



Guía Metodológica

Guía metodológica sobre la aplicación de metodologías de economía experimental y del comportamiento para la evaluación de políticas públicas

Preparada por Diego Aycinena

Esta Guía Metodológica se ha realizado con aportación del proyecto de Apoyo a la Cooperación Triangular para la promoción de la Evaluación de Políticas Públicas en América Latina y el Caribe, **EVALÚA**, financiado por la Unión Europea y administrado por la FIIAPP F.S.P., que tiene como objetivo compartir experiencias en evaluación generadas en el ámbito de las instituciones públicas de América Latina, entre cuyos socios se encuentra el Departamento Nacional de Planeación.

Este documento ha sido elaborado con el apoyo financiero de la Unión Europea.

Este documento es resultado de un proyecto cofinanciado por la Unión Europea. Los contenidos de esta publicación en ningún caso reflejan las opiniones de la Unión Europea.

Este proyecto forma parte de:



Proyecto coordinado por:



Tabla de Contenidos

I. Introducción	4
II. Economía del Comportamiento	4
1. ¿Qué es la economía del comportamiento?	4
2. Preferencias no estándar	5
A. Aversión a las pérdidas	5
B. Preferencias relacionadas con el riesgo	5
C. Preferencias temporales	6
D. Preferencias y normas sociales	6
3. Creencias no estándar	7
A. Creencias motivadas	7
B. Sesgos de inferencia	7
C. Sobre-proyección o sesgo de proyección	9
4. Toma de decisiones no estándar	9
A. Contabilidad Mental	9
B. Marco de decisión	10
C. Atención limitada	10
D. Heurísticas	10
E. Persuasión y presión social	11
F. Emociones	11
5. Intervenciones Comportamentales y Nudges	11
III. Economía Experimental	14
1. ¿Qué es economía experimental?	14
2. Tipos de experimentos	17
<i>Experimentos en el Laboratorio</i>	17
<i>Experimentos de Campo Artefactuales</i>	19
<i>Experimentos de Campo Enmarcados (Framed)</i>	20
<i>Experimentos de Campo Naturales</i>	21
3. ¿Cuándo utilizar cada tipo de experimentos?	21
4. Juegos y experimentos tradicionales	23
A. Juego de la confianza	23

B. Juego del ultimátum	23
C. Juego del dictador	24
D. Juegos de cooperación (bienes públicos por contribuciones voluntarias)	24
E. Juegos de coordinación (esfuerzo mínimo o del eslabón más débil)	25
F. Decisiones individuales (preferencias temporales, riesgo, etc.)	25
5. Principios de diseño experimental	26
Bibliografía.....	31

Guía Metodológica

Guía metodológica sobre la aplicación de metodologías de economía experimental y del comportamiento para la evaluación de políticas públicas

I. Introducción

Esta guía presenta una introducción de los conceptos básicos de la economía experimental y la economía del comportamiento. Aunque muchas veces se utilice economía experimental y la economía del comportamiento como sinónimos, el lector verá que no lo son y se aclarará las diferencias y cómo se relacionan.

La economía del comportamiento es una rama que busca modificar y enriquecer los modelos económicos tradicionales incorporando elementos de psicología. Como verá en esta guía, la economía del comportamiento es mucho más que sesgos y *nudges*.

La economía experimental es una disciplina primordialmente metodológica para generar datos en forma controlada. Los métodos experimentales empezaron a llegar a la ciencia económica en forma consistente y sistemática, aunque gradual, entre mediados de 1950 y principios de los 1980's. Los primeros experimentos (en el laboratorio) en enfocarse en examinar políticas públicas se remontan a principios de la década de 1980, examinando políticas de transparencia en precios anunciados (Hong y Plott, 1982).

II. Economía del Comportamiento

1. ¿Qué es la economía del comportamiento?

La economía del comportamiento es un campo de la economía que usa evidencia de psicología y otras disciplinas para enriquecer los modelos teóricos económicos al incluir límites a la racionalidad, a la fuerza de voluntad y al interés propio.¹ No busca rechazar el modelo canónico, sino modificarlo y enriquecerlo.

Cuando se habla de economía del comportamiento, especialmente en círculos de hacedores de políticas públicas, la conversación usualmente se reduce a sesgos y *nudges*. Como se verá en esta guía, la economía del comportamiento es mucho más rica. Esta abarca el mismo campo de estudio que la economía tradicional, pero utiliza modelos teóricos de comportamiento modificados con el objetivo de enriquecerlos para mejorar el poder explicativo.

Para organizar los hallazgos de la economía del comportamiento, la presente guía utiliza en un marco conceptual basado en las desviaciones del modelo canónico que estudia al *homo*

¹ Loewenstein (2003) sugiere que la esencia de la economía del comportamiento es la aplicación de conocimientos psicológicos penetrantes a la economía. Dada la naturaleza de los enfoques de la psicología, estos elementos se encajan naturalmente en modelos teóricos microeconómicos, aunque también en algunos modelos macroeconómicos con micro-fundamentos.

economicus.² Esto se resume en que los individuos se desvían del modelo estándar en tres aspectos: preferencias no estándar, creencias no estándar y toma de decisiones no estándar. Las siguientes tres secciones examinan cada una de estas en mayor detalle.

2. Preferencias no estándar

Las preferencias juegan un papel central en economía, tanto para las elecciones de los agentes como en su bienestar (percibido). Esto va más allá de las preferencias por distintos bienes o servicios cuando éstos tienen distintos costos y beneficios. También abarca la forma como se sopesan beneficios y costos en distintos momentos en el tiempo (preferencias temporales), cuando los costos y beneficios son inciertos (decisiones ante incertidumbre y preferencias de riesgo), o cuando las decisiones involucran costos y beneficios que afectan distintas personas (preferencias y normas sociales). A continuación, se presentan los hallazgos de la economía del comportamiento relacionado con desviaciones de las preferencias estándar del modelo canónico.

A. Aversión a las pérdidas

Uno de los hallazgos más sólidos de la economía del comportamiento está relacionado con la aversión a las pérdidas (Tversky & Kahneman, 1991). Es tan sólido, que se encuentra presente no solo en humanos, sino también en otros primates (Chen et al., 2006). El concepto de aversión a las pérdidas indica que las desviaciones negativas respecto a un punto de referencia (pérdidas) son más poderosas que las desviaciones positivas. En palabras más sencillas, el dolor de las pérdidas es mayor que el placer de las ganancias.

Efecto dotación y sesgo del Status Quo. Un fenómeno que se manifiesta con la aversión a las pérdidas es el efecto dotación (Kahneman et al., 1990, 2012). De acuerdo al efecto dotación, existe la tendencia a valorar más un bien o servicio por el simple hecho de poseerlo. Es decir, para que una persona renuncie a un bien que posee, requiere que lo compensen más de lo que estaría dispuesto a pagar por obtenerlo.³ Esto se ve reflejado en el también conocido sesgo del *status quo* (Samuelson & Zeckhauser, 1988). Al evaluar cambios (por ejemplo, en programas sociales o políticas públicas), las personas son más sensibles a la dimensión en la que están perdiendo con relación al *status quo* (punto de referencia). Esto puede dificultar la implementación de nuevas políticas y cambios, pues las desventajas del mismo pueden percibirse mayores a las ventajas.

B. Preferencias relacionadas con el riesgo

La forma como se analiza la actitud frente al riesgo también ha cambiado en las últimas décadas gracias a las contribuciones a la economía del comportamiento desde la psicología. Entre éstas cabe

² Se dice que el modelo canónico estudia al mítico *homo economicus*, especie que tiene un egoísmo impenitente, fuerza de voluntad inquebrantable y capacidad de razonamiento ilimitada para calcular instantáneamente soluciones a problemas de optimización difíciles, cuyas soluciones determinan sus decisiones.

³ Kahneman et al. (2012, p. 194) ofrecen un ejemplo basado en un economista amante del vino que conocen, quien compró unos buenos vinos hace algunos años a precios bajos. El valor de mercado de dichos vinos se ha apreciado mucho de valor, por lo que una botella que originalmente le costó \$ 10 ahora costaría \$ 200 en una subasta. Este economista bebe ocasionalmente algo de este vino, pero no estaría dispuesto a vender el vino a dicho precio. Para renunciar a una botella requeriría que lo compensen más que el valor actual de mercado (el cual no estaría dispuesto a pagar por una botella).

destacar la teoría de prospectos (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). La teoría de prospectos postula que la actitud al riesgo no es igual para pérdidas y para ganancias.⁴

Otra contribución de la economía del comportamiento está relacionada con la *aversión miope a las pérdidas* (Benartzi & Thaler, 1995; Haigh & List, 2005; Thaler et al., 1997). La aversión miope a las pérdidas (*myopic loss aversion*) es la combinación de la aversión a las pérdidas en situaciones que involucran riesgo, con la frecuente evaluación de resultados (o lo que se conoce como *narrow bracketing*).⁵ Por ejemplo, este concepto nos dice que por la aversión a las pérdidas, evaluar un portafolio de inversión diariamente vs. hacerlo anualmente, puede llevar a elegir portafolios distintos. Al evaluar diariamente, para evitar el dolor de ver pérdidas (aunque pequeñas o infrecuentes) puede llevarnos a elegir portafolios más conservadores que en el largo plazo generan retornos mucho menores.

C. Preferencias temporales

El modelo tradicional asume que las personas tienen una fuerza de voluntad ilimitada para poder llevar a cabo sus planes en el tiempo y actuar de forma consistente. En la práctica, por otro lado, se observa que padecemos de sesgo presente o problemas de auto-control, y actuamos en forma inconsistente: planeamos empezar una dieta, un régimen de ejercicio o dejar de fumar, pero posponemos y lo postergamos constantemente. Nuevos modelos (Laibson, 1997; O'Donoghue & Rabin, 1999) permiten separar el grado de paciencia del grado de inconsistencia temporal (lo que se manifiesta en procrastinación o dilación y problemas de auto control). Esto explica por qué a veces utilizamos instrumentos de ahorro ilíquidos como mecanismo de compromiso para ayudarnos a sobreponernos a los problemas de autocontrol.

D. Preferencias y normas sociales

El modelo tradicional asume que los individuos intentan maximizar su bienestar sin tomar en consideración el de otras personas. Este supuesto es útil y realista en muchos contextos (como en mercados financieros), pero en otros no lo es. Los modelos de *preferencias sociales* (Bolton & Ockenfels, 2000; Cox et al., 2007; Dufwenberg & Kirchsteiger, 2004; Falk & Fischbacher, 2006; Fehr & Schmidt, 1999; Levine, 1998) permiten incorporar preferencias asociadas al bienestar de terceros, por ejemplo, altruismo, malevolencia, reciprocidad y retribución, etc.

Nuevos modelos de normas sociales (Bicchieri, 2005, 2010; Kimbrough & Vostroknutov, 2016; Krupka & Weber, 2013) permiten estudiar casos en los que las preferencias sociales sean condicionales al contexto (lo cual, por ejemplo, permite entender por qué en mercados financieros no se manifiestan preferencias sociales, y en otros casos sí). Los modelos de normas sociales sugieren que al elegir una acción no solo importan los beneficios y costos que se obtienen de dicha acción. También se toma en cuenta que tanto es acción es acorde a lo que el grupo de referencia (círculo social) normalmente hace y lo que éste considera apropiado hacer en esa situación. Estos

⁴ Además, postula que el peso que se le da a las probabilidades en las decisiones es distorsionado, dándole sobre-importancia las probabilidades pequeñas (efecto posibilidad) y sub-dimensionando la importancia de las probabilidades altas (efecto certeza).

⁵ Una anécdota famosa describe una conversación durante un almuerzo entre el Premio Nobel Paul Samuelson y un colega a quien le ofreció una apuesta tentadora: tira una moneda; cara ganas 200 dólares, cruz pierdes 100 dólares. El colega declinó la oferta de Samuelson pero expresó su disposición a aceptar una serie de 100 apuestas idénticas.

modelos deben ser tenidos en cuenta en el análisis de políticas públicas, pues sugieren que la percepción sobre el comportamiento de una masa crítica puede ser fundamental para el éxito o fracaso de algunas políticas públicas. Es decir, si las expectativas de cómo se comportan los demás y de cómo los demás creen que es correcto comportarse tienen influencia en el comportamiento individual, es importante tener en cuenta esto para la exitosa implementación de políticas públicas (por ejemplo, contribuciones a régimen pensional, pago en sistemas de salud de cobertura universal, etc.).

3. Creencias no estándar

Las decisiones que se toman se ven influenciadas por las creencias que tenemos, tanto normativas y del comportamiento de otros, como las creencias factuales: qué tan probables son ciertos resultados. Los modelos económicos tradicionales asumen racionalidad casi perfecta para adquirir información y procesarla de forma adecuada (de acuerdo al teorema de Bayes). A continuación, exploramos varias de estas desviaciones sistemáticas.

A. Creencias motivadas

Los humanos no somos máquinas buscadoras de la verdad (Bénabou & Tirole, 2016). Si bien no podemos escoger libremente nuestras creencias, estas no son independientes de nuestros gustos y preferencias (Epley & Gilovich, 2016). Nuestros gustos y preferencias influyen en la forma en que recolectamos información, procesamos argumentos y recuperamos las memorias. Esto nos permite *razonar* hacia conclusiones más acorde a nuestros gustos y preferencias, llegando a creencias sesgadas que se perciben objetivas. En especial, cuando nuestras creencias involucran algún aspecto sobre nuestra identidad, nuestro ego o afectan nuestra auto-imagen, nuestra racionalidad se ve comprometida. En estas situaciones tendemos a asimilar nueva información de manera sesgada. Nos enfocamos en la información que confirma nuestras creencias, especialmente aquella que fortalece nuestro ego o afecta positivamente nuestra imagen, mientras despreciamos la información que las contradice (Zimmermann, 2020). En otros casos, esta asimilación sesgada de la información puede llevar a que, para personas con creencias iniciales distintas, ¡la misma información los lleve a polarizar sus creencias! Resultados de esto se manifiestan como la persistencia de creencias, la asimilación sesgada de información y en algunos casos la polarización de las creencias (Lord et al., 1979; Rabin & Schrag, 1999).

Exceso de confianza. Hay evidencia considerable de que las personas (y especialmente los hombres) padecen de exceso de confianza tanto en sus habilidades, como en la precisión de sus creencias. El exceso de confianza es en muchos casos una manifestación de creencias motivadas, aplicadas a nuestras habilidades o nuestro conocimiento. Este exceso de confianza ha sido documentado y estudiado en mercados financieros (y sugiere que es esta diferencia de género en exceso de confianza la que hace que los hombres tengan menores retornos que las mujeres; Barber & Odean, 2001). Es importante reconocer que, como responsables de políticas públicas, no estamos exentos de sesgos. Esto es especialmente importante pues, cuando se toman decisiones en ambientes poco predecibles, los expertos son incluso más susceptibles a exceso de confianza.

B. Sesgos de inferencia

Hay muchos sesgos sobre cómo utilizamos información para formar y actualizar nuestras creencias que no necesariamente están relacionados con nuestras preferencias o nuestro ego (Kahneman, 2011). Este tipo de sesgos son importantes por su prevalencia generalizada y ayudan a explicar

fenómenos de muy diversa índole: cómo formamos estereotipos sociales que pueden perjudicar a ciertas poblaciones, cómo cometemos errores sistemáticos en inversiones financieras, o cómo médicos expertos pueden cometer errores al caer presa a estos sesgos. A continuación, se presentan algunos de los principales sesgos cognitivos que se han documentado que identifican como las personas nos desviamos del estándar de racionalidad perfecta, es decir, tenemos creencias poco racionales.

Representatividad. Utilizamos la representatividad cuando debemos evaluar la probabilidad de que un evento, persona o instancia pertenezca a una clase, tipo o proceso. Por ejemplo, para evaluar la probabilidad que un analista que realiza ciertas predicciones sea de un tipo habilidoso (en vez de mediocre que ha tenido suerte), que la persona que padece ciertos síntomas padezca cierta enfermedad. Por ejemplo, podemos catalogar a una persona como criminal o violenta por estereotipos de criminales violentos, sin importar que la mayoría de personas no son criminales ni violentos. Una instancia de representatividad está relacionada con *desatender la tasa base*. Al incorporar nueva información para nuestras creencias, debemos tomar en cuenta la tasa base o la probabilidad a priori de los resultados. Por ejemplo, al observar un resultado positivo en una prueba médica para una enfermedad, debemos considerar si en la población general dicha enfermedad es relativamente común o si es extremadamente rara. En casos de enfermedades extremadamente raras, el resultado de positivo en la prueba es más probable que se deba a un resultado falso positivo. Si no tomamos en cuenta lo raro de la enfermedad (la tasa base), vamos a sobreestimar la posibilidad de enfermedad.⁶

Disponibilidad. La disponibilidad (Tversky & Kahneman, 1974) sugiere que tendemos juzgar la frecuencia relativa o la probabilidad de ocurrencia de un evento por la facilidad con que instancias de dicho evento están disponibles en nuestra mente. Es decir, le damos un peso desproporcionado a casos que son prominentes, memorables o vívidos a costa de estadísticas abstractas y generales, pero más relevantes. Por ejemplo, tendemos a juzgar como más probables las muertes por hechos violentos o accidentes dramáticos (asesinatos, accidentes aéreos, etc.) que las muertes por causas naturales (infartos, problemas respiratorios, accidentes de tránsito, etc.) porque los primeros son más memorables y están fácilmente disponibles en nuestra memoria o en nuestra imaginación.

Anclaje y ajuste. Cuando debemos hacer una estimación o predicción numérica y tenemos algún valor disponible, frecuentemente nos *anclamos* a dicho valor. Incluso si los valores iniciales que obtenemos son arbitrarios (y sabemos esto), no ajustamos lo suficiente y estos tienden a ser muy pequeños e insuficientes (Slovic & Lichtenstein, 1971; Tversky & Kahneman, 1974).

Sesgo de comprensión retrospectiva (hindsight bias). Es la tendencia de las personas con conocimiento de un resultado, a creer falsamente que hubieran predicho el resultado informado de un evento (Fischhoff & Beyth, 1975; Hawkins & Hastie, 1990). Es decir, implica una exageración (no consciente) de la similitud entre sus creencias ex post (actuales) y las que serían/eran sus creencias antes de un evento informativo. Una vez que conocemos el resultado (que antes era desconocido o

⁶ Otro sesgo relacionado con la representatividad es la ley de los números pequeños. De acuerdo a ésta, tratamos información basada en muestras muy pequeñas como si fuera una muestra representativa y las tratamos como si fueran más precisas de lo que son. Esto nos lleva a sobre-inferir de fenómenos puramente aleatorios. Por ejemplo, podemos sobreestimar la habilidad de un analista mediocre, debido a que ha tenido tres aciertos (por pura suerte).

incierto), tendemos a creer que “*ya lo sabíamos*”, que era “*obvio*” o incluso “*inevitable*”. Por lo tanto, una vez que obtenemos cierta información, tendemos a subestimar el efecto que dicha información tuvo en nosotros.

C. Sobre-proyección o sesgo de proyección.

Otra forma en la que nuestras creencias se desvían sistemáticamente es que tendemos a juzgar nuestras preferencias futuras como muy parecidas a las actuales (Loewenstein et al., 2003). Es decir, proyectamos nuestras preferencias actuales al futuro.⁷ Por ejemplo, la mayoría de personas sabe qué sucede cuando vamos de compras al supermercado con hambre –muchos, a pesar de saberlo, terminamos comprando de más.

4. Toma de decisiones no estándar

Además de las preferencias y las creencias no estándar, a veces vemos que el proceso de toma de decisiones no es estándar y no se asemeja a lo que podemos analizar con los modelos tradicionales.

A. Contabilidad Mental

Contabilidad mental se refiere al conjunto de operaciones cognitivas utilizadas por los individuos y los hogares para codificar, clasificar y evaluar las actividades financieras (Thaler, 1985, 1999, 2012). La idea es que de la misma forma que las empresas utilizan sistemas contables, las personas utilizan un sistema contable mental para manejar sus “cuentas” financieras y sus decisiones. El área de contabilidad mental ha sido muy utilizada para comprender las decisiones de los consumidores. Si bien ha sido poco utilizada para políticas públicas, el potencial de su utilización en estos temas es muy grande.

El sistema de contabilidad mental tiene 3 componentes. El primero se refiere a la asignación de actividades a cuentas específicas (*budgeting & fungibility*). Este componente sugiere que las personas tienen “cuentas mentales” específicas a las que asignan ingresos, gastos, tiempo, etc. Esto, por ejemplo, se traduce en que quizá van a ser más reacios a usar ingresos “comunes” para gastos “superfluos”, pero que serán más propensos a realizar gastos superfluos y/o extraordinarios de ganancias que vienen de una fuente extraordinaria y superflua.⁸

Otro componente es la determinación del ciclo contable (*choice bracketing and dynamics*). Este hace referencia a la frecuencia con la que evaluamos los costos y los beneficios o con la que “cerramos” las cuentas. Por ejemplo, si compramos una acción, cae el precio y la mantenemos, la cuenta sigue abierta; la pérdida está en papel, aún no se realiza (pérdidas / ganancias en papel se sienten menos que las realizadas). Esto explica la disposición a vender acciones que suben de precio y mantener las que caen (Odean, 1998). También sugiere que tendemos a tener una preferencia por precios fijos en vez de precios por uso.⁹

⁷ Se puede interpretar como que *anclamos* en nuestras preferencias actuales y no ajustamos suficiente para proyectar las preferencias futuras.

⁸ Por ejemplo, gastar en una gran fiesta o en una vacación, fondos que vienen de un bono inesperado o de una ganancia de lotería o juego de azar.

⁹ ¿Cuántas personas se registrarían para asistir a un gimnasio que no tiene membresía, pero cobra una tarifa muy baja por kilómetro recorrido en la banda –aunque el gasto mensual en promedio sea el mismo?

B. Marco de decisión

La teoría económica tradicional asume que lo que determina las decisiones son los efectos o las consecuencias, no el marco como son presentadas o descritas (es decir, asume invarianza descriptiva). Sin embargo, un gran cuerpo de evidencia acumulada muestra lo contrario. Dos planteamientos lógicamente equivalentes (pero enmarcados en forma distinta) de un problema o decisión pueden llevar a los tomadores de decisiones a elegir diferentes opciones (Tversky & Kahneman, 1981). En términos coloquiales, una decisión puede verse afectada si el vaso que tiene agua hasta la mitad se presenta medio lleno en vez de medio vacío –aunque la cantidad de agua no cambie. Experimentos han demostrado que este efecto del marco de presentación de una decisión afecta incluso a médicos, dependiendo de si se les presenta como vidas salvadas/probabilidad de salvar vidas o fallecidos/probabilidad de muertes.

C. Atención limitada

El modelo tradicional asume que no hay restricciones de atención para poder obtener y procesar toda la información necesaria. Sin embargo, amplia evidencia de psicología demuestra que las personas tienen límites de atención. Esto es importante para estimar, por ejemplo, la respuesta de la demanda a cambios en impuestos indirectos que no son transparentes, por ejemplo, cuando no se incluyen sino hasta en el momento de pagar (Chetty et al., 2009). Dependiendo del objetivo de un impuesto –sea para recaudar fondos o para reducir demanda de bienes o servicios considerados no apropiados (alcohol, tabaco, etc.)- la prominencia y la transparencia del impuesto pueden ser fundamentales en la toma de decisiones de los individuos. Por otro lado, es fundamental tomar en cuenta la atención limitada en la forma para presentar información pública para que los ciudadanos tomen mejores decisiones (por ejemplo, decisiones relacionadas a la calidad de instituciones educativas, de salud, financieras, etc.).

D. Heurísticas

Cuando las personas deben tomar decisiones sobre un menú de opciones grande (por ejemplo, opciones de inversión, de planes de salud o de votación), la evidencia de psicología sugiere que utilizamos una serie de heurísticas o atajos mentales para simplificar el proceso de toma de decisión. La mayoría de la evidencia de estas heurísticas viene de decisiones financieras, debido a la abundante información sobre estas decisiones. Sin embargo, estas se usan de forma más general y aplican a varias posibles decisiones que involucren menús de opciones grandes. A continuación, evaluamos algunas de las principales.

Sobre-diversificación o heurística de $1/n$. Cuando enfrentan decisiones que permiten asignar cantidades de recursos entre múltiples opciones, las personas tienden a asignarlos en forma más o menos proporcional al número de opciones. Esta heurística es común en decisiones sobre ahorros o inversión, incluso ante expertos en finanzas: Harry Markowitz, quien recibió el nobel en economía de 1990 por el desarrollo de la teoría de elección de portafolios, admitió que dividía sus inversiones 50% en bonos y 50% en acciones!¹⁰

¹⁰ "Debí haber calculado las covarianzas históricas de las clases de activos y dibujar una frontera eficiente. En vez de eso... dividí mis contribuciones al cincuenta por ciento entre bonos y acciones", citado en Benartzi & Thaler (2007, p. 86).

Preferencia por lo familiar. Para simplificar decisiones complejas, frecuentemente se elige una opción que sea familiar –aunque la dimensión de familiaridad no sea relevante para la decisión. Ejemplos de esto son la preferencia por invertir en acciones de empresas del país o región del inversionista.

Preferencia por lo prominente. A pesar de no ser una estrategia racional, las personas tienden a enfocar su atención en aquellas opciones que son más prominentes; es decir, eligen entre las opciones que sobresalen, prestando menor atención al resto. Las opciones que sobresalen dependen del contexto. Por ejemplo, los candidatos a elección que aparecen como primera opción en la boleta, tienden a tener mayor intención de voto –especialmente cuando son candidatos “pequeños”, con una baja intención de voto.

Evitar decisiones. Otra heurística para lidiar con decisiones complejas es, sorprendentemente, evitar la decisión (Iyengar & Lepper, 2000). Por ejemplo, cuando se ofrecen muchas alternativas de planes inversión para ahorro para retiro. Ante esta situación, la opción por defecto (el *status quo*) tiene una importancia muy relevante, pues se ha encontrado que una alternativa es mucho más popular al ser el *status quo*, y la ventaja de una opción por defecto (default) aumenta conforme aumenta el número de alternativas.

E. Persuasión y presión social

En ocasiones, nuestras decisiones pueden verse influenciadas por terceros que intentan persuadirnos (Cialdini, 2006). Esta influencia de terceros puede darse a pesar de que tengan un interés en persuadirnos que puede estar en conflicto con nuestro bienestar o con decirnos la verdad. A pesar de esto, no descontamos lo suficiente la influencia de estos persuasores, que pueden ser agentes de ventas y mercadeo comerciales, agentes de mercadeo político, solicitantes para organizaciones sin fines de lucro, agencias de noticias o agencias para campañas de acciones públicas. Otro tipo de casos relacionados se dan en situaciones en las que la presión social influye en las decisiones. Así, nuestros pares influyen en decisiones positivas (mejoras en el desempeño (Mas & Moretti, 2009)) o negativas (consumo de alcohol (Kremer & Levy, 2008)).

F. Emociones

Muchos de los fenómenos antes mencionados (problemas de auto-control, preferencias sociales, normas sociales, sobre-proyección de preferencias) se deben a nuestras emociones y sentimientos. La literatura psicológica muestra el papel de las emociones como mecanismo mediador de nuestras acciones. Sin embargo, recientemente se han empezado a incorporar estos hallazgos psicológicos a fenómenos económicos (Loewenstein, 2010). Por ejemplo, el estado emocional o de ánimo de los agentes puede tener consecuencias en los mercados financieros o en elecciones (Breaban & Noussair, 2018; Healy et al., 2010; Hirshleifer & Shumway, 2003).

5. Intervenciones Comportamentales y *Nudges*

La Tabla 1 presenta una recapitulación de los principales hallazgos de la economía del comportamiento. Los hallazgos de la economía del comportamiento pueden ser utilizados para diseñar programas y políticas públicas efectivos y relevantes para decisiones económicas importantes. Thaler & Benartzi (2004) ilustra cómo hallazgos de la economía del comportamiento se han utilizado para diseñar intervenciones comportamentales exitosas que incrementan los

ahorros voluntarios en planes de largo plazo para el retiro. Para hacerlo, parten del problema de auto-control para reducir consumo e incrementar los ahorros en el presente. El modelo de elección inter-temporal (ver Preferencias temporales, p. 6) basado en sesgo presente aplicado a ahorros predice que las personas quieren incrementar sus ahorros, pero no en el presente sino más adelante. Cuando se llega al “más adelante”, este se convierte en el nuevo presente y se sigue posponiendo (indefinidamente) el sacrificio que implica el incremento en ahorros. Aprovechando el problema de autocontrol y sesgo presente ampliamente documentado, diseñaron un programa que ofrecía la posibilidad en el presente de comprometerse a ahorrar porcentajes de incrementos salariales futuros. El plan de ahorros llamado “*Save More Tomorrow*” mostró ser exitoso: más del 75% de los que recibieron la oferta del plan se unieron a este, y más del 80% de quienes lo hicieron permanecieron en el plan (pasados cuatro aumentos de sueldos). Para los participantes, la tasa de ahorro se incrementó casi cuatro veces, pasando del 3.5% al 13.6% del salario.

Tabla 1: Panorama general de los principales hallazgos de la Economía del Comportamiento

Preferencias no estándar	Creencias no estándar	Toma de decisiones no estándar
<ul style="list-style-type: none"> • Aversión a las pérdidas, sesgo del <i>status quo</i> • Descuento hiperbólico, inconsistencia temporal y problemas de auto-control • Preferencias sociales (altruismo, reciprocidad, etc.) y normas sociales (descriptivas y prescriptivas) 	<ul style="list-style-type: none"> • Creencias motivadas, exceso de confianza • Sesgos de representatividad, disponibilidad, anclaje, y de comprensión retrospectiva • Sesgo de proyección 	<ul style="list-style-type: none"> • Contabilidad mental • Efecto de marco (<i>framing</i>) • Atención limitada • Heurísticas de decisión (sobrediversificación, preferencia por lo familiar, por lo prominente y evitar decisiones) • Presión social y persuasión • Emociones

Una de las intervenciones comportamentales más frecuentemente utilizada o apelada para políticas públicas son los llamados *nudges*. Los *nudges* son intervenciones que buscan influir en las decisiones modificando la arquitectura de la toma de decisiones, sin modificar sustancialmente los incentivos ni restringir opciones, explotando los diversos hallazgos de la economía del comportamiento (Thaler & Sunstein, 2009).

Un ejemplo de este tipo de intervenciones, aplicado nuevamente a incrementar el ahorro voluntario en planes de pensiones, viene de Madrian & Shea (2001). Basado en hallazgos comportamentales de atención limitada y heurísticas que evitan la decisión, examinan el cambio de la opción por defecto (*default*) en los planes de ahorro para empleados en una corporación. En vez de que los empleados tuvieran que elegir activamente participar en el plan (y elegir entre las múltiples alternativas de tasas de contribución y asignaciones de portafolio), la empresa define un plan por defecto en el que todos participan automáticamente, a menos que explícitamente elijan no participar (o participar con un plan distinto). Es decir, las personas son libres de salirse de ese plan o modificar cualquiera de los componentes (la tasa de contribución y el portafolio para asignación de activos). A pesar que ninguna de las características económicas del plan cambió (y que las

personas eran libres de salirse), la participación en dicho plan incrementó significativamente.¹¹ Además, las tasas de contribución y los portafolios predeterminados elegidas por la compañía bajo la inscripción automática tuvieron una fuerte influencia en el comportamiento de ahorro.

Los *nudges* se han vuelto muy populares pues permiten hacer intervenciones de bajo costo y basadas en un *paternalismo suave o libertario* –sin recurrir a prohibiciones ni restricciones (Thaler & Sunstein, 2003). Además, se han documentado varios casos exitosos de intervenciones usando *nudges*, por ejemplo, para incrementar las donaciones de órganos (Johnson & Goldstein, 2003), dejar de fumar mediante contratos de compromiso (Giné et al., 2010), reducir muertes en quirófanos mediante *checklists* (Haynes et al., 2009), incrementar la tasa de repagos de préstamos mediante recordatorios personalizados (Karlan et al., 2012), incrementar el cumplimiento de pago de impuestos (Hallsworth et al., 2017; Kettle et al., 2016), entre otros.¹²

Sin embargo, los *nudges* deben diseñarse con cuidado, de lo contrario pueden ser inefectivos o incluso contraproducentes (Bicchieri & Dimant, 2019; Bolton et al., 2018; Wilson et al., 2015) o utilizarse con otros fines (*sludges*) (Thaler, 2018). Ericson (2017), por ejemplo, indica que bajo ciertas situaciones, “los recordatorios anticipados pueden provocar demoras adicionales, reduciendo tanto el bienestar como la probabilidad de que la tarea se complete”.

También es importante ser conscientes de que algunos de los sesgos que acabamos de ver (disponibilidad, asimilación sesgada y sesgo de confirmación) nos pueden llevar a sobreestimar la probabilidad de éxito de los *nudges*. Las intervenciones que muestran efectos positivos tienden a recibir más atención y aquellas que no muestran efectos tienden a ser ignoradas o a quedar almacenadas para el olvido. Además, se ha encontrado que para algunas intervenciones los efectos tienden a ser de corto plazo y se ignora efectos nulos a largo plazo (Ferraro & Price, 2013).

Es importante tener siempre presente que los *nudges* no son una panacea para resolver cualquier problema de políticas públicas. La probabilidad de éxito de una intervención comportamental o de un *nudge* va a ser tan buena como tan sólido sea el fundamento comportamental en que se base.

¹¹ La teoría económica tradicional nos dice que si las personas tienen preferencias, y el costo de cambiar de opción es bajo, no debería importar cuál es la opción por defecto, pues las personas cambiarán y elegirán su opción preferida. En este caso, a pesar que el costo de cambiar era muy bajo, el hecho de no necesitar tomar una decisión activa para tomar un plan podría interpretarse como si hubieran cambiado sus preferencias por tomar algún plan.

¹² Es importante resaltar que estos casos exitosos de nudges tiene un fundamento en al menos un hallazgo de la economía del comportamiento. Por ejemplo, Giné et al. (2010) se basa en hallazgos de descuento hiperbólico. Haynes et al. (2009) se basa en atención limitada, Karlan et al. (2012) se basa en reciprocidad u obligaciones personales, y los resultados de Hallsworth et al. (2017) y Kettle et al. (2016) se basan en conformismo a normas descriptivas o expectativas empíricas. También es importante aclarar que el hecho que una intervención tenga un fundamento teórico basado en hallazgos comportamentales no garantiza su éxito automático.

III. Economía Experimental

1. ¿Qué es economía experimental?

Los experimentos son el método científico, por excelencia. La economía experimental es el uso de métodos experimentales para generar datos bajo condiciones controladas.¹³ Por ejemplo, si queremos ver el impacto que tiene el número de empresas oferentes en el precio, un experimento nos permitiría variar sistemáticamente el número de empresas, manteniendo todos los demás factores constantes, y observar como el comportamiento de estas incide en el precio.

Friedman y Sunder (1994) definen los tipos de datos según el entorno en el que se obtiene y el proceso por el cual se genera. Los datos pueden ser recopilados en el entorno natural donde ocurre (datos de campo), o en un entorno artificial diseñado para examinar la pregunta que se requiere (datos de laboratorio). Según el proceso, los datos pueden generarse en forma natural por procesos no controlados (datos casuales o *happenstance*), o por un proceso deliberado, diseñado en un ambiente controlado. La Tabla 2 muestra las distintas combinaciones de tipos de datos que pueden generarse de acuerdo con esta taxonomía.¹⁴

Tabla 2: Tipos de datos (Friedman y Sunder, 1994)

		Proceso de generación de datos	
		Casual: generada en forma natural por procesos no controlados	Experimental: deliberada, generada en ambiente controlado
Entorno en el que se generan los datos	Campo: recopilada en el entorno natural donde ocurre	Datos macro (PIB, inflación, tasa de cambio, retornos financieros, etc.) Datos administrativos (resultados de pruebas estandarizadas educativas, criminalidad, ingresos, etc.) Encuestas de hogares y de firmas	<i>RCT's</i> y Experimentos para evaluación de impacto (P. ej: educación, cumplimiento fiscal, ahorros y remesas, educación) <i>Nudges</i> (Donación de órganos, ahorro en consumo de agua & energía, contribuciones a pensiones, etc.)
	Laboratorio: recopilada en un ambiente "artificial"		Juegos para medir pro-socialidad (ultimátum, dictador, bienes públicos, confianza, esfuerzo mínimo) Subastas, diseño de mercados complejos, mercados para permisos de emisiones Preferencias de riesgo y temporales

¹³ Smith (1976, p. 275) define el control como la esencia de la metodología experimental.

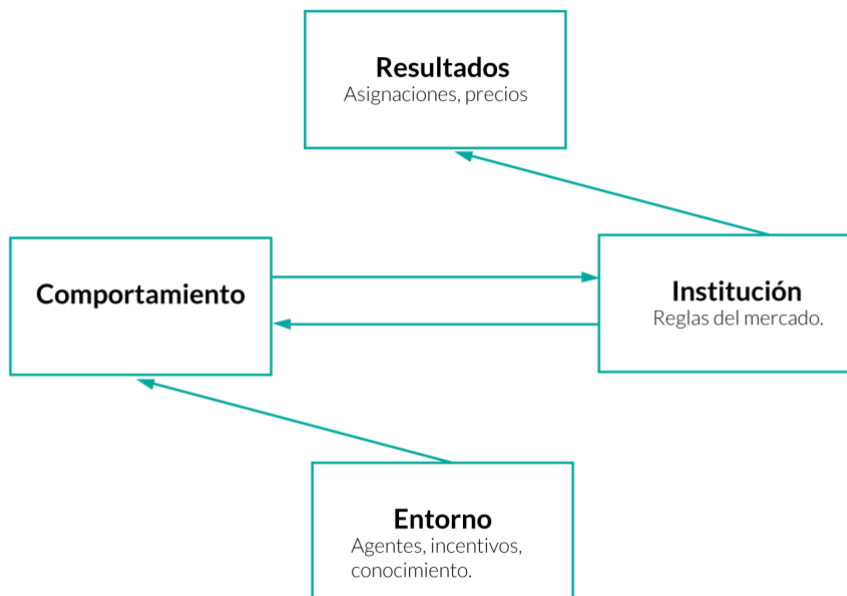
¹⁴ Los casos mostrados en la tabla no son exhaustivos. Por ejemplo, los *nimble RCTs*, son un subconjunto de los RCTs que combinan datos de intervenciones experimentales de campo con datos de registros administrativos, por la tanto, abarcan más de una categoría.

El enfoque de esta guía es en la metodología experimental que genera datos en forma deliberada en un proceso diseñado y controlado por el investigador. Los datos generados por métodos experimentales pueden hacerse en distintos entornos. En un extremo está el laboratorio, en el cual se crea un entorno artificial diseñado para estudiar un problema de interés específico; en el otro está el campo, en el cual se manipulan los procesos de tal forma que puede estudiar el problema en el entorno natural en el que ocurre. Entre esos dos extremos existen entornos intermedios, y cada uno presenta algunas ventajas y algunas desventajas. Más adelante (p. 17) se explorarán los distintos tipos de experimentos con sus respectivas ventajas y limitaciones.

Marco Conceptual de los Experimentos de Laboratorio

En un experimento de laboratorio, se introduce en un entorno artificial una situación en la que individuos tomando decisiones motivadas por ganancias monetarias generan datos en forma controlada para examinar una variable de interés. Un experimento económico genera, en un ambiente controlado, la situación económica que queremos analizar para luego poder hacer variantes y comparar los resultados entre dichas variantes. Los resultados experimentales observados son la consecuencia de las elecciones de los individuos, que se dan en un entorno económico, mediadas por reglas definidas por las instituciones. La Ilustración 1 muestra el marco conceptual para los experimentos de laboratorio de Smith (1982), basado en los elementos que definen un sistema microeconómico: el **entorno**, la **institución**, además del **comportamiento** de los agentes en dicho entorno e institución, y los resultados que esto produce.

Ilustración 1: Elementos de un Sistema Microeconómico (adaptación de Smith, 1982)



El *entorno* económico consiste en los agentes que interactúan en el sistema, las preferencias y/o costos de los agentes, sus conocimientos y habilidades, y las restricciones que enfrentan.¹⁵ En el

¹⁵ Un monopolista o un oligopolio, ofreciendo bienes homogéneos o heterogéneos, con curvas de costos iguales o distintas; varios compradores interesados en adquirir un bien cuyo valor es el mismo para

laboratorio, el entorno es controlado a través de recompensas monetarias para inducir la estructura de valores/costos deseada (Smith, 1976, 1982).

La *institución*, se refiere al mecanismo o algoritmo que define las reglas y los procedimientos bajo los cuales los agentes pueden interactuar. Esto incluye reglas referentes a las condiciones y los tipos de *mensajes*, así como procedimientos para procesar los mensajes y convertirlos en transacciones, asignar bienes y recursos, imputar precios y costos, y manejar la información. Por ejemplo, en una subasta inglesa, los oferentes pueden enviar como *mensaje* cualquier precio mayor al último ofertado; en una subasta holandesa o descendente, el precio inicia a un nivel arbitrariamente alto y va descendiendo hasta que el primer oferente envía como *mensaje* que acepta comprar a dicho precio. La regla de asignación indica cómo se asigna el bien de acuerdo con los mensajes recibidos: la persona con la oferta más alta (la última) en el caso de la subasta inglesa, o la primera en ser aceptada en la subasta holandesa. La regla para imputar precios indica qué precio debe pagar cada oferente según los *mensajes* enviados.¹⁶

El *comportamiento* hace referencia a las acciones estratégicas que toman agentes en función del entorno y la institución. Éstos se manifiestan por medio de los *mensajes* enviados (ofertas, solicitudes, aceptaciones, precios publicados, compras, etc.) según las reglas de la institución. A través de las reglas vigentes, la institución define como el comportamiento se traduce en los resultados. Es decir, la institución es el algoritmo cuyas reglas procesan los "mensajes" de los participantes, y los traducen en los resultados del sistema microeconómico –asignación final, precios, costos, eficiencia del sistema, etcétera.

Ventajas de los datos experimentales

Los experimentos proporcionan dos ventajas importantes sobre los datos que se genera en forma casual: control y replicabilidad. Por un lado, los datos generados por experimentos permiten mayor control para tener una inferencia causal clara y robusta, y para aislar en forma directa e indirecta de otros efectos. En los experimentos es posible controlar directamente las instituciones (P. Ej.: reglas de votación, reglas de subasta, reglas de pagos, comunicación, etc.), los incentivos (pagos de los sujetos) y el entorno. A pesar de esto, no siempre se tiene el control directo completo (por ejemplo, sobre el altruismo, rencor y preferencias sociales, normas sociales y creencias de los sujetos). En esos casos, se utiliza la aleatorización como control indirecto, lo cual evita la correlación de tratamientos con características individuales. Es decir, la asignación aleatoria a tratamientos permite la variación controlada sin la confusión que puede generar variables importantes omitidas y eventos que ocurren en forma simultánea.

Supongamos que utilizamos un experimento para examinar la relación causal de x (numero de empresas) sobre y (precio). En este caso, intentamos aislar el impacto de x (la variable de enfoque) sobre y (la variable de resultado), del resto de factores w, z , etc. (las variables estorbo, por ejemplo,

cualquiera –valor común- pero no lo conocen con exactitud, o cuyo valor es privado e independiente –cada comprador conoce con exactitud cuánto valora el bien, son ejemplos que describen distintos entornos.

¹⁶ En las subastas a sobre cerrado, todos los participantes realizan sus ofertas en forma simultánea –por ejemplo, cada oferente anota su oferta y la envía en un sobre. El subastador asigna el bien a quien haya realizado la oferta más alta, pagando su propia oferta –en el caso de subastas de primer precio (SPP)- o pagando la segunda oferta más alta –en el caso de subastas de segundo precio (SSP).

el tipo de productos que se ofrecen, la posibilidad de comunicación, las reglas institucionales, el número de compradores, etc.) que también afectan *y*. El control directo de las variables importantes es lo que distingue datos casuales (*happenstance*) de datos experimentales. Los datos experimentales nos permiten controlar y medir variables de estorbo para aislar las variables de enfoque.

Por otra parte, los datos experimentales son replicables en la medida que permiten construir conocimiento basándose en hallazgos previos. Esto facilita un proceso acumulativo y sistemático de aprendizaje experimental. Esta ventaja puede parecer como específica a la academia y poco importante para las políticas públicas. Sin embargo, es vital poder construir conocimiento generalizado que respalde las políticas, de lo contrario cada política necesitaría una justificación específica.

Otra ventaja de los datos experimental es que muchas veces nos permite observar directamente variables que de otra forma no sería posible observar. Por ejemplo, en el laboratorio es posible:

- Inducir y observar los valores que tienen los compradores y los costos que tienen los vendedores,
- Observar los precios de todas las transacciones, incluso las ofertas que no son aceptadas,
- Observar la comunicación entre agentes o los votos (secretos)
- Observar acciones antisociales que fuera del laboratorio se buscaría mantener ocultas (y solo sería posible observar las denuncias), como tomar dinero ganado por terceros, o sub-declaraciones de ingresos para evitar “impuestos” y obligaciones.

2. Tipos de experimentos

Existen varios tipos de experimentos y varias taxonomías para clasificarlos (Charness et al., 2013; Gneezy & Imas, 2017; Plott, 1987). La más utilizada (Harrison & List, 2004), define cuatro grandes grupos de experimentos. La Tabla 3 resume las características de los distintos tipos de experimentos. A continuación, se presentan junto con varios ejemplos de cada uno.

*Experimentos en el Laboratorio*¹⁷

Los experimentos de laboratorio tradicionales típicamente se llevan a cabo, como su nombre lo indica, en un laboratorio con computadores.¹⁸ En este tipo de experimentos se utiliza un entorno

¹⁷ Es común que los hacedores de políticas públicas y tomadores de decisión suelen ver la evidencia de experimentos de laboratorio con suspicacia. Sin embargo, el uso de experimentos en el laboratorio para informar decisiones relacionadas a políticas públicas no es nuevo. Grether & Plott (1984) y Hong & Plott (1982) son quizás los primeros experimentos de laboratorio diseñados para responder a problemas concretos de políticas públicas. De hecho, ya para finales de la década de 1980's Estados Unidos estaba utilizando experimentos de laboratorio para informar decisiones de políticas públicas de la *Federal Trade Commission*, para el departamento de transporte, etc. Plott (1987) ya hacía una primera compilación de experimentos en el laboratorio relacionados a políticas públicas.

¹⁸ En muchas ocasiones los participantes en los experimentos de laboratorio toman sus decisiones e interactúan a través de los computadores (comúnmente utilizando software especial para el propósito, como puede ser *oTree* o *zTree*, (Chen, et al., 2016; Fischbacher, 2007)). Sin embargo, nada de esto es necesario. Existen muchos experimentos de laboratorio que se llevan a cabo en salones de clase y las decisiones se toman con papel y lápiz.

inducido, con terminología/marco abstracto, reglas/instituciones definidas e implementadas por el experimentador, y típicamente utiliza estudiantes universitarios como participantes.

Tabla 3: Resumen de las características de los distintos tipos de experimentos

Características	Experimento de Laboratorio	Experimento de Campo Artefactual	Experimento de Campo Enmarcado	Experimento de Campo Natural
¿Quién participa?	Usualmente estudiantes universitarios (muestra de conveniencia)	Población “natural” (muestra de la población de interés)	Población “natural” (muestra de la población de interés)	Población “natural” (muestra de la población de interés)
¿Dónde se lleva a cabo?	Laboratorio experimental	En un “laboratorio en el campo”	En el campo	En el campo
¿Qué información o experiencia llevan los participantes?	Usualmente ninguna específica o relevante (el aprendizaje se da en el laboratorio)	Pueden, o no, llevar experiencia en la tarea, con las reglas o con las características de los bienes/servicios	Llevan experiencia en la tarea, con las reglas o con las características de los bienes/servicios en cuestión	Llevan experiencia en la tarea, con las reglas o con las características de los bienes/servicios en cuestión
¿Cuál es el entorno microeconómico? (Naturaleza del bien o servicios)	Valores y/o costos, inducidos por el experimentador, usualmente con bien/servicios y un marco/terminología abstractos	Valores y/o costos, inducidos por el experimentador, usualmente con bien/servicios y un marco/terminología abstractos	Pueden ser inducidos por el experimentador, pero con un marco/terminología acorde al entorno natural	Valores y/o costos naturales (<i>homegrown</i>) con un marco o terminología acorde al entorno natural
¿Cuál es la institución, las reglas o la tarea?	Definidas e implementadas exógenamente por el experimentador	Definidas e implementadas exógenamente por el experimentador	Pueden ser definidas y/o implementadas exógenamente por el experimentador, o las que se dan naturalmente en el ambiente	Usualmente son definidas e implementadas naturalmente en el ambiente
¿Cuál es la naturaleza del ambiente en el que operan los participantes?	Artificial	Artificial	Combinaciones de elementos naturales y artificiales	Es en el que los participantes participan o interactúan naturalmente

Si bien muchos se enfocan en estudiar mecanismos formales como subastas y otras instituciones de mercado, las aplicaciones son muy variadas incluyendo varias relevantes para políticas públicas:

- Diseños de mercados para agua (Murphy et al., 2000; Murphy et al., 2009), de gas (McCabe et al., 1989; Rassenti et al., 1994) y de energía eléctrica (Abbink et al., 2003; Brandts et al., 2008; Rassenti et al., 2003a, 2003b). Por ejemplo, Rassenti et al. (2003a) examinan el impacto de la demanda activa en precios y control del poder de mercado local en mercados de energía eléctrica.
- Esquemas de pagos e incentivos por servicios médicos (Bejarano et al., 2017; Henning-Schmidt et al., 2011). Por ejemplo, Brosig-Koch et al. (2017) examinan el efecto de planes de pago mixtos en la sobre- y sub-provisión de servicios médicos.
- Regulaciones en mercados de gasolina (Deck & Wilson, 2008; Haucap & Müller, 2014).
- Evasión fiscal y sub-reporte de ingresos (Cummings et al., 2009). Por ejemplo, Coricelli et al. (2010) examinan la relación entre las emociones, falsos reportes y la toma de decisiones racional a través de un experimento de laboratorio sobre auto-reporte de ingresos y evasión fiscal.
- Mecanismos de subastas para el espectro radioeléctrico (Banks et al., 2003; Kagel et al., 2010)
- Tasa de repago y opciones de préstamos para estudios universitarios (Abraham et al., 2018a, 2018b; Cox, Kreisman, & Dynarski, 2018).
- Mercados de bienes y servicios ambientales (Cherry et al., 2008; Friesen & Gangadharan, 2013) y mercados para permisos de emisión de contaminantes (Cason, 2010).
- Políticas para prevenir o combatir la corrupción (Abbink, 2004; Abbink & Serra, 2012; Abbink & Wu, 2017; Ryvkin et al., 2017; Serra, 2012).¹⁹ Por ejemplo, Abbink et al. (2014) examinan los efectos de la penalización asimétrica de las partes involucradas en un soborno y Butler et al. (2017) los efectos de incentivos monetarios para denunciar la participación en una transacción ilegal que involucra un soborno.
- Cox, et al. (2002) evalúan, a pedido del ministerio de transporte de Holanda, políticas alternativas para incrementar la competencia en el transporte de trenes.

Experimentos de Campo Artefactuales

Al igual que el anterior, los experimentos de campo artefactuales se dan con un entorno inducido, utilizando terminología/marco abstracto, con reglas/instituciones definidas e implementadas por el experimentador, pero con una población distinta (no estudiantes universitarios). Es decir, la diferencia primordial de estos experimentos, con respecto a los de laboratorio, es que se hacen con la población de interés o con una población que se considera “más relevante” para el problema en cuestión. Comúnmente, esto implica tener que acercarse a la población de interés, en vez de llevar a la población de interés al laboratorio.

Un uso muy común de estos experimentos es para medir parámetros de interés en poblaciones particulares. Por ejemplo, parámetros de preferencia temporal (paciencia, descuento hiperbólico) o riesgo (Binswanger, 1980; Tanaka et al., 2010). Aycinena et al., (2019) examinan las preferencias temporales y control intra-hogar en una muestra de beneficiarias de programas de transferencias monetarias condicionadas en Guatemala. Callen et al. (2014) y Jakiela & Ozier (2015) examinan el

¹⁹ Una posible crítica a los experimentos de corrupción en el laboratorio es la validez externa de los mismos. Armantier & Boly (2013) examinan esto comparando resultados del mismo experimento en tres distintos entornos: en el laboratorio en Canadá y en el laboratorio y campo en Burkina-Faso. No encuentran diferencias estadísticamente significativas y concluyen que los experimentos de laboratorio son informativos y tienen relevancia empírica.

impacto de la exposición a la violencia en la actitud ante el riesgo/incertidumbre en Afghanistan y Kenia, respectivamente. Attanasio et al. (2012) examinan el papel de las redes sociales, y la actitud al riesgo en las decisiones de agrupar y diversificar el riesgo (*risk pooling*) en una muestra de participantes del programa de transferencias monetarias condicionadas en Colombia (Familias en Acción).

Otro ejemplo de su uso es para medir parámetros relacionados con el capital social (como confianza, reciprocidad, altruismo, etc.) en poblaciones de interés (Karlan, 2005). Por ejemplo, Friebel et al. (2018) miden confiabilidad en personas que aplican para el cuerpo policial en Alemania. Para políticas públicas, estos experimentos son sumamente útiles cuando estas poblaciones han sido expuestas a diversas intervenciones, como programas de transferencias condicionadas (Attanasio et al., 2008), o en combinación con RCTs (en español, experimentos aleatorios controlados). Jakiela et al. (2015) introducen un experimento de campo artefactual (diseñado para medir respeto a derechos de propiedad de otros), en la muestra de un RCT que asignaba aleatoriamente a niñas a recibir una beca educativa. Con esta combinación pueden medir el impacto causal de mayores niveles de educación (generados por la beca) sobre capital social (respeto a derechos de propiedad de otros).

Experimentos de Campo Enmarcados (Framed)

Los experimentos de campo enmarcados, al igual que los experimentos de campo artefactuales, se realizan con una población no estudiantil, con reglas/instituciones definidas e implementadas por el experimentador, pero llevados al contexto de campo, ya sea en la tarea, bien/mercancía o información que los sujetos pueden usar.

Un ejemplo común son experimentos de contribuciones a bienes públicos o extracciones de recursos de uso común (Cardenas, 2000, 2004; Santis & Chávez, 2015; Velez, et al., 2009, 2010). Si bien estos iniciaron en el laboratorio con estudiantes universitarios bajo un marco general –es decir, como experimentos tradicionales de laboratorio- ha sido común “sacarlos” al campo para comprender cómo las poblaciones de interés, utilizando decisiones contextualizadas, responden a distintos esquemas e intervenciones para incrementar las contribuciones o mantener las extracciones a niveles sostenibles. Por ejemplo, (Janssen et al., 2013) examinan distintos mecanismos para solucionar dilemas de uso de agua para irrigación mediante experimentos de campo enmarcados con diversas poblaciones en Colombia y Tailandia.

Brosig-Koch et al. (2016) realizan experimentos de campo enmarcado con médicos y estudiantes de medicina sobre la provisión de servicios médicos bajo distintos esquemas de incentivos. Cox, et al. (2016a, 2016b) utilizan experimentos con estudiantes de medicina y residentes para comparar mecanismos de pago por desempeño que intentan alinear los incentivos de médicos y hospitales para reducir costos de readmisiones, sin sacrificar calidad. Hill & Viceisza (2012) implementaron un experimento enmarcado en Etiopía para examinar la hipótesis de que los micro-seguros inducen a los pequeños agricultores a tomar mayores riesgos rentables. Ambler (2015) examina el impacto en asimetrías de información sobre las decisiones de remesas de migrantes. Barr et al. (2009) realizan un experimento de campo enmarcado en Etiopía con un juego de corrupción (apropiación indebida) con diversos agentes (incluyendo fiscalizadores), enmarcado en un contexto de prestación de servicios de salud.

Experimentos de Campo Naturales

En los experimentos de campo naturales, al igual que los experimentos de campo enmarcados, se usa el contexto de campo, ya sea en la tarea, bien/mercancía o información que los sujetos pueden usar, con la población en su entorno natural. Sin embargo, en este caso no se induce el entorno, sino que se da en el entorno donde los sujetos naturalmente realizan estas tareas y, típicamente, los sujetos no saben que están en un experimento.

Muchos *RCTs* (*randomized controlled trials* o, en español, experimentos aleatorios controlados), frecuentemente utilizados como método de evaluación de impacto de intervenciones para el desarrollo, son un caso común (un subconjunto) de experimentos de campo naturales. Por ejemplo, solamente en temas de educación, cientos de experimentos han examinado intervenciones para aumentar la demanda por servicios educativos (Ambler et al., 2014a; Barrera-Osorio et al., 2011; Jensen, 2010) o evaluar impacto de insumos educativos (Beuermann et al., 2015; Glewwe et al., 2009), de cambios en la pedagogía (Banerjee et al., 2016; Muralidharan et al., 2019), o en la gobernanza de las escuelas (Duflo et al., 2015; Glewwe & Maïga, 2011).

En temas relacionados con economía política, se han utilizado experimentos de campo para evaluar distintas intervenciones orientadas a incrementar la participación electoral (Gerber & Green, 2000; Green et al., 2013). Armantier & Boly (2011, 2013) diseñan experimentos de campo naturales para medir la corrupción en donde se proponían sobornos a calificadoros de exámenes. Otra aplicación es la medición de reducciones en consumo de agua o energía (Ferraro & Price, 2013; Schultz et al., 2016; Schultz et al., 2007) o el incremento del reciclaje (Schultz, 1998). También han sido muy utilizados para determinar la respuesta a distintas intervenciones para mejorar el cumplimiento fiscal (Castro & Scartascini, 2015; Kettle et al., 2016; Kleven et al., 2011). Por ejemplo, Hallsworth et al. (2017) realizan dos experimentos de campo naturales en el Reino Unido variando mensajes de normas y mensajes de servicios públicos en cartas de recordatorio de pagos por impuestos no pagados.

3. ¿Cuándo utilizar cada tipo de experimentos?

No hay reglas fijas con respecto a cuándo utilizar un tipo de experimento u otro. Sin embargo, las ventajas y desventajas de cada uno proveen una guía. Los experimentos del laboratorio (y en algunos casos, los artefactuales de campo) tienen varias ventajas: relativos bajos costos (tanto financieros como de logística), y son altamente flexibles. En general, es conveniente usar este tipo de experimentos cuando:

- 1) No se puede observar la variable de interés en un entorno natural. Por ejemplo, variables sobre corrupción, o sobre decisiones judiciales (donde se debe averiguar el verdadero estado de culpabilidad/inocencia de una persona). O en donde las valoraciones/costos son cruciales (y no es creíble que los individuos/firmas las revelen verazmente), por ejemplo, comparando mecanismos de subasta o reglas de mercado distintas.
- 2) Quiere evaluarse el desempeño de distintas instituciones (o el diseño de nuevas instituciones), o determinar qué sucedería en un entorno distinto. Estos casos son muy difíciles sino imposibles de evaluar en el campo (al menos a un costo razonable). Ejemplos de experimentos al respecto examinan (i) cambios a reglas de precios en ciertos mercados (mercados de gasolina, mercados de servicios de salud), (ii) diseños institucionales con nuevos mecanismos de mercado (mercados mayoristas de energía eléctrica, gas, agua y

mercados de servicios medioambientales), o (iii) diseños de subastas con aplicaciones concretas (subastas para el espectro radioeléctrico o para permisos intercambiables de emisión de contaminantes), entre otros.

- 3) Se tiene un presupuesto limitado y/o plazos cortos (y no existe información que este fácilmente disponible para responder la pregunta de interés). Es muy útil cuando se tiene poco tiempo y desea reducir el número de opciones de políticas públicas a considerar o se desea tomar decisiones más informadas.
- 4) Como prueba de concepto. Cuando se desea demostrar a tomadores de decisión la viabilidad de implementar nuevas reglas, políticas o mecanismos. En este sentido, los experimentos de laboratorio son sumamente útiles pues no solamente permiten mostrarles evidencia clara, sino que es posible diseñar experimentos demostrativos en los que los tomadores de decisión participen y observen los resultados. Por ejemplo, en Australia, investigadores de la Universidad del Nuevo Gales del Sur han utilizado experimentos como herramienta demostrativa con tomadores de decisiones políticas para demostrar el poder de distintos mecanismos de mercado para mercados regulados (energía, gas, etc.).

Los experimentos de campo son muy útiles cuando se tiene un especial interés en examinar el comportamiento de una población específica –que se sospecha sería distinto a una población estándar- y es factible hacerlo a un costo razonable. Los principales retos de este tipo de experimentos están relacionados con aspectos logísticos y con el costo de oportunidad del tiempo de los participantes, por lo cual muchas veces es necesario simplificar el experimento.

Los experimentos de campo enmarcados son útiles cuando no solo se tiene un especial interés en examinar el comportamiento de una población específica, sino que el contexto de campo del problema a estudiar puede ser importante.²⁰ Esto es primordial si la población que se está estudiando para el tipo de problema en cuestión recurre a heurísticas específicas o incorporan preferencias distintas a las que se usan en el laboratorio o en experimentos artefactuales.

A veces se suele agrupar los experimentos de campo artefactuales y enmarcados (Gneezy & Imas, 2017; Viceisza, 2016).²¹ En general, es conveniente usar estos tipos de experimentos en circunstancias similares a las de los experimentos de laboratorio, pero cuando se quiere aprender algo específico a una población particular. Ejemplos concretos son:

- 1) Como herramienta de medición; para identificar y estimar parámetros asociados a características de interés de una población particular, por ejemplo, medición de preferencias inter-temporales, actitud al riesgo, capital social o prosocialidad, etc.
- 2) Para evaluar los resultados sobre una población específica de interés. Un ejemplo de esto son los experimentos sobre manejo de recursos de uso común.
- 3) Para probar teorías sobre mecanismos o heurísticas de decisión de poblaciones específicas.

Los experimentos de campo naturales son más apropiados cuando el entorno natural completo en el que se dan las decisiones es crucial. Por ejemplo, Ambler et al., (2014b) y Aycinena et al., (2010)

²⁰ El contexto de campo puede ser el bien o servicio o el tipo de participantes con el que la población de interés está familiarizada.

²¹ Estos experimentos son también llamados *lab-in-field* experiments, pues para ambos implica “llevar” o “acercar” el laboratorio a los participantes.

examinan qué tan sensibles son al precio los envíos de remesas de migrantes centroamericanos en Estados Unidos.²² Sería prácticamente imposible obtener estimaciones de esto con alguno de los otros tipos de experimentos. En estos casos, el entorno no es inducido a través de recompensas, sino que se toma el entorno natural y se “manipula” exógenamente (en el ejemplo citado, el precio para enviar remesas) para generar variación controlada del experimento.

Para resumir, los experimentos de laboratorio, artefactuales y enmarcados tienen menores costos y ventaja comparativa en ayudarnos a comprender los mecanismos y en darnos información valiosa antes de implementar una política o programa mediante pruebas a escala.²³ Los experimentos de campo naturales tienen ventaja comparativa en darnos retroalimentación y evaluar el impacto de políticas o programas implementadas (*policy experiments*), o para medir tamaño de efectos en la población específica de interés en su contexto natural. Sin embargo, debe quedar claro que no hay una jerarquía de experimentos –simplemente tienen distintas ventajas y limitaciones. Además, estos enfoques tienden a ser complementarios, como se han podido ver ejemplos de combinaciones de experimentos de campo artefactuales con experimentos de campo naturales.

4. Juegos y experimentos tradicionales

Los juegos y tareas experimentales son tan variados como la creatividad de los investigadores. Sería prácticamente imposible hacer un listado completo de todos los juegos y tareas que se han utilizado en experimentos. Sin embargo, resulta conveniente hacer un resumen de los juegos y tareas más comunes.

A. Juego de la confianza

La confianza es un aspecto fundamental de las ciencias sociales. Algunos se refieren a la confianza como un primitivo universal (que está presente en todas las culturas humanas conocidas). En economía, usualmente se examina la confianza y la reciprocidad mediante el juego de la confianza (o juego de la inversión), original de Berg et al. (1995). Este es un juego secuencial en el que interactúan 2 participantes; el jugador uno, que recibe un monto inicial de dinero (M) y decide cuánto de ese monto transfiere al jugador 2. El jugador 2 recibe el monto que envió el jugador 1 incrementado (típicamente, el tripe), y decide cuánto enviar de regreso al jugador 1. La predicción teórica tradicional, basada en inducción hacia atrás, sugiere que el jugador 2 se quedará con todo el dinero. Anticipando esa decisión, el jugador 1 no transfiere nada al jugador 2. Típicamente, el monto que decide enviar el jugador 1 se interpreta como confianza del jugador 1 en el jugador 2, debido a que, si el jugador 1 prefiere tener más dinero, el envío de un monto positivo significa que confía en que el jugador 2 le va a devolver un monto de dinero tal que el jugador 1 recupere su “inversión”. El monto que el jugador 2 decide enviar de regreso usualmente se interpreta como reciprocidad positiva del jugador 2 por la confianza del jugador 1.

B. Juego del ultimátum

El juego del ultimátum (Güth et al., 1982) es un juego de negociación en el que dos personas están decidiendo cómo repartir un monto de dinero entre ambos. En el juego, el jugador 1 hace una

²² Ashraf et al., (2015) muestran que darles acceso a cuentas de ahorro compartidas en el país de origen a migrantes centroamericanos incrementa los ahorros del hogar transnacional.

²³ Estos son los que Congdon et al. (2017) llaman *mechanism experiments* y a los cuales les atribuye entre otras ventajas, la de generar “mas informacion relevante para politicas publicas por dólar de investigacion que los *policy evaluations*”.

propuesta final de cómo se reparten el dinero. El jugador 2 puede aceptar o rechazar la propuesta del jugador 1. Si la acepta, se reparten el dinero como lo propuso el jugador 1; si la rechaza, ambos se quedan con 0. La predicción teórica tradicional es que el jugador 2 preferiría tener más dinero a menos, por lo que aceptaría cualquier monto. Ante esa situación, el jugador 1 preferiría enviar el mínimo monto posible. Típicamente, se observa que el jugador 1 envía montos sustanciales (entre 40%-50%), y hay altas tasas de rechazo a ofertas relativamente bajas (20%-30%). El comportamiento central de interés es el del jugador 2, el cual se interpreta como reciprocidad negativa. El comportamiento del jugador 1 no es fácilmente interpretable, pues mezcla factores estratégicos con factores pro-sociales.

C. Juego del dictador

El juego del dictador (Forsythe et al., 1994; Kahneman et al., 1986) surge para poder interpretar el resultado del jugador 1 en el juego anterior. Al igual que en el del ultimátum, en este el jugador 1 (el dictador) hace una propuesta final de cómo se reparten el dinero. La diferencia es que esa decisión es final y el jugador 2 es un receptor pasivo. El dictador puede elegir darle 0 al otro jugador y quedarse con todo el dinero; esta es la predicción teórica del modelo tradicional sin preferencias sociales. Frecuentemente se observan ofertas positivas (entre 20%-40%). La interpretación original era altruismo puro (Forsythe et al., 1994; Kahneman et al., 1986). Luego surgieron interpretaciones de otros tipos de motivaciones pro-sociales, por ejemplo, aversión a la desigualdad (Fehr & Schmidt, 1999). Trabajos más recientes cuestionan este tipo de motivaciones basados en preferencias prosociales (Bardsley, 2008; Cherry et al., 2002; List, 2007), y algunos otros (Bicchieri & Xiao, 2009; Krupka & Weber, 2013) sugieren que una mejor interpretación está dada por normas sociales: el contexto o la situación hace que se generen expectativas sobre lo que haría la mayoría de personas y lo que la mayoría piensa que es lo apropiado. Muchos apegan su comportamiento a lo que se espera que hagan.

D. Juegos de cooperación (bienes públicos por contribuciones voluntarias)

Para medir cooperación, típicamente se utilizan juegos en los que existe un dilema social: individualmente es mejor no cooperar, aunque colectivamente sea lo mejor. Para esto usualmente se utiliza el juego de bienes públicos con mecanismo de contribuciones voluntarias (Isaac & Walker, 1988; Isaac et al., 1984; Marwell & Ames, 1979)²⁴. Este es un juego de grupos, que suele hacerse en forma simultánea, en el que cada uno de los participantes recibe una cantidad de dinero inicial, y deben decidir entre asignarlo a una cuenta grupal (de la cual todos se benefician) o quedarse con el dinero. La clave del juego es que el retorno individual que obtiene cada jugador por contribuir a la cuenta grupal es menor que el de quedarse con el dinero. Sin embargo, el retorno social de contribuir a la cuenta grupal es más alto, pues todos se benefician de cada contribución. Es decir, la suma de todos los beneficios individuales se maximiza cuando todos contribuyen el máximo a la cuenta grupal. Sin embargo, cada participante maximiza sus propios beneficios cuando no contribuye a la cuenta grupal, sino que se queda con todo el dinero. Bajo los supuestos del modelo tradicional (sin preferencias sociales), se espera que nadie contribuya nada. La literatura encuentra patrones consistentes y robustos, con niveles de cooperación o contribuciones entre 40-60% del

²⁴ Algunas veces se utiliza el dilema del prisionero, el cual es una versión equivalente, con “grupos” de 2 jugadores, en donde solo pueden tomar 2 posibles acciones. Dicho de otra forma, el juego de contribuciones voluntarias es como un dilema del prisionero para grupos de más de dos personas, con un campo de acción más rico.

monto inicial en promedio.²⁵ La interpretación de los resultados es una aproximación al nivel de cooperación de los individuos que participan. Algunas variantes (Fischbacher, Gächter, & Fehr, 2001; Fischbacher & Gaechter, 2006) permiten clasificar o identificar distintos tipos de personas (cooperadores puros, cooperadores condicionales, *free riders*, etc.).

E. Juegos de coordinación (esfuerzo mínimo o del eslabón más débil)

Para medir coordinación es común utilizar el juego del esfuerzo mínimo o del eslabón más débil (Van Huyck et al., 1990, 1991). En este juego, los participantes miembros de un grupo deben elegir un nivel de “esfuerzo”. El pago que reciben los participantes está determinado por el *mínimo* esfuerzo de todos los miembros del grupo, menos el costo del esfuerzo individual.²⁶ De esta forma, nadie tiene el incentivo a hacer un esfuerzo por encima del mínimo del grupo (pues le representa un mayor costo, sin ningún beneficio); pero tampoco tienen el incentivo de hacer un esfuerzo menor al del mínimo del grupo, pues esto reduciría su propio beneficio (en una proporción mayor a la de la reducción del costo). Todos tienen el incentivo, entonces, a coordinar en el mismo nivel de esfuerzo, y todos prefieren un nivel de esfuerzo más alto. Sin embargo, al momento de elegir su nivel de esfuerzo, cada participante desconoce el nivel de esfuerzo que escogerán los otros miembros de su grupo. En este juego, cualquier nivel en el que los participantes logren coordinar para elegir el mismo esfuerzo es un equilibrio.²⁷ Todos los miembros del grupo están mejor si todos realizan el nivel máximo de esfuerzo, pero como mayor nivel de esfuerzo implica mayor costo y el pago está determinado por el nivel de esfuerzo mínimo realizado entre los miembros del grupo, elegir el nivel de esfuerzo máximo implica asumir un costo alto sin la certeza de que el pago a obtener sea el más alto posible.

F. Decisiones individuales (preferencias temporales, riesgo, etc.)

Hay varias tareas de decisiones individuales experimentales en las que un individuo toma decisiones y estas determinan sus pagos independientemente de las decisiones de terceros. Estas se utilizan comúnmente para medir características individuales o parámetros relacionados con dichas características. Por ejemplo, para medir la actitud al riesgo, existen varias tareas de elección de loterías incentivadas variando los riesgos y retornos de las distintas posibles elecciones (Binswanger, 1980; Charness et al., 2013; Eckel & Grossman, 2008; Holt & Laury, 2002).

Otras tareas experimentales están enfocadas en medir las preferencias intertemporales, es decir, relacionadas con la distintos costos y beneficios en distintos períodos de tiempo (Andersen et al., 2006, 2008; Andreoni & Sprenger, 2012a, 2012b; Burks et al., 2012; Harrison et al., 2002; Tanaka et al., 2010). Para medir estas preferencias inter-temporales hay varias tareas de elecciones incentivadas que permiten ver, por ejemplo, cuánto una persona está dispuesta a sacrificar en el presente por tener más en el futuro (paciencia o descuento del futuro). Algunas tareas, además de lo anterior, permiten determinar si existe un sesgo por el presente (ver sección C. Preferencias

²⁵ Este nivel de contribuciones es típicamente observado en juegos sin repetición o en el período inicial de juegos repetidos. Las contribuciones tienden a decaer con repetición, a menos que haya comunicación; en algunos casos, la posibilidad de sancionar o castigar permite sostener niveles altos de contribuciones.

²⁶ A mayor esfuerzo individual, mayor costo individual; y mientras mayor es el *mínimo* esfuerzo del grupo, mayores son los ingresos que reciben todos los miembros del grupo.

²⁷ Este juego tiene múltiples equilibrios; cualquier nivel de esfuerzo en el que los participantes se logren coordinar (i.e, donde todos los miembros del grupo escojan el mismo nivel de esfuerzo) es un equilibrio de Nash, aunque hay un equilibrio dominante en pagos (el del máximo nivel de esfuerzo posible).

temporales en el capítulo II Economía del Comportamiento). Este sesgo por el presente lleva a inconsistencias inter-temporales las cuales se pueden manifestar en la tendencia a posponer indefinidamente planes como ir al gimnasio o empezar a ahorrar. Finalmente, algunas tareas permiten, además de lo anterior, identificar si las personas prefieren tener flujos constantes de ingresos/consumo vs. flujos menos frecuentes, pero más grandes (preferencia por suavizar el consumo).

Otras tareas permiten identificar diferentes tipos de características individuales, como la preferencia por competir con otras personas (Niederle & Vesterlund, 2007), la motivación intrínseca a decir ser honesto, incluso cuando mentir es rentable (Fischbacher & Föllmi-Heusi, 2013), o las expectativas normativas –relacionadas con las normas sociales (Bicchieri & Xiao, 2009; Krupka & Weber, 2013).

5. Principios de diseño experimental

¿Cómo diseñar un experimento? Para empezar, debe plantearse el problema y los objetivos del experimento, poniendo al frente una pregunta de investigación claramente definida. Con base en esta pregunta de investigación se debe definir la variable de resultado (*outcome*): ¿a partir de los resultados de qué variable es posible responder la pregunta de investigación? Una vez se define la variable de interés, se procede a establecer la variable de enfoque que se va a analizar, como distintos tratamientos, así como las otras variables que se espera influyan en la variable de interés.

Antes de empezar con el diseño experimental, es importante comprender bien el entorno y la institución de interés: ¿Quiénes son los agentes que participan? ¿Cuál es la estructura de incentivos para cada uno de los agentes (valores, costos fijos, costos variables, recursos disponibles, importancia de normas, etc.)? ¿Cuáles son las reglas (formales e informales) y los procedimientos bajo los cuales los agentes interactúan actualmente? ¿Cuáles elementos parecen ser cruciales? Una descripción profunda y completa del entorno e institución actual ayudan a responder estas preguntas y permiten tener un conocimiento sistemático y organizado del problema que se quiere abordar. Además, esta descripción del entorno ayudará a determinar qué papel pueden jugar distintos hallazgos de la economía del comportamiento.

Una vez se tiene definido el problema y la pregunta de investigación, y se comprende el entorno e institución actual, inicia con la etapa del diseño experimental. Con base en la descripción del problema y objetivos, se elige un diseño experimental que responda la pregunta que nos interesa, aislando la variable de enfoque y controlando las variables estorbo –todas aquellas que puedan afectar la variable de resultado, pero que no son de nuestro interés.²⁸ En general se va a querer controlar todas las variables controlables para que los datos sean lo más informativos posible. La variable de enfoque se controla como variable de tratamiento. Idealmente, se usa niveles ampliamente separados para magnificar los contrastes y se les varía en forma independiente. Por ejemplo, supongamos que queremos responder la pregunta de si el monto de los incentivos afecta el comportamiento del jugador 2 en el juego del ultimátum. La variable de resultado sería la tasa de aceptación o de rechazo por parte de los jugadores 2, y la variable de enfoque sería el monto de los incentivos. En este caso, la variable de enfoque será la de tratamiento, la cual intentaríamos variar

²⁸ Si tuviéramos data generada en forma natural, diríamos que son todas las variables por las que es importante controlar.

en niveles ampliamente separados. Por ejemplo, Cameron (1999) utilizó tres montos a negociar ampliamente separados en el juegos del ultimátum implementado en Indonesia: Rp 5.000, Rp 40.000 y Rp 200.000.²⁹ Tras definir los tratamientos, se busca controlar las variables estorbo para aislar el efecto de la variable de enfoque sobre la variable de resultado. Las variables estorbo pueden mantenerse constantes (para reducir la complejidad), variarlas sistemáticamente como tratamiento (si se cree que puedan existir interacciones), o aleatorizarlas. Como sugerencia, especialmente cuando se estén diseñando experimentos por primera vez, se recomienda tener diseños simples. Es más valioso tener diseños simples basados en el trabajo previo de otros autores (como línea base o punto de partida) que permitan responder bien la pregunta de investigación, a diseños ingeniosos y creativos que pueden generar ruido o no permitir interpretaciones claras.

Una vez que se ha elegido un diseño experimental que responda la pregunta, se procede a definir los protocolos experimentales, los parámetros a utilizar y la unidad de observación. En forma paralela, en caso aplique, se trabaja en el modelo del entorno y la institución. Con base en la descripción del entorno, se debe también modelar los agentes con sus respectivas restricciones económicas, físicas y tecnológicas (costos, valores, tecnología, recursos, información, etc.) que determinan el sistema de incentivos de este entorno. Con base a la descripción de la institución, se diseña y/o modela el algoritmo con las reglas y procedimientos que describen cómo los agentes interactúan. Esto sirve de partida para planear los primeros experimentos piloto.

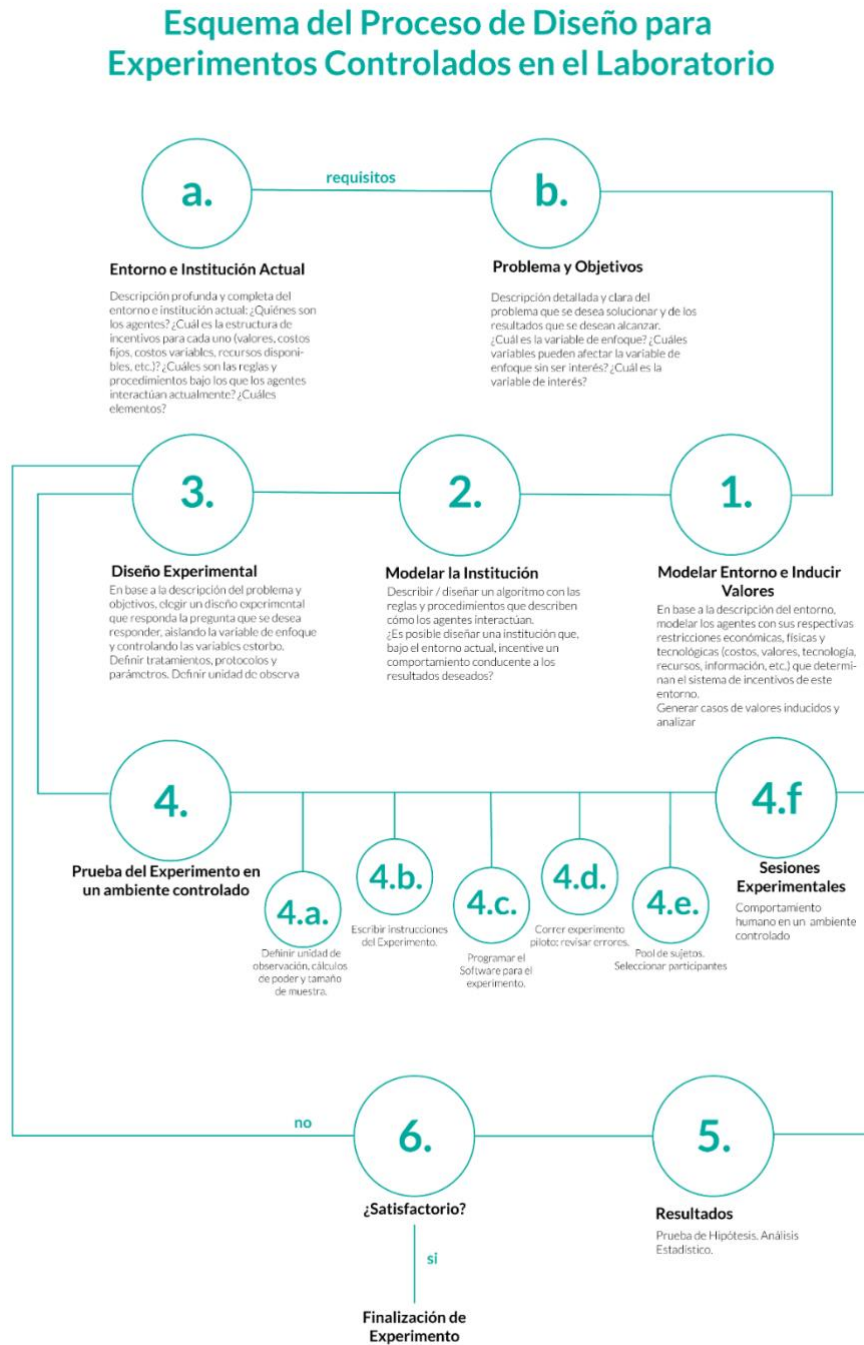
El siguiente paso es definir la unidad de observación, cálculos de poder y tamaño de muestra, así como la población objetivo. Se procede a escribir las instrucciones y, en caso de ser necesario o relevante, a programar el software para el experimento. Las instrucciones deben explicar cómo interactúan las personas y cómo las decisiones se traducen en pagos, de tal forma que los participantes comprendan claramente las consecuencias e implicaciones de sus acciones. Una vez se tenga listo el software/hoja de decisión, instrucciones, parámetros y población objetivo, se procede a llevar a cabo un experimento piloto.

El objetivo de los experimentos pilotos es revisar qué está funcionando bien y detectar problemas que se pueden corregir. Por ejemplo, revisar el correcto almacenamiento de datos, si existe confusión por parte de los sujetos, la duración de la sesión, fuentes de errores, etc. A la vez, los resultados de las pruebas piloto sirven para refinar los análisis de poder y revisar que el tamaño muestral sea adecuado.

Una vez se está satisfecho con los resultados de los pilotos, se finaliza el plan (número de sesiones, número de participantes, protocolos experimentales, etc.) y se procede a realizar las sesiones experimentales.

²⁹ Otros ejemplos que utilizan variaciones ampliamente separadas en el monto de los incentivos como variable de tratamiento son Hoffman et al. (1996), Slonim y Roth (1998), Carpenter et al. (2005), Fehr et al. (2014), y Holt y Laury (2002) para aversión al riesgo, Quizás el ejemplo documentado más extremo es un experimento de Andersen et al. (2011) en el que utilizan cuatro distintos niveles (Rs20, Rs200, Rs2,000, Rs20,000) en los experimentos en el noreste de India. Por supuesto que usar distintos niveles de la variable de tratamiento ampliamente separados no aplica solo para montos de incentivos, sino para cualquier variable de tratamiento. El objetivo general de usar niveles ampliamente separados es generar variación que permita observar efectos distintos, si estos existen.

Ilustración 2: Proceso de Diseño de Experimentos de Laboratorio (elaboración propia)



Estos principios de diseño experimental son bastante generales y están orientados en experimentos de laboratorio o *lab-in-field*, en los cuales hay interacciones entre sujetos. Sin embargo, los principios pueden aplicarse a otros tipos de experimentos. Al final, la respuesta a la pregunta de *¿cómo diseñar un buen experimento?* depende de lo que se desea investigar y para quién. Todo

buen diseño experimental parte de tener una clara definición del problema y objetivos de la investigación claros y precisos.

Terminología de diseño experimental

Entre-sujetos (between-subjects) vs. Intra-sujetos (within-subjects). Los distintos tratamientos pueden ser aplicados a distintas personas (diseño entre-sujetos), o la misma persona puede participar en más de un tratamiento (diseño intra-sujetos). Sobre-simplificando, los diseños entre-sujetos tienen la ventaja de ser más limpios, es decir cuya interpretación de los resultados es más fácil y requiere menos supuestos. Los tratamientos intra-sujetos tienen la ventaja de ofrecer más poder estadístico pues permiten controlar por todas las características individuales de los sujetos. Cuando se utiliza un tratamiento intra-sujetos, lo más usual es variar el orden de los tratamientos (y alternar el orden en distintos sujetos).

Método de respuesta directa vs. el método de la estrategia (*strategy method*). En experimentos que involucran decisiones estratégicas o interdependientes, se puede utilizar como alternativa al método de respuesta directa (la acción que elige tomar un participante), el método estratégico. Dicho método obtiene la acción condicional de un participante a cada posible(s) decisión(es) del otro(s) participantes. Esto tiene la ventaja de que nos permite observar todo el posible perfil de acciones que los participantes tomarían ante cada una de las posibles acciones que la contraparte podría tomar. La desventaja es cuando las decisiones involucran una alta carga emocional, puede que los resultados no sean los mismos en un estado “frío” que en “caliente”.

Pruebas repetidas (*repeated trials*) vs. una sola (*one shot*). La comprensión de los juegos (por muy sencillos que sean) es sumamente importante. Dado que el aprendizaje de los participantes lo obtienen por experiencia, una opción es utilizar repeticiones de los juegos. Sin embargo, frecuentemente nos interesan las decisiones que un participante tomaría en una sola interacción. Ante estas situaciones, surge un dilema entre utilizar una sola o pruebas repetidas. Si se decide utilizar pruebas repetidas, se debe decidir entre el protocolo de emparejamiento o agrupamiento de los participantes que interactuaran. Este puede variar entre interactuar siempre con la(s) misma(s) personas(s) (*fixed matching*), reasignar los grupos aleatoriamente cada período (*random rematching*), o reasignar grupos de tal forma que los participantes no interactúen más de una vez con alguien (*perfect stranger matching*).³⁰

Anonimato: sencillo (single-blind) vs. doble (double-blind). El anonimato de las decisiones entre participantes surgió desde los inicios de la economía experimental para aislar de factores externos (reputación, retaliación, etc. fuera del laboratorio) las interacciones en el laboratorio. A diferencia del anonimato sencillo (entre participantes), el protocolo de anonimato doble garantiza que las decisiones de un participante se mantienen anónimas también entre los participantes y el experimentador, de tal forma que el experimentador no puede mapear las decisiones con la identidad de los participantes. La ventaja de esto es que aísla las decisiones de efectos de percepción o reputación social ante el experimentador; la desventaja son los costos logísticos que se requieren para implementarlo.

³⁰ En este último caso, se garantiza que nunca repetirán interacciones con otra persona.

Prácticas y convenciones experimentales:

Las prácticas y convenciones experimentales en economía se diferencian de las prácticas experimentales en otras ciencias sociales, en particular en psicología. Las principales prácticas y convenciones experimentales en economía son:

Usar incentivos (monetarios). En economía, a diferencia de psicología, la convención indica que las decisiones deben tener consecuencias y no ser puramente hipotéticas. Por esta razón, se paga por desempeño con base en los resultados de las decisiones de los participantes. Esto tiene varias ventajas, entre las que sobresalen un mayor esfuerzo y atención de los participantes, y menor ruido en los datos. Su desventaja es que los experimentos se vuelvan más costosos.

No usar el engaño. Otra práctica común en economía es no utilizar el engaño. Esto se hace para mantener el control del experimento. Si los participantes en el experimento no creen en las instrucciones (por ejemplo, si les dice a los participantes que están jugando un dilema del prisionero, pero ellos creen que están jugando un juego de coordinación), o intentan ser más astutos que el experimentador, este pierde el control. Esta práctica tiene la ventaja de ayudar a mantener el control y evitar externalidades negativas a otros investigadores. Las desventajas son que dificulta o restringe el estudio de situaciones complejas o extremas, o dificulta el diseño de un experimento e incrementa los costos.

Por último, las teorías económicas son generales, no son específicas a un contexto particular. Por este motivo, desde los inicios de la metodología experimental en el laboratorio, la convención es utilizar experimentos con un marco neutral y abstracto. Esta convención se da para experimentos de laboratorio y/o experimentos de campo artificiales. Un enmarcado concreto puede ayudar a los sujetos a comprender el experimento y acercar el experimento a la pregunta de investigación. Sin embargo, puede hacer que el investigador pierda el control porque ya no sabe cómo los sujetos perciben su rol o que los sujetos utilicen heurísticas de su percepción de comportamiento fuera del laboratorio de ciertos roles/contextos. Es de notar que esta es una de las convenciones que más se suele relajar en experimentos sobre políticas públicas, pues en muchos casos las preguntas de interés son específicas al contexto.

Bibliografía

- Abbink, K. (2004). Staff rotation as an anti-corruption policy: an experimental study. *European Journal of Political Economy*, 20(4), 887–906. <https://doi.org/10.1016/J.EJPOLECO.2003.10.008>
- Abbink, K., Brandts, J., & McDaniel, T. (2003). Asymmetric Demand Information in Uniform and Discriminatory Call Auctions: An Experimental Analysis Motivated by Electricity Markets. *Journal of Regulatory Economics*, 23(2), 125–144. <https://doi.org/10.1023/A:1022202929469>
- Abbink, K., Dasgupta, U., Gangadharan, L., & Jain, T. (2014). Letting the briber go free: An experiment on mitigating harassment bribes. *Journal of Public Economics*, 111, 17–28. <https://doi.org/10.1016/J.JPUBECO.2013.12.012>
- Abbink, K., & Serra, D. (2012). Anticorruption policies: Lessons from the lab. In *Research in Experimental Economics* (Vol. 15, pp. 77–115). [https://doi.org/10.1108/S0193-2306\(2012\)0000015006](https://doi.org/10.1108/S0193-2306(2012)0000015006)
- Abbink, K., & Wu, K. (2017). Reward self-reporting to deter corruption: An experiment on mitigating collusive bribery. *Journal of Economic Behavior & Organization*, 133, 256–272. <https://doi.org/10.1016/J.JEBO.2016.09.013>
- Abraham, K., Filiz-Ozbay, E., Ozbay, E., & Turner, L. (2018a). *Behavioral Effects of Student Loan Repayment Plan Options on Borrowers' Career Decisions: Theory and Experimental Evidence* (NBER Working Paper Series No. 24804). *NBER Working Paper Series*. Cambridge, MA. <https://doi.org/10.3386/w24804>
- Abraham, K., Filiz-Ozbay, E., Ozbay, E., & Turner, L. (2018b). *Framing Effects, Earnings Expectations, and the Design of Student Loan Repayment Schemes* (NBER Working Paper Series No. 24484). *NBER Working Paper Series*. Cambridge, MA. <https://doi.org/10.3386/w24484>
- Ambler, K. (2015). Don't tell on me: Experimental evidence of asymmetric information in transnational households. *Journal of Development Economics*, 113, 52–69.
- Ambler, K., Aycinena, D., & Yang, D. (2014a). *Channeling Remittances to Education: A Field Experiment Among Migrants from El Salvador* (NBER WP) (NBER WP No. 20262). *NBER Working Paper Series*.
- Ambler, K., Aycinena, D., & Yang, D. (2014b). *Remittance Responses to Temporary Discounts: A field Experiment Among Central American Migrants*. <https://doi.org/10.3386/w20522>
- Andersen, S., Ertac, S., Gneezy, U., Hoffman, M., & List, J. A. (2011). Stakes Matter in Ultimatum Games. *American Economic Review*, 101(7), 3427–3439. <https://doi.org/10.1257/aer.101.7.3427>
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. (2008). Eliciting Risk and Time Preferences. *Econometrica*, 76(3), 583–618. <https://doi.org/10.1111/j.1468-0262.2008.00848.x>
- Andersen, S., Harrison, G. W., Lau, M. I., & Rutström, E. E. (2006). Elicitation using multiple price list formats. *Experimental Economics*, 9(4), 383–405. <https://doi.org/10.1007/s10683-006-7055-6>
- Andreoni, J., & Sprenger, C. D. (2012a). Estimating Time Preferences from Convex Budgets. *American Economic Review*, 102(7), 3333–3356. <https://doi.org/10.1257/aer.102.7.3333>
- Andreoni, J., & Sprenger, C. D. (2012b). Risk Preferences Are Not Time Preferences. *American Economic Review*, 102(7), 3357–3376. <https://doi.org/10.1257/aer.102.7.3357>
- Armantier, O., & Boly, A. (2011). A controlled field experiment on corruption. *European Economic Review*, 55(8), 1072–1082. <https://doi.org/10.1016/J.EUROECOREV.2011.04.007>
- Armantier, O., & Boly, A. (2013). Comparing Corruption in the Laboratory and in the Field in Burkina Faso and in Canada. *The Economic Journal*, 123(573), 1168–1187. <https://doi.org/10.1111/eoj.12019>
- Ashraf, N., Aycinena, D., Martínez A., C., & Yang, D. (2015). Savings in transnational households: A field experiment among migrants from El Salvador. *Review of Economics and Statistics*, 97(2), 332–351. https://doi.org/10.1162/REST_a_00462
- Attanasio, O., Barr, A., Camilo Cardenas, J., Genicot, G., & Meghir, C. (2012). Risk Pooling, Risk Preferences, and Social Networks. *American Economic Journal: Applied Economics*, 4(2). <https://doi.org/10.1257/app.4.2.134>
- Attanasio, O., Pellerano, L., & Polanía, S. (2008). Building trust? Conditional cash transfers and social capital. *The Institute For Fiscal Studies*, 30(March), 1–41.
- Aycinena, D., Blazsek, S., & Rentschler, L. (2019). Smoothing, discounting and demand for intra-household control for recipients of conditional cash transfer. *Journal of Applied Economics*, 22(1), 218–241. <https://doi.org/10.1080/15140326.2019.1596641>
- Aycinena, D., Yang, D., & Martínez A., C. (2010). *The Impact of Remittance Fees on Remittance Flows: Evidence from a Field Experiment Among Salvadoran Migrants*.
- Banerjee, A. V., Banerji, R., Berry, J., Duflo, E., Kannan, H., Mukherji, S., ... Walton, M. (2016). *Mainstreaming an Effective Intervention: Evidence from Randomized Evaluations of "Teaching at the Right Level" in India* (NBER Working Paper No. 22746). *NBER Working Paper Series*.
- Banks, J., Olson, M., Porter, D., & Rassenti, Stephen Smith, V. L. (2003). Theory, experiment and the federal communications commission spectrum auctions. *Journal of Economic Behavior & Organization*, 51(3), 303–350.
- Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *Quarterly Journal of Economics*, 116(1), 261–292. <https://doi.org/10.1162/003355301556400>
- Bardsley, N. (2008). Dictator game giving: altruism or artefact? *Experimental Economics*, 11(2), 122–133.

- <https://doi.org/10.1007/s10683-007-9172-2>
- Barr, A., Lindelow, M., & Serneels, P. (2009). Corruption in public service delivery: An experimental analysis. *Journal of Economic Behavior & Organization*, 72(1), 225–239.
- Barrera-Osorio, F., Bertrand, M., Linden, L. L., & Perez-Calle, F. (2011). Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia. *American Economic Journal: Applied Economics*, 3(2), 167–195. <https://doi.org/10.1257/app.3.2.167>
- Bejarano, H., Green, E. P., & Rassenti, S. (2017). Payment scheme self-selection in the credence goods market: An experimental study. *Journal of Economic Behavior and Organization*, 142, 396–403. <https://doi.org/10.1016/j.jebo.2017.08.004>
- Bénabou, R., & Tirole, J. (2016). Mindful economics: The production, consumption, and value of beliefs. *Journal of Economic Perspectives*, 30(3), 141–164. <https://doi.org/10.1257/jep.30.3.141>
- Benartzi, S., & Thaler, R. H. (1995). Myopic Loss Aversion and the Equity Premium Puzzle. *The Quarterly Journal of Economics*, 110(1), 73–92. <https://doi.org/10.2307/2118511>
- Benartzi, S., & Thaler, R. H. (2007). Heuristics and Biases in Retirement Savings Behavior. *Journal of Economic Perspectives*, 21(3), 81–104. <https://doi.org/10.2469/dig.v38.n1.43>
- Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, Reciprocity, and Social History. *Games and Economic Behavior*. <https://doi.org/10.1006/game.1995.1027>
- Beuermann, D. W., Cristia, J., Cueto, S., Malamud, O., & Cruz-Aguayo, Y. (2015). One Laptop per Child at Home: Short-Term Impacts from a Randomized Experiment in Peru. *American Economic Journal: Applied Economics*, 7(2), 53–80. <https://doi.org/10.1257/app.20130267>
- Bicchieri, C. (2005). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.
- Bicchieri, C. (2010). Norms, preferences, and conditional behavior. *Politics, Philosophy & Economics*, 9(3), 297–313. <https://doi.org/10.1177/1470594X10369276>
- Bicchieri, C., & Dimant, E. (2019). Nudging with Care: The Risks and Benefits of Social Information. *Public Choice*, (July). <https://doi.org/10.1007/s11127-019-00684-6>
- Bicchieri, C., & Xiao, E. (2009). Do the Right Thing: But Only if Others Do So. *Journal of Behavioral Decision Making*, 22(October 2008), 191–208. <https://doi.org/10.1002/bdm>
- Binswanger, H. P. (1980). Attitudes toward Risk: Experimental Measurement in Rural India. *American Journal of Agricultural Economics*, 62(3), 395. <https://doi.org/10.2307/1240194>
- Bolton, G., Dimant, E., & Schmidt, U. (2018). *When a Nudge Backfires: Using Observation with Social and Economic Incentives to Promote Pro-Social Behavior*. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.3294375>
- Bolton, G. E., & Ockenfels, A. (2000). ERC: A Theory of Equity, Reciprocity, and Competition. *American Economic Review*, 90(1), 166–193. <https://doi.org/10.1257/aer.90.1.166>
- Brandts, J., Pezaris-Christou, P., & Schram, A. (2008). Competition with Forward Contracts: A Laboratory Analysis Motivated By Electricity Market Design. *The Economic Journal*, 118(525), 192–214. <https://doi.org/10.1111/j.1468-0297.2007.02109.x>
- Breaban, A., & Noussair, C. N. (2018). Emotional State and Market Behavior. *Review of Finance*, 22(1), 279–309. <https://doi.org/10.1093/rof/rfx022>
- Brosig-Koch, J., Hennig-Schmidt, H., Kairies-Schwarz, N., & Wiesen, D. (2016). Using artefactual field and lab experiments to investigate how fee-for-service and capitation affect medical service provision. *Journal of Economic Behavior & Organization*, 131, 17–23. <https://doi.org/10.1016/J.JEBO.2015.04.011>
- Brosig-Koch, J., Hennig-Schmidt, H., Kairies-Schwarz, N., & Wiesen, D. (2017). The Effects of Introducing Mixed Payment Systems for Physicians – Experimental Evidence. *Health Economics*, 26(2), 243–262. <https://doi.org/10.1002/hec.3292>
- Burks, S. V., Carpenter, J. P., Götte, L., & Rustichini, A. (2012). Which measures of time preference best predict outcomes: Evidence from a large-scale field experiment. *Journal of Economic Behavior & Organization*, 84(1), 308–320. <https://doi.org/10.1016/J.JEBO.2012.03.012>
- Butler, J. V., Serra, D., & Spagnolo, G. (2017). Motivating Whistleblowers. *Ssrn*, (February). <https://doi.org/10.2139/ssrn.3086671>
- Callen, M., Isaqzadeh, M., Long, J. D., & Sprenger, C. D. (2014). Violence and Risk Preference: Experimental Evidence from Afghanistan. *American Economic Review*, 104(1), 123–148. <https://doi.org/10.1257/aer.104.1.123>
- Cameron, L. A. (1999). Raising the stakes in the ultimatum game : Experimental evidence from Indonesia. *Economic Inquiry*, 37(1), 47–59.
- Cardenas, J.-C. (2000). How Do Groups Solve Local Commons Dilemmas? Lessons from Experimental Economics in the Field. *Environment, Development and Sustainability*, 2(3/4), 305–322. <https://doi.org/10.1023/A:1011422313042>
- Cardenas, J.-C. (2004). Norms from outside and from inside: an experimental analysis on the governance of local ecosystems. *Forest Policy and Economics*, 6(3–4), 229–241. <https://doi.org/10.1016/J.FORPOL.2004.03.006>
- Carpenter, J., Verhoogen, E., & Burks, S. (2005). The effect of stakes in distribution experiments. *Economics Letters*, 86(3), 393–398. <https://doi.org/10.1016/j.econlet.2004.08.007>

- Cason, T. N. (2010). What Can Laboratory Experiments Teach Us About Emissions Permit Market Design? *Agricultural and Resource Economics Review*, 39(2), 151–161. <https://doi.org/10.1017/S1068280500007218>
- Castro, L., & Scartascini, C. (2015). Tax compliance and enforcement in the pampas evidence from a field experiment. *Journal of Economic Behavior & Organization*, 116, 65–82. <https://doi.org/10.1016/j.jebo.2015.04.002>
- Charness, G., Gneezy, U., & Imas, A. (2013). Eliciting risk preferences. *Journal of Economic Behavior and Organization*, 87, 43–51. <https://doi.org/10.1016/j.jebo.2012.12.023>
- Charness, G., Gneezy, U., & Kuhn, M. A. (2013). Experimental methods: Extra-laboratory experiments-extending the reach of experimental economics. *Journal of Economic Behavior and Organization*, 91, 93–100. <https://doi.org/10.1016/j.jebo.2013.04.002>
- Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree-An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9, 88–97. <https://doi.org/10.1016/j.jbef.2015.12.001>
- Chen, M. K., Lakshminarayanan, V., & Santos, L. R. (2006). How Basic Are Behavioral Biases? Evidence from Capuchin Monkey Trading Behavior. *Journal of Political Economy*, 114(3), 517–537. <https://doi.org/10.1086/503550>
- Cherry, T. L., Frykblom, P., & Shogren, J. F. (2002). Hardnose the Dictator. *American Economic Review*, 92(4), 1218–1221. <https://doi.org/10.1257/00028280260344740>
- Cherry, T. L., Kroll, S., & Shogren, J. F. (2008). *Environmental economics, experimental methods*. Routledge.
- Chetty, R., Looney, A., & Kroft, K. (2009). Salience and Taxation: Theory and Evidence. *American Economic Review*, 99(4), 1145–1177. <https://doi.org/10.1257/aer.99.4.1145>
- Cialdini, R. B. (2006). *Influence: Psychology of Persuasion* (Revised ed). New York: Harper Business.
- Congdon, W. J., Kling, J. R., Ludwig, J., & Mullainathan, S. (2017). Social Policy: Mechanism Experiments & Policy Evaluations. In E. Duflo & A. V. Banerjee (Eds.), *Handbook of Field Experiments*. Elsevier.
- Coricelli, G., Joffily, M., Montmarquette, C., & Villeval, M. C. (2010). Cheating, emotions, and rationality: an experiment on tax evasion. *Experimental Economics*, 13(2), 226–247. <https://doi.org/10.1007/s10683-010-9237-5>
- Cox, J. C., Friedman, D., & Gjerstad, S. (2007). A tractable model of reciprocity and fairness. *Games and Economic Behavior*, 59(1), 17–45. <https://doi.org/10.1016/j.geb.2006.05.001>
- Cox, J. C., Kreisman, D., & Dynarski, S. (2018). *Designed to Fail: Effects of the Default Option and Information Complexity on Student Loan Repayment* (NBER Working Paper Series No. 25258). *NBER Working Paper Series*. Cambridge, MA. <https://doi.org/10.3386/w25258>
- Cox, J. C., Offerman, T., Olson, M., & Schram, A. (2002). Competition For versus On the Rails: A Laboratory Experiment. *International Economic Review*, 43(3), 831–855.
- Cox, J. C., Sadiraj, V., Schnier, K. E., & Sweeney, J. F. (2016a). Higher quality and lower cost from improving hospital discharge decision making. *Journal of Economic Behavior & Organization*, 131, 1–16. <https://doi.org/10.1016/j.jebo.2015.03.017>
- Cox, J. C., Sadiraj, V., Schnier, K. E., & Sweeney, J. F. (2016b). Incentivizing Cost-Effective Reductions in Hospital Readmission Rates. *Journal of Economic Behavior & Organization*, 131, 24–35. <https://doi.org/10.1002/cncr.27633>. Percutaneous
- Cummings, R. G., Martinez-Vazquez, J., McKee, M., & Torgler, B. (2009). Tax morale affects tax compliance: Evidence from surveys and an artefactual field experiment. *Journal of Economic Behavior & Organization*, 70(3), 447–457. <https://doi.org/10.1016/j.jebo.2008.02.010>
- de Brauw, A., & Eozenou, P. (2014). Measuring risk attitudes among Mozambican farmers. *Journal of Development Economics*, 111, 61–74. <https://doi.org/10.1016/j.jdeveco.2014.08.002>
- Deck, C. A., & Wilson, B. J. (2008). Experimental gasoline markets. *Journal of Economic Behavior & Organization*, 67(1), 134–149. <https://doi.org/10.1016/j.jebo.2006.11.004>
- Duflo, E., Dupas, P., & Kremer, M. (2015). School governance, teacher incentives, and pupil–teacher ratios: Experimental evidence from Kenyan primary schools. *Journal of Public Economics*, 123, 92–110. <https://doi.org/10.1016/j.jpubeco.2014.11.008>
- Dufwenberg, M., & Kirchsteiger, G. (2004). A theory of sequential reciprocity. *Games and Economic Behavior*, 47(2), 268–298. <https://doi.org/10.1016/j.geb.2003.06.003>
- Eckel, C. C., & Grossman, P. J. (2008). Men, women and risk aversion: experimental evidence. *Handbook of Experimental Economics Results*, 1, 1061–1073.
- Epley, N., & Gilovich, T. (2016). The mechanics of motivated reasoning. *Journal of Economic Perspectives*, 30(3), 133–140. <https://doi.org/10.1257/jep.30.3.133>
- Ericson, K. M. (2017). On the Interaction of Memory and Procrastination: Implications for Reminders, Deadlines, and Empirical Estimation. *Journal of the European Economic Association*, 15(3), 692–719. <https://doi.org/10.1093/jeea/jvw015>
- Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54(2), 293–315. <https://doi.org/10.1016/j.geb.2005.03.001>
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*,

- 114(August), 817–868. <https://doi.org/10.1162/003355399556151>
- Fehr, E., Tougareva, E., & Fischbacher, U. (2014). Do high stakes and competition undermine fair behaviour? Evidence from Russia. *Journal of Economic Behavior and Organization*, *108*, 354–363. <https://doi.org/10.1016/j.jebo.2013.09.005>
- Ferraro, P. J., & Price, M. K. (2013). Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment. *Review of Economics and Statistics*, *95*(1), 64–73. https://doi.org/10.1162/REST_a_00344
- Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171–178.
- Fischbacher, U., & Föllmi-Heusi, F. (2013). Lies in Disguise – an Experimental Study on Cheating. *Journal of the European Economic Association*, *11*(3), 525–547. <https://doi.org/10.1111/jeea.12014>
- Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are People Conditionally Cooperative? Evidence from a Public Goods Experiment. *Economics Letters*, *71*(3), 397–404. <https://doi.org/10.2139/ssrn.203288>
- Fischbacher, U., & Gaechter, S. (2006). Heterogeneous Social Preferences and the Dynamics of Free Riding in Public Goods, 541–556.
- Fischhoff, B., & Beyth, R. (1975). I knew it would happen: Remembered probabilities of once—future things. *Organizational Behavior and Human Performance*, *13*(1), 1–16. [https://doi.org/10.1016/0030-5073\(75\)90002-1](https://doi.org/10.1016/0030-5073(75)90002-1)
- Forsythe, R., Horowitz, J. L., Savin, N. E., & Sefton, M. (1994). Fairness in Simple Bargaining Experiments. *Games and Economic Behavior*, *6*(3), 347–369. <https://doi.org/10.1006/GAME.1994.1021>
- Friebel, G., Kosfeld, M., & Thielmann, G. (2018). *Trust the Police? Self-Selection of Motivated Agents into the German Police Force*.
- Friedman, D., & Sunder, S. (1994). *Experimental Methods: A Primer for Economists*. Cambridge University Press.
- Friesen, L., & Gangadharan, L. (2013). Environmental Markets: What Do We Learn from the Lab? In *A Collection of Surveys on Market Experiments* (Vol. 27, pp. 135–157). Chichester, UK: John Wiley & Sons Ltd. <https://doi.org/10.1002/9781118790700.ch7>
- Gerber, A. S., & Green, D. P. (2000). The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment. *American Political Science Review*, *94*(3), 653–663. <https://doi.org/10.2307/2585837>
- Giné, X., Karlan, D., & Zinman, J. (2010). Put Your Money Where Your Butt Is: A Commitment Contract for Smoking Cessation. *American Economic Journal: Applied Economics*, *2*(4), 213–235. <https://doi.org/10.1257/app.2.4.213>
- Glewwe, P., Kremer, M., & Moulin, S. (2009). Many Children Left Behind? Textbooks and Test Scores in Kenya. *American Economic Journal: Applied Economics*, *1*(1), 112–135. <https://doi.org/10.1257/app.1.1.112>
- Glewwe, P., & Maïga, E. W. H. (2011). The impacts of school management reforms in Madagascar: do the impacts vary by teacher type? *Journal of Development Effectiveness*, *3*(4), 435–469. <https://doi.org/10.1080/19439342.2011.604729>
- Gneezy, U., & Imas, A. (2017). Lab in the Field: Measuring Preferences in the Wild. In A. V. Banerjee & E. Duflo (Eds.), *Handbook of Field Experiments* (1st ed.). Elsevier.
- Green, D. P., McGrath, M. C., & Aronow, P. M. (2013). Field Experiments and the Study of Voter Turnout. *Journal of Elections, Public Opinion and Parties*, *23*(1), 27–48. <https://doi.org/10.1080/17457289.2012.728223>
- Grether, D. M., & Plott, C. R. (1984). The Effects of Market Practices in Oligopolistic Markets: An Experimental Examination of the Ethyl Case. *Economic Inquiry*, *22*(4), 479–507. <https://doi.org/10.1111/j.1465-7295.1984.tb00700.x>
- Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization*, *3*(4), 367–388. [https://doi.org/10.1016/0167-2681\(82\)90011-7](https://doi.org/10.1016/0167-2681(82)90011-7)
- Haigh, M. S., & List, J. A. Do professional traders exhibit myopic loss aversion? an experimental analysis, 60 *Journal of Finance* § (2005). John Wiley & Sons, Ltd (10.1111). <https://doi.org/10.1111/j.1540-6261.2005.00737.x>
- Hallsworth, M., List, J. A., Metcalfe, R. D., & Vlaev, I. (2017). The behavioralist as tax collector: Using natural field experiments to enhance tax compliance. *Journal of Public Economics*, *148*, 14–31. <https://doi.org/10.1016/J.JPUBECO.2017.02.003>
- Harrison, G. W., Lau, M. I., & Williams, M. B. (2002). Estimating Individual Discount Rates in Denmark: A Field Experiment. *American Economic Review*, *92*(5), 1606–1617. <https://doi.org/10.1257/000282802762024674>
- Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature*, *42*(4), 1009–1055. <https://doi.org/10.1257/0022051043004577>
- Haucap, J., & Müller, H. C. (2014). The Effects of Gasoline Price Regulations: Experimental Evidence. *Ssrn*, (47). <https://doi.org/10.2139/ssrn.2482288>
- Hawkins, S. A., & Hastie, R. (1990). Hindsight: Biased judgments of past events after the outcomes are known. *Psychological Bulletin*, *107*(3), 311–327.
- Haynes, A. B., Weiser, T. G., Berry, W. R., Lipsitz, S. R., Breizat, A.-H. S., Dellinger, E. P., ... Gawande, A. A. (2009). A Surgical Safety Checklist to Reduce Morbidity and Mortality in a Global Population. *New England Journal of Medicine*, *360*(5), 491–499. <https://doi.org/10.1056/NEJMsa0810119>
- Healy, A. J., Malhotra, N., & Mo, C. H. (2010). Irrelevant events affect voters' evaluations of government performance.

- Proceedings of the National Academy of Sciences*, 107(29), 12804–12809.
<https://doi.org/10.1073/pnas.1007420107>
- Hennig-Schmidt, H., Selten, R., & Wiesen, D. (2011). How payment systems affect physicians' provision behaviour—An experimental investigation. *Journal of Health Economics*, 30(4), 637–646.
<https://doi.org/10.1016/J.JHEALECO.2011.05.001>
- Hill, R. V., & Viceisza, A. (2012). A field experiment on the impact of weather shocks and insurance on risky investment. *Experimental Economics*, 15(2), 341–371. <https://doi.org/10.1007/s10683-011-9303-7>
- Hirshleifer, D., & Shumway, T. (2003). Good Day Sunshine: Stock Returns and the Weather. *The Journal of Finance*, 58(3), 1009–1032. <https://doi.org/10.1111/1540-6261.00556>
- Hoffman, E., McCabe, K. A., & Smith, V. L. (1996). On Expectations and the Monetary Stakes in Ultimatum Games. *International Journal of Game Theory*, 25, 289–301.
- Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644–1655.
<https://doi.org/10.1257/000282802762024700>
- Hong, J. T., & Plott, C. R. (1982). Rate Filing Policies for Inland Water Transportation: An Experimental Approach. *The Bell Journal of Economics*, 13(1). <https://doi.org/10.2307/3003426>
- Huyck, J. B. Van, Battalio, R. C., Beil, R. O., Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1990). Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure. *American Economic Review*, 80(1), 234–248.
- Isaac, R. M., & Walker, J. M. (1988). Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism. *The Quarterly Journal of Economics*, 103(1), 179. <https://doi.org/10.2307/1882648>
- Isaac, R. M., Walker, J. M., & Thomas, S. H. (1984). Divergent evidence on free riding: An experimental examination of possible explanations. *Public Choice*, 43(2), 113–149. <https://doi.org/10.1007/BF00140829>
- Iyengar, S. S., & Lepper, M. R. (2000). When Choice is Demotivating: Can One Desire Too Much of a Good Thing?, 79(6), 995–1006. <https://doi.org/10.1037/0022-3514.79.6.995>
- Jakiela, P., Miguel, E., & te Velde, V. L. (2015). You've earned it: estimating the impact of human capital on social preferences. *Experimental Economics*, 18(3), 385–407. <https://doi.org/10.1007/s10683-014-9409-9>
- Jakiela, P., & Ozier, O. (2015). The Impact of Violence on Individual Risk Preferences: Evidence from a natural experiment. *World Bank Policy Research Paper*, (9870). <https://doi.org/10.1007/s10902-015-9633-9>
- Janssen, M. A., Bousquet, F., Cardenas, J.-C., Castillo, D., & Worrapimphong, K. (2013). Breaking the elected rules in a field experiment on forestry resources. *Ecological Economics*, 90, 132–139.
<https://doi.org/10.1016/J.ECOLECON.2013.03.012>
- Jensen, R. (2010). The (Perceived) Returns to Education and the Demand for Schooling. *Quarterly Journal of Economics*, 125(2), 515–548. <https://doi.org/10.1162/qjec.2010.125.2.515>
- Johnson, E. J., & Goldstein, D. (2003). Do defaults save lives? *Science*, 302(5649), 1338–1339.
- Kagel, J. H., Lien, Y., & Milgrom, P. R. (2010). Ascending prices and package bidding: A theoretical and experimental analysis. *American Economic Journal: Microeconomics*, 2(3), 160–185.
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness as a Constraint on Profit Seeking: Entitlements in the Market. *American Economic Review*, 76(4), 728–741. <https://doi.org/10.1017/cbo9780511803475.019>
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1990). Experimental Tests of the Endowment Effect and the Coase Theorem. *Journal of Political Economy*, 98(6), 1325–1348. <https://doi.org/10.2307/j.ctvc4j8j.7>
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The Endowment Effect, Loss Aversion, and Status Quo Bias. *Journal of Economic Perspectives*, 5(1), 193–206. <https://doi.org/10.1257/jep.5.1.193>
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263–291.
<https://doi.org/10.2307/1914185>
- Karlan, D., Morten, M., & Zinman, J. (2012). *A Personal Touch: Text Messaging for Loan Repayment* (NBER Working Paper Series No. 17952) (Vol. 5). Cambridge, MA. <https://doi.org/10.3386/w17952>
- Karlan, D. S. (2005). Using Experimental Economics to Measure Social Capital and Predict Financial Decisions. *American Economic Review*, 95(5), 1688–1699. <https://doi.org/10.1257/000282805775014407>
- Kettle, S., Hernandez, M., Ruda, S., & Sanders, M. (2016). Behavioral Interventions in Tax Compliance: Evidence from Guatemala. *World Bank Policy Research*, (June), 1–40. <https://doi.org/10.1103/PhysRevLett.77.4891>
- Kimbrough, E. O., & Vostroknutov, A. (2016). Norms Make Preferences Social. *Journal of the European Economic Association*, 14(3), 608–638. <https://doi.org/10.1111/jeea.12152>
- Kleven, H. J., Knudsen, M. B., Kreiner, C. T., Pedersen, S., & Saez, E. (2011). Unwilling or Unable to Cheat? Evidence From a Tax Audit Experiment in Denmark. *Econometrica*, 79(3), 651–692. <https://doi.org/10.3982/ECTA9113>
- Kremer, M., & Levy, D. (2008). Peer Effects and Alcohol Use among College Students. *Journal of Economic Perspectives*, 22(3), 247–257. <https://doi.org/10.1257/jep.22.3.189>
- Krupka, E. L., & Weber, R. A. (2013). Identifying social norms using coordination games: Why does dictator game sharing vary? *Journal of the European Economic Association*, 11(3), 495–524. <https://doi.org/10.1111/jeea.12006>
- Laibson, D. (1997). Golden Eggs and Hyperbolic Discounting. *Quarterly Journal of Economics*, 112(2), 443–477.

- <https://doi.org/10.1162/003355397555253>
- Larson, T., & Capra, C. M. (2009). Exploiting moral wiggle room: Illusory preference for fairness? A comment. *Judgment and Decision Making*, 4(6), 467–474.
- Levine, D. K. (1998). Modeling Altruism and Spitefulness in Experiments. *Review of Economic Dynamics*, 1(3), 593–622. <https://doi.org/10.1006/redy.1998.0023>
- List, J. A. (2007). On the Interpretation of Giving in Dictator Games. *Journal of Political Economy*, 115(3), 482–493. <https://doi.org/10.1086/519249>
- Loewenstein, G. (2003). Experimental Economics From the Vantage-point of Behavioural Economics. *The Economic Journal*, 109(453), 25–34. <https://doi.org/10.1111/1468-0297.00400>
- Loewenstein, G. (2010). Emotions in Economic Theory and Economic Behavior. *American Economic Review*, 90(2), 426–432. <https://doi.org/10.1257/aer.90.2.426>
- Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003). Projection bias in predicting future utility. *The Quarterly Journal of Economics*, 118(4), 1209–1248. <https://doi.org/10.1162/003355303322552784>
- Lord, C. G., Ross, L., & Lepper, M. R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology*, 37(11), 2098–2109. <https://doi.org/http://dx.doi.org/10.1037/0022-3514.37.11.2098>
- Madrian, B. C., & Shea, D. F. (2001). The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior. *The Quarterly Journal of Economics*, 116(4), 1149–1187. <https://doi.org/10.1162/003355301753265543>
- Marwell, G., & Ames, R. E. (1979). Experiments on the Provision of Public Goods. I. Resources, Interest, Group Size, and the Free-Rider Problem. *American Journal of Sociology*, 84(6), 1335–1360. <https://doi.org/10.1086/226937>
- Mas, A., & Moretti, E. (2009). Peers at Work. *American Economic Review*, 99(1), 112–145. <https://doi.org/10.1257/aer.99.1.112>
- McCabe, K. A., Rassenti, S. J., & Smith, V. L. (1989). Designing 'smart' computer-assisted markets: An experimental auction for gas networks. *European Journal of Political Economy*, 5(2–3), 259–283. [https://doi.org/10.1016/0176-2680\(89\)90049-9](https://doi.org/10.1016/0176-2680(89)90049-9)
- Muralidharan, K., Singh, A., & Ganimian, A. J. (2019). Disrupting Education? Experimental Evidence on Technology-Aided Instruction in India. *American Economic Review*, 109(4), 1426–1460. <https://doi.org/10.1257/aer.20171112>
- Murphy, J. J., Dinar, A., Howitt, R. E., Rassenti, S. J., & Smith, V. L. (2000). The Design of "Smart" Water Market Institutions Using Laboratory Experiments. *Environmental and Resource Economics*, 17(4), 375–394. <https://doi.org/10.1023/A:1026598014870>
- Murphy, J. J., Dinar, A., Howitt, R. E., Rassenti, S. J., Smith, V. L., & Weinberg, M. (2009). The design of water markets when instream flows have value. *Journal of Environmental Management*, 90(2), 1089–1096. <https://doi.org/10.1016/J.JENVMAN.2008.04.001>
- Niederle, M., & Vesterlund, L. (2007). Do Women Shy Away From Competition? Do Men Compete Too Much? *The Quarterly Journal of Economics*, 122(3), 1067–1101.
- O'Donoghue, T., & Rabin, M. (1999). Doing It Now or Later. *American Economic Review*, 89(1), 103–124. <https://doi.org/10.1007/s11606-008-0688-8>
- Odean, T. (1998). Are Investors Reluctant to Realize Their Losses? *The Journal of Finance*, 53(5), 1775–1798. <https://doi.org/10.1111/0022-1082.00072>
- Plott, C. R. (1987). Dimensions of parallelism: some policy applications of experimental methods. In A. E. Roth (Ed.), *Laboratory experimentation in economics* (pp. 193–219). Cambridge University Press. <https://doi.org/10.1017/cbo9780511528316.007>
- Rabin, M., & Schrag, J. L. (1999). First impressions matter: a model of confirmatory bias. *Quarterly Journal of Economics*, 114(1), 37–83. <https://doi.org/10.1162/003355399555945>
- Rassenti, S. J., Reynolds, S. S., & Smit, V. L. (1994). Cotenancy and competition in an experimental auction market for natural gas pipeline networks. *Economic Theory*, 4(1), 41–65. <https://doi.org/10.1007/BF01211118>
- Rassenti, S. J., Smith, V. L., & Wilson, B. J. (2003a). Controlling market power and price spikes in electricity networks: Demand-side bidding. *Proceedings of the National Academy of Sciences*, 100(5), 2998–3003. <https://doi.org/10.1073/pnas.0437942100>
- Rassenti, S. J., Smith, V. L., & Wilson, B. J. (2003b). Discriminatory Price Auctions in Electricity Markets: Low Volatility at the Expense of High Price Levels. *Journal of Regulatory Economics*, 23(2), 109–123. <https://doi.org/10.1023/A:1022250812631>
- Rytkin, D., Serra, D., & Tremewan, J. (2017). I paid a bribe: An experiment on information sharing and extortionary corruption. *European Economic Review*, 94, 1–22. <https://doi.org/10.1016/J.EUROECOREV.2017.02.003>
- Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1), 7–59. <https://doi.org/10.1007/BF00055564>
- Santis, O., & Chávez, C. (2015). Quota compliance in TURFs: An experimental analysis on complementarities of formal and informal enforcement with changes in abundance. *Ecological Economics*, 120, 440–450. <https://doi.org/10.1016/J.ECOLECON.2015.11.017>

- Schultz, P. W. (1998). Changing Behavior With Normative Feedback Interventions: A Field Experiment on Curbside Recycling. *Basic and Applied Social Psychology*, 21(1), 25–36. https://doi.org/10.1207/s15324834basps2101_3
- Schultz, P. W., Messina, A., Tronu, G., Limas, E. F., Gupta, R., & Estrada, M. (2016). Personalized Normative Feedback and the Moderating Role of Personal Norms. *Environment and Behavior*, 48(5), 686–710. <https://doi.org/10.1177/0013916514553835>
- Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The Constructive, Destructive, and Reconstructive Power of Social Norms. *Psychological Science*, 18(5), 429. <https://doi.org/10.1111/j.1467-9280.2007.01917.x>
- Serra, D. (2012). Combining top-down and bottom-up accountability: Evidence from a bribery experiment. *Journal of Law, Economics, and Organization*, 28(3), 569–587. <https://doi.org/10.1093/jleo/ewr010>
- Slonim, R., & Roth, A. E. (1998). Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic. *Econometrica*, 66(3), 569–596. <https://doi.org/10.2307/2998575>
- Slovic, P., & Lichtenstein, S. (1971). Comparison of Bayesian and regression approaches to the study of information processing in judgment. *Organizational Behavior and Human Performance*, 6(6), 649–744. [https://doi.org/10.1016/0030-5073\(71\)90033-X](https://doi.org/10.1016/0030-5073(71)90033-X)
- Smith, V. L. (1976). Experimental Economics: Induced Value Theory. *American Economic Review*, 66(2), 274–279. <https://doi.org/10.1017/cbo9780511528354.008>
- Smith, V. L. (1982). Microeconomic Systems as an Experimental Science. *American Economic Review*, 72(5), 923–955. <https://doi.org/10.1017/cbo9780511528354.018>
- Tanaka, T., Camerer, C. F., & Nguyen, Q. (2010). Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam. *American Economic Review*, 100(1), 557–571. <https://doi.org/10.1257/aer.100.1.557>
- Thaler, R. H. (1985). Mental Accounting and Consumer Choice. *Marketing Science*, 4(3), 199–214. <https://doi.org/10.1287/mksc.4.3.199>
- Thaler, R. H. (1990). Anomalies: Saving, Fungibility, and Mental Accounts. *Journal of Economic Perspectives*, 4(1), 193–205. <https://doi.org/10.1257/jep.4.1.193>
- Thaler, R. H. (1999). Mental Accounting Matters. *Journal of Behavioral Decision Making*, 12(3), 183–206. [https://doi.org/10.1002/\(SICI\)1099-0771\(199909\)12:3<183::AID-BDM318>3.0.CO;2-F](https://doi.org/10.1002/(SICI)1099-0771(199909)12:3<183::AID-BDM318>3.0.CO;2-F)
- Thaler, R. H. (2018). Nudge, not sludge. *Science*, 361(6401), 431. <https://doi.org/10.1126/science.aau9241>
- Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow™: Using Behavioral Economics to Increase Employee Saving. *Journal of Political Economy*, 112(S1), S164–S187. <https://doi.org/10.1086/380085>
- Thaler, R. H., & Sunstein, C. R. (2003). Libertarian Paternalism. *American Economic Review*, 93(2), 175–179. <https://doi.org/10.1257/000282803321947001>
- Thaler, R. H., & Sunstein, C. R. (2009). *Nudge: Improving decisions about health, wealth, and happiness*. Penguin.
- Thaler, R. H., Tversky, A., Kahneman, D., & Schwartz, A. (1997). The Effect of Myopia and Loss Aversion on Risk Taking: An Experimental Test. *Quarterly Journal of Economics*, 112(2), 647–661. <https://doi.org/10.1162/003355397555226>
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211(4481), 453–458. <https://doi.org/10.1126/science.7455683>
- Tversky, A., & Kahneman, D. (1991). Loss Aversion in Riskless Choice: A Reference-Dependent Model. *The Quarterly Journal of Economics*, 106(4), 1039–1061. <https://doi.org/10.2307/2937956>
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5(4), 297–323. <https://doi.org/10.1007/Bf00122574>
- Van Huyck, J. B., Battalio, R. C., & Beil, R. O. (1991). Strategic Uncertainty, Equilibrium Selection, and Coordination Failure. *The Quarterly Journal of Economics*, 106(3), 885–910.
- Velez, M. A., Murphy, J. J., & Stranlund, J. K. (2010). Centralized and decentralized management of local common pool resources in the developing world: experimental evidence from fishing communities in Colombia. *Economic Inquiry*, 48(2), 254–265. <https://doi.org/10.1111/j.1465-7295.2008.00125.x>
- Velez, M. A., Stranlund, J. K., & Murphy, J. J. (2009). What motivates common pool resource users? Experimental evidence from the field. *Journal of Economic Behavior & Organization*, 70(3), 485–497. <https://doi.org/10.1016/J.JEBO.2008.02.008>
- Viceisza, A. C. G. (2016). Creating a Lab in the Field: Economics Experiments for Policymaking. *Journal of Economic Surveys*, 30(5), 835–854. <https://doi.org/10.1111/joes.12118>
- Wilson, A. L., Bogomolova, S., & Buckley, J. D. (2015). Lack of efficacy of a salience nudge for substituting selection of lower-calorie for higher-calorie milk in the work place. *Nutrients*, 7(6), 4336–4344. <https://doi.org/10.3390/nu7064336>
- Zimmermann, F. (2020). The Dynamics of Motivated Beliefs. *American Economic Review*, 110(2), 337–361.