales. Anderson y Sunder (1989) es una de las excepciones al encontrar diferencias significativas entre el comportamiento de sujetos sin experiencia y profesionales de instituciones financieras en un sofisticado mercado experimental de valores en el que se intercambian activos mediante una subasta doble.

Dyer, Kagel y Levin (1989) utilizaron a profesionales con un mínimo de 20 años de experiencia en el sector de la construcción en una subasta experimental diseñada de manera muy similar a las que los profesionales estaban acostumbrados a enfrentarse en sus trabajos para encontrar que, a pesar de emplear sistemáticamente patrones de pujas distintos a los empleados por los estudiantes, no eran capaces de evitar las pérdidas asociadas a la sobrepuja del ganador, objeto del experimento. DeJong, Forsythe y Uecker (1988) emplearon a profesionales del Consejo Profesional de Contables norteamericano y del Departamento de Contabilidad de la Universidad de Iowa en un experimento de auditoría junto a estudiantes sin experiencia para encontrar que no había diferencias significativas de comportamiento en términos de precios, beneficios o eficiencia del mercado. Cech (1988) emplea a un grupo de ejecutivos de la industria de las telecomunicaciones y asesores de organismos públicos de defensa de la competencia junto a estudiantes en un experimento sobre regulación de mercados para concluir que no había diferencias significativas entre los dos tipos de muestras.

En entornos de decisión política Potters y Van Winden (1997) analizan las decisiones de funcionarios públicos en un juego de señalización para el estudio del funcionamiento de los lobbys para no encontrar diferencias con experimentos previos en términos agregados, aunque sí en términos individuales. En un entorno similar, Fatás, Neugebauer y Tamborero (2004) analizan el efecto punto de referencia y el efecto certidumbre usando una muestra de políticos con un doctorado en economía y experiencia directa en la toma de decisiones políticas para encontrar los mismos patrones de conducta existentes para los estudiantes sin experiencia.

Ball y Cech (1996) realizan un recorrido sistemático sobre los experimentos que han reclutado sus sujetos entre los *agentes económicos reales*, en lugar de recurrir a estudiantes., para concluir que *excepto en un caso*¹⁴, *no encontramos evidencia de este efecto entre estudiantes y profesionales de mercado participando en un mismo entorno experimental en el que se han respetado cuidadosamente los preceptos de la economía experimental y se han contrastado asuntos económicos relevantes* (pág. 257)¹⁵.

Plott (1982) probablemente acierta al resaltar que la validez externa está relativamente a salvo del problema de la experiencia ya que *aunque los procesos de laboratorio son simples en comparación con los naturales, son reales en el sentido de que en ellos participan sujetos reales tratando de obtener beneficios sustanciales siguiendo normas reales. Precisamente porque son reales es por lo que son interesantes*¹⁶.

4.2. *El tamaño de la recompensa*

Crítica 2: Si las apuestas fueran lo suficientemente grandes no observaríamos los mismos resultados. Por mucho que los incentivos ofrecidos a los estudiantes sean superiores a su coste de oportunidad medio, si se recompensaran las decisiones a una escala mayor, los resultados cambiarían.

¹⁴ El mencionado de Anderson y Sunder (1989)

¹⁵ Es curioso que una gran parte de los psicólogos que desarrollan análisis experimentales de conducta (ver Lowestein (2000) como ejemplo reciente) encuentran esta crítica metodológica irrelevante, bajo el argumento de que si el entorno experimental no permite un proceso de aprendizaje lo suficientemente intenso, el mundo real muchas veces tampoco, dado que la experiencia no garantiza que las decisiones sean cualitativamente mejores para sus propios intereses.

¹⁶ Pág. 1486.

Frente a esta segunda crítica, se puede empezar recurriendo a la respuesta empleada en la primera. Dado que en la mayoría de los experimentos citados el coste de oportunidad de los profesionales fue empleado como método para determinar el tamaño de los incentivos, la ausencia de diferencias significativas entre los dos grupos de sujetos es una primera señal de que empíricamente la evidencia a favor de este razonamiento también es débil¹⁷.

Diversos experimentos han analizado las decisiones de sujetos mediante el recurso a incentivos muy superiores como manera de controlar este tipo de variable. Slovic y Roth (1998) multiplicaron en algunos grupos por 25 el tamaño del incentivo en Eslovaquia para no encontrar diferencias significativas en juegos sencillos de negociación. Cadsby y Maynes (1999) llegan a similares conclusiones para el caso de una provisión voluntaria de un bien público. Cameron (1999) estudia en Indonesia los resultados de un juego del ultimátum en el que el tamaño del pastel a repartir es el equivalente al salario de 3 meses para encontrar sólo cambios de segundo orden en el comportamiento de los sujetos que recibían la oferta y ningún cambio significativo en el de los que las hacían.

La utilización de grupos de control con una estructura de incentivos muy superior al resto de grupos analizados es empleado de manera sistemática en algunos laboratorios norteamericanos. En concreto, en el *Economic Science Laboratory* de la Universidad de Arizona se celebran cuatro talleres anuales en los que estudiantes de licenciatura y estudiantes de postgrado participan en experimentos en los que las ganancias individuales máximas de cada día superan con mucho el salario mensual de los estudiantes. A pesar de ello, la evidencia es abrumadoramente favorable a la ausencia de diferencias sistemáticas entre los grupos de control con pagos elevados y los grupos con una estructura de incentivos normalizada.

Adviértase que este argumento está relacionado con la prevención de los posibles efectos demanda antes mencionados. Aun siendo los incentivos los apropiados, existe el riesgo de que los sujetos quieran además parecer listos, bondadosos, inteligentes o, simplemente, ganadores a los ojos del experimentador. Tal y como se mencionó anteriormente, el anonimato y la ausencia de información sobre el objetivo real del experimento mitigan al menos este tipo de problemas.

¹⁷ Por ejemplo, Cech (1988) empleó como incentivo experimental el salario que efectivamente cobraban en sus empresas los ejecutivos por hora (unos €200/hora).

4.3. *El tamaño de la muestra*

Crítica 3: La muestra analizada tiende a tener un tamaño insuficiente en los experimentos. Los resultados experimentales son sensibles al número de observaciones sobre los que se establecen y por lo tanto no son extrapolables.

Para responder a esta crítica la economía experimental ha utilizado fundamentalmente la replicación estacionaria. Gracias a la proliferación de laboratorios de economía experimental, y al requisito impuesto a todos los trabajos experimentales publicados de especificar suficientemente los detalles concretos del trabajo para permitir siempre su replicación controlada, la repetición masiva de los principales tipos de experimentos es un hecho. Ya desde el trabajo de Popper (1959) es evidente que la replicación es correctamente percibida como una condición necesaria de la validez de los resultados experimentales, aunque las repeticiones experimentales nunca sean exactas. En cierto sentido, sin embargo, estas diferencias en el diseño experimental son importantes para incrementar nuestra confianza en los resultados, siempre que se mantengan dentro de unos niveles aceptables para la comunidad científica relevante. Una vez que unos resultados experimentales determinados están bien asentados, la replicación no aporta nada.

En el caso concreto de la economía experimental, la replicación de resultados incorpora el problema práctico, tal y como recoge Siakantaris (2000), de que existe una serie de factores humanos específicamente numerosos que deben ser controlados, debido a que el laboratorio no es un entorno socialmente neutro. A pesar de ello, la replicación estacionaria de los experimentos ha adquirido tal dimensión que determinados resultados básicos están más allá de cualquier duda razonable relativa a la replicabilidad de los mismos.

5. CONCLUSIONES

A lo largo de este trabajo se ha pretendido recorrer algunos de los elementos esenciales de la metodología experimental, no desde el punto de vista del filósofo de la ciencia sino del investigador en economía o en ciencias sociales. Probablemente este camino se hubiera podido recorrer de manera alternativa acudiendo a un estudio más sistemático de los logros o resultados concretos de esta metodología en alguno de los campos de investigación en los que ha sido utilizada.

Excede con mucho el ámbito de este trabajo el pretender resumir tales resultados, especialmente cuando existen trabajos excelentes que cubren dichos avances y que

pueden servir de punto de arranque de aquellos investigadores que deseen profundizar. Tanto en organización industrial (Plott, 1982), la economía pública (Ledyard, 1995, o Chen, 2000), la negociación (Guth, 1995), los experimentos macroeconómicos (Duffy, 1998), el cumplimiento fiscal (Andreoni, 1999, o para un reciente estudio en castellano Fatás y Roig, 2004), o las preferencias sociales (por ejemplo, véase Fehr y Gächter, 1998) existen numerosos trabajos recientes y completos.

Aunque el debate metodológico sobre la economía experimental está vivo, y puede seguirse en recientes trabajos como Camerer (1999) o Starmer (1999), en las secciones precedentes se han establecido las características esenciales de todo experimento económico, basadas en los principios de monotonicidad, preponderancia y dominancia, así como sus principales objetivos: construir sólidas regularidades empíricas, contrastar y mejorar los modelos teóricos existentes y servir de banco de pruebas a medidas de política económica.

¿Qué hemos aprendido acerca del comportamiento individual en los experimentos económicos? Podemos utilizar las conclusiones de Ledyard (1995) sobre experimentos de economía pública para remarcar los siguientes puntos:

(i) La teoría convencional no puede explicar los datos procedentes de los experimentos. Muchas anomalías en este sentido han sido detectadas y sistematizadas por los experimentos económicos. Estas anomalías tienen que ver tanto con las limitaciones cognitivas de los agentes (su racionalidad limitada) como con la existencia de preferencias sociales no incluidas en los modelos convencionales (relacionadas con la equidad, la aversión a la desigualdad o la reciprocidad).

(ii) Las decisiones observadas en los experimentos responden sistemáticamente a las instituciones y los incentivos monetarios contenidos en los mismos, reflejando que la metodología experimental está dotada de la suficiente validez interna como para analizar comportamientos relevantes de sujetos lo suficientemente motivados. En este sentido, el análisis comparativo de instituciones en entornos controlados como el laboratorio parece ser una metodología no sólo de futuro sino de presente para buena parte de los programas de investigación en economía.

(iii) Tomados uno a uno, los modelos teóricos alternativos a los tradicionales basados bien en agentes con racionalidad limitada (a la Camerer, 1999) o en la existencia de algún tipo de preferencias sociales (a la Charness y Rabin, 2002) distan todavía de ofrecer una potencia de fuego equiparable a los modelos tradicionales basados en una capacidad cognitiva sin límites y en una función de utilidad bien definida con el pago monetario individual como único componente. Sin embargo, la distancia existente en términos de universalidad y simplicidad entre los nuevos modelos

surgidos al amparo de resultados experimentales y los modelos tradicionales se ha reducido tanto en los últimos años, que no es descabellado afirmar que la nueva teoría económica del comportamiento va a pasar a ser el paradigma dominante en economía y ciencias sociales.

Probablemente ya lo sea a la hora de acometer proyectos de investigación, y esté empezando a serlo en la manera en que se explica la economía en muchas universidades y escuelas de negocios. Es más que posible que las generaciones futuras de estudiantes de economía estudien como una anomalía histórica, en alguna asignatura dedicada al devenir del pensamiento económico, la época en que todavía se consideraba la economía como una ciencia no experimental. Posiblemente entonces, desde la humildad del rigor y el método científico, se les explicarán los límites y virtudes de todo estudio económico experimental, en pie de igualdad con el resto de ciencias.

6. REFERENCIAS

- ALLAIS, M. (1953), «Le comportement de l'homme rationnel devant le risque», *Econometrica*, 21.
- ANDERSON, S. y SUNDER, S. (1989), «Professional traders as intuitive Bayesian», Canergie Mellon WP, 8851.
- ANDREONI, J. (1999), «Tax Compliance», *Journal of Economic Literature*, 36.
- BALL, S. B. y CECHE, P. (1996), «Subject Pool Choice and Treatment Effects in Economic Laboratory Research», en M. Isaac (ed.), *Research in Experimental Economics*, vol. 6, JAI Press.
- BALL, S. B. y HOLT, C. A. (1998) «Classroom Games: Bubbles in an Asset Market», *Journal of Economic Perspectives*, 12:1 (Winter), 207-218.
- BOLTON, G. (1991), «A comparative model of bargaining: theory and evidence» *American Economic Review*, 81.
- BRANDTS, J. y ABBINK, K. (2001), «The Spanish Auction for Government Securities: A Laboratory Analysis», mimeo, Instituto de Análisis Económico.
- BROSETA, B.; FATÁS, E. y NEUGEBAUER, T. (2003), «Asset Markets and Equilibrium Selection in Public Goods Games with a Provision Point. An Experimental Study», *Economic Inquiry*, 41, 4.
- CADSBY, Charles B. y MAYNES, Elizabeth (1999), «Voluntary provision of threshold public goods with continuous contributions: experimental evidence», *Journal of Public Economics*, 71, 53-73.
- CAMERER, C. (1999), «Behavioral Economics: Reunifying Psychology and Economics», *Proceedings of the National Academy of Sciences*, 96(September), 10575-10577.

- CAMERON, L. (1999), «Raising the Stakes in the Ultimatum Game: Experimental Evidence from Indonesia», *Economic Inquiry*, 37.
- CECH, P. A. (1988), «Removal of regulatory barriers to entry», *University of Arizona WP-8817*.
- CHAMBERLIN, E. (1948), «An experimental imperfect market», *Journal of Political Economy*, 56, 2.
- CHAMBERLIN, J. (1974), «Provision of Collective Goods as a function of group size», *American Political Science Review*, 68.
- CHARNESS, G. y RABIN, M. (2002), «Understanding social preferences with simple tests», *Quarterly Journal of Economics*, 117, 3.
- CHEN, Y. (2000), «Incentive-Compatible Mechanisms for Pure Public Goods: A Survey of Experimental Research», en Plott y Smith (eds.), *The Handbook of Experimental Economics*, Princeton University Press.
- COHEN, J. (1992), «White consumer response to Asian models of advertising», *Journal of Consumer Marketing*, 9, 2.
- CUMMINGS, R. G.; HARRISON G. y RUTSTROM, E. E. (1992), «Homegrown Values and Hypothetical Surveys: Is the Dichotomus Choice Approach Incentive Compatible?», *University of New Mexico Working Paper*, October.
- DAVIS, D. D. y HOLT, C. (1993), *Experimental Economics*, Princeton University Press.
- DEJONG, D.; FORSYTHE, R. y UECKER, W. (1988), «A note on the use of businessmen as subjects in sealed offer markets», *Journal of Economic Behavior and Organization*, 9.
- DUFFY, J. (1998), «Monetary Theory in the Laboratory», *Federal Reserve Bank of St. Louis Review*, Septiembre/October
- FATÁS, E.; NEUGEBAUER, T. y TAMBORERO, P. (2004), «How politicians make choices. A political choice experiment», *Journal of Economics*, en prensa.
- FATÁS, E. y ROIG, J. M. (2004), «Equidad y evasión fiscal. Un test experimental», *Revista de Economía Aplicada*, XII, 34, 17-37.
- FATÁS, E. y PALLARDÓ, V. (2004), «An experimental analysis of the ECB as a monetary policy committee», LINEEX WP 66-2004.
- FEHR, E. y GÄCHTER, S. (1998), «Reciprocity and Economics. The Economic Implications of Homo Reciprocans», *European Economic Review* 42, 845-859.
- FIORINA, M. y PLOT, C. (1991), «Committee Decisions under Majority Rule: An experimental study», *American Political Science Review*, 72.
- FOURAKER, L. E. y SIEGEL, S. (1963), «Bargaining Behavior», *Research Bulletin*, 9, 3.
- FRIEDMAN, D. y RUST, J. (1992), «Competitvity in Auction Markets: An Experiment and Theoretical Investigation», *UCLA Working Paper*, February.
- FRIEDMAN, D. y SUNDER, S. (1994), *Experimental Methods. A primer for Economists*, Cambridge University Press.
- GRETHER, D. y PLOTT, C. (1984), «The effects of market practices in oligopolistic markets: an experimental evaluation of the Ethyl case», *Economic Inquiry*, 22, 4.
- GÜTH, W. (1995), «On Ultimatum Bargaining Experiments –A personal Review», *JEBO*, 27.

- HEY, J. D. (1996), *Experimentos en Economía*, Fondo de Cultura Económica, México.
- HOFFMAN, E. y SPITZER, M. (1985), «Entitlements, Rights, and Fairness: An Experimental Examination of Subjects' Concepts of Distributive Justice», *Journal of Legal Studies*, 14, 2.
- HONG, J. T. y PLOTT, C. (1982), «Rate filling policies for inland water transportation: An experimental approach», *Bell Journal of Economics*, 13.
- KAGEL, J. H. (1995), «Auctions: A Survey of Experimental Research», en J. H. Kagel y A. E. Roth (eds.), *The Handbook of Experimental Economics*, Princeton University Press.
- KAGEL, J. H. y ROTH, A. E. (eds.) (1995), *The Handbook of Experimental Economics*, Princeton University Press.
- LEDYARD, J. (1995), «Public goods», en J. H. Kagel y A. E. Roth (eds.), *The Handbook of Experimental Economics*, Princeton University Press.
- LOEWENSTEIN, G. (1999), «Experimental Economics from the Vantage Point of Behavioral Economics», *The Economic Journal*, 109.
- McCABE, K. (1991), «Smart computer assisted markets», *Science*, 254.
- McCABE, K. A.; RASSENTI, S. J. y SMITH, V. L. (1994), «Auction Design for Composite Goods: The Natural Gas Industry», *Journal of Economic Behavior and Organization* 14** (June), 127-149.
- PENCAVEL, J. (1986), «Labor supply of men», en Ashenfelter y Layard (eds.), *Handbook in Labor Economics*, North Holland.
- PLOTT, C. (1982), «Industrial Organization Theory and Experimental Economics», *Journal of Economic Literature*, 20.
- (1987), «Dimensions of Parallelism: Some Policy Applications of Experimental Methods», en Roth (ed.), *Laboratory Experimentation in Economics*, Cambridge University Press.
- POPPER, K. (1959), *The Logic of the Scientific Discovery*, Routledge.
- POTTERS, J. y WINDEN, F. van, (1998), «The Performance of Professionals and Students in an Experimental Study of Lobbying», Discussion paper Tinbergen Instituut (TI 98-008/1).
- RASSENTI, S.; REYNOLDS, S. y SMITH, Vernon L. (1988), «Cotenency and Competition in an Experimental Auction Market for Natural Gas Pipeline Networks», *Economic Theory* 4: 41-65.
- ROTH, A. E. (1987), «Bargaining Phenomena and Bargaining Theory», en Roth (ed.), *Experimentation in Economics*, Cambridge University Press.
- (1995), «Introduction to Experimental Economics», en J. H. Kagel y A. E. Roth (eds.), *The Handbook of Experimental Economics*, Princeton University Press.
- SCHURR, P. H. (1987), «Effects of gain and loss decision frames on risky purchase negotiations», *Journal of Applied Psychology*, 72.
- SIAKANTARIS, N. (2000), «Experimental Economics under the Microscope», *Cambridge Journal of Economics*, 24.
- SLONIM, R. y ROTH, A. (1998), «Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic», *Econometrica*, 66.
- SMITH, V. (1982), «Microeconomic systems as an Experimental Science», *American Economic Review*, 72.

- SMITH, V. (1987), «Experimental methods in Economics», en Eatwell, Milgate y Newman (eds.), *The New Palgrave: A Dictionary in Economics*, vol. 2, Stockton Press.
- (1989), «Theory, Experiments and Economics», *Journal of Economic Perspectives*, winter.
- (1991), «Experimental Economics: Behavioral Lessons for Microeconomic Theory and Policy», en V. Smith (ed.), *Papers in Experimental Economics*, Cambridge University Press.
- (1994), «Economics in the Laboratory», *Journal of Economic Perspectives*, 8:1 (winter).
- SMITH, V.; McCABE, K. A. y RASSENTI, S. J. (1991), «Lakatos and Experimental Economics», en Smith, V. (ed.), *The Criterion of Empirical Progress*, Princeton University Press.
- STARMER, C. (1999), «Experimental Economics: Hard Science of Wasteful Tinkering?», *The Economic Journal*, 109.
- THALER, R. (1991), *Quasi Rational Economics*, Russell Sage Foundation.
- THURSTONE, L. L. (1931), «The Indifference Function», *Journal of Social Psychology*, 2.
- VAN WINDEN, F. (2001), «An experimental study of the Elswijk plan», mimeo CREED.
- VON NEUMANN, J. y MORGERSTERN, O. (1953), *Theory of Games and Economic Behaviour*, Princeton University Press, N.J..
- WELLS, D. (1991), «Laboratory experiments for undergraduate instruction in Economics», *Journal of Economic Education*, 22.
- WILLIAMS, A. y WALKER, J. (1993), «Computerized laboratory exercises for microeconomics education: Three applications motivated by the methodology of Experimental Economics», *Journal of Economics Education*, 14.