MINISTERIO DE INDUSTRIA,
COMERCIO Y TURISMO

CUADERNOS ECONÓMICOS
ICE
INFORMACIÓN COMERCIAL ESPAÑOLA

# CUADERNOS ECONÓMICOS

**Número 99 • 2020/I**

## ECONOMÍA DEL COMPORTAMIENTO

**Presentación**
Nuria Rodríguez-Priego

**Evaluación de políticas públicas con métodos cuasi-experimentales: una aplicación a las políticas activas de empleo en la Comunidad Valenciana**
José Vila, Yolanda Gómez, Juan Martínez y José Luis Cervera-Ferri

**Experimental Trends in Corruption and Tax Evasion**
Nuria Rodríguez-Priego and Ágnes Pintér

**Economic Models of Optimism: What Does the Evidence Say?**
Adrián Caballero and Raúl López-Pérez

**Sobornos y contratación pública: un experimento económico**
Álvaro Javier Vargas Villamizar y Luis Alejandro Palacio García

**Por qué donar y cuánto donar: determinantes de la generosidad en el juego del dictador**
Antonio M. Espín, Filippos Exadaktylos y Teresa García-Muñoz

**Reshuffling Students to Fight Discrimination and Foster Pro-Social Behavior: A Field Experiment**
Marina Bosque Mercader and Andrés Gago

**Country Differences in Ultimatum Wage Bargaining with a Real Task: Evidence from Greece, Spain and the UK**
Aurora García-Gallego, Nikolaos Georgantzís and Ainhoa Jaramillo-Gutiérrez

**Conflicto 2 × 2: una lección de teoría de juegos en el laboratorio**
Luis Alejandro Palacio García y Alexandra Cortés Aguilar

## TRIBUNA DE ECONOMÍA

**Impacto económico de la pandemia en Europa**
Emilio Ontiveros Baeza

**EDITOR**
**Javier Salinas Jiménez**
Universidad Autónoma de Madrid

**CONSEJO CIENTÍFICO**
**Michelle Boldrin,** Washington University in St Louis
**Mariam Camarero Olivas,** Universitat Jaume I
**Juan José Dolado,** U. Carlos III de Madrid
**Álvaro Escribano,** U. Carlos III de Madrid
**María Paz Espinosa,** U. del País Vasco
**Jesús Gonzalo Muñoz,** U. Carlos III de Madrid
**Andrew Jones,** University of York
**Timothy J. Kehoe,** University of Minnesota
**José María Labeaga,** UNED
**Inés Macho Stadler,** U. Autònoma de Barcelona
**Inmaculada Martínez Zarzoso,** Universitat Jaume I
**Matilde Mas Ivars,** Universitat de València
**Franco Peracchi,** Georgetown University
**Marta Rey García,** Universidade da Coruña
**Joan Enric Ricart,** IESE, U. de Navarra

**CONSEJO EDITORIAL**
**Carmen Arguedas Tomás,** U. Autónoma de Madrid
**Luis Ayala Cañón,** U. Rey Juan Carlos
**Hugo Benítez-Silva,** Stony Brook University N.Y.
**Carmen Díaz Roldán,** U. de Castilla-La Mancha
**Florentino Felgueroso,** U. de Oviedo
**Ángel López Nicolás,** U. Politécnica de Cartagena
**Ana Lozano Vivas,** U. de Málaga
**Javier Suárez Pandiello,** U. de Oviedo

**DIRECCIÓN EJECUTIVA**
**Bernardo Hernández San Juan**

**JEFA DE REDACCIÓN**
**Cristina Gómez de la Torre Curt**

**EDICIÓN Y REDACCIÓN**
Ministerio de Industria, Comercio y Turismo
Secretaría de Estado de Comercio
Paseo de la Castellana, 162. Planta 12. 28046 Madrid
Teléfonos 91 349 6053 / 91 349 3853

**DISTRIBUCIÓN**
Ministerio de Industria, Comercio y Turismo
Centro de Publicaciones
Calle Panamá, 1. Planta 0, despacho 3. 28036 Madrid
Teléfonos: 91 349 5129 / 91 349 5133 / 91 349 4335
CentroPublicaciones@mincotur.es
*www.revistasice.com*

Esta publicación se encuentra en las siguientes bases de datos bibliográficas:
*REBIUN, ECONLIT, DIALNET, LATINDEX, DICE, InDICES CSIC e ISOC*.
*Sus índices de impacto aparecen en IN-RECS*.

Los análisis, opiniones y conclusiones expuestos en los artículos de esta publicación son los de los autores y no representan opiniones oficiales de la Secretaría de Estado de Comercio, con las que no tienen por qué coincidir necesariamente.

Editor: S.G. de Estudios y Evaluación de Instrumentos de Política Comercial. Secretaría de Estado de Comercio
Diseño de cubierta: TexTec
Composición y maquetación: TexTec
Impresión y encuadernación: Centro de Impresión Digital y Diseño de la Secretaría de Estado de Comercio
Papel:
Exterior: estucado mate ecológico (250 g)
Interior: offset ecológico FSC/TCF (80 g)
ECPMINCOTUR: 1.ª ed./220/0620
PVP: 18 € + IVA
DL: M 22392-1977
NIPO: 112-19-008-4
e-NIPO: 112-19-009-X
ISSN: 0210-2633
e-ISSN: 2340-9037

**Catálogo general de publicaciones oficiales:**
**https://publicacionesoficiales.boe.es/**

# *CUADERNOS ECONÓMICOS DE ICE*

MINISTERIO DE INDUSTRIA, COMERCIO Y TURISMO

*NÚMERO 99 • 2020/I*

## *SUMARIO*

• Presentación: Economía del comportamiento ..... Nuria Rodríguez-Priego 5

*Políticas públicas*

• Evaluación de políticas públicas con métodos cuasi-experimentales: una aplicación a las políticas activas de empleo en la Comunidad Valenciana ..... José Vila, Yolanda Gómez, Juan Martínez y José Luis Cervera-Ferri 11

• Experimental Trends in Corruption and Tax Evasion ... Nuria Rodríguez-Priego and Ágnes Pintér 25

• Economic Models of Optimism: What Does the Evidence Say? ..... Adrián Caballero and Raúl López-Pérez 49

• Sobornos y contratación pública: un experimento económico ..... Álvaro Javier Vargas Villamizar y Luis Alejandro Palacio García 77

*Teoría de juegos*

• Por qué donar y cuánto donar: determinantes de la generosidad en el juego del dictador ..... Antonio M. Espín, Filippos Exadaktylos y Teresa García-Muñoz 103

• Reshuffling Students to Fight Discrimination and Foster Pro-Social Behavior: A Field Experiment ..... Marina Bosque Mercader and Andrés Gago 121

• Country Differences in Ultimatum Wage Bargaining with a Real Task: Evidence from Greece, Spain and the UK ..... Aurora García-Gallego, Nikolaos Georgantzís and Ainhoa Jaramillo-Gutiérrez 151

• Conflicto 2×2: una lección de teoría de juegos en el laboratorio ..... Luis Alejandro Palacio García y Alexandra Cortés Aguilar 171

*TRIBUNA DE ECONOMÍA*

Impacto económico de la pandemia en Europa ..... Emilio Ontiveros Baeza 191

# Presentación
# Economía del comportamiento

Nuria Rodríguez-Priego
Universidad Autónoma de Madrid

Tradicionalmente se ha asumido que el ser humano es un ser racional, que actúa de acuerdo a una determinada lógica y toma sus decisiones atendiendo a dicho raciocinio. Así, los modelos económicos tradicionales, siguiendo esta línea de pensamiento, tienen entre sus objetivos la formulación de predicciones económicas sobre la producción y el consumo a través del cálculo de los costes y beneficios marginales. Para realizar dichas predicciones, estos modelos asumen varios supuestos básicos tales como la propia racionalidad del individuo, el hecho de que los agentes económicos tengan unos gustos y preferencias determinados, el objetivo de maximizar el beneficio en el caso de las empresas y la utilidad en el caso de los consumidores, y la existencia de información completa.

Algunos economistas han cuestionado estos supuestos, comenzando por la existencia de información completa y dejando ver que, la existencia de asimetrías en la información, puede llevar a los agentes económicos a tomar decisiones sub-óptimas y redundar en fallos de mercado. Por otro lado, existe un intenso debate en torno a la racionalidad limitada de los individuos, base de la que parte la economía del comportamiento, entre otras. La atención y capacidad de cálculo de los agentes se muestran limitados en numerosas ocasiones, lo que conlleva al uso de la heurística y la intuición para satisfacer las necesidades de empresas y consumidores.

Los estudios en economía del comportamiento han estado propiciamente apoyados por la psicología, testando el efecto de las normas sociales, la reciprocidad, el altruismo, la honestidad, la aversión a las pérdidas y otros múltiples sesgos de comportamiento, descubiertos en una y otra ramas de la ciencia, que se alejan en ocasiones de la racionalidad y el individualismo asumidos inicialmente. La necesaria transversalidad de estos estudios se hace patente especialmente en sus aplicaciones políticas, ya que los retos y objetivos marcados en las agendas de los gobiernos requieren con frecuencia la identificación y análisis de elementos comportamentales con el objetivo de llevar a cabo un cambio de comportamiento de los ciudadanos o de las empresas, y dicho cambio puede tener una naturaleza muy diversa que requiera la participación de expertos en economía, psicología, estadística, educación, medio ambiente, o salud, sin ánimo de ser exhaustivos.

Parece suficientemente justificado, por tanto, dedicar este número de *Cuadernos Económicos de ICE* al estudio de las diversas aplicaciones que puede comprender la economía del comportamiento. Para ello, el monográfico se ha estructurado en dos bloques:

I) Políticas públicas
II) Teoría de juegos

El primer bloque se centra en la utilidad de los experimentos económicos para la toma de decisiones políticas. Los responsables políticos cada vez recurren más a las ciencias del comportamiento para testar diversas opciones y abordar complejos desafíos políticos. El diseño de las políticas públicas puede realizarse de forma más efectiva examinando cómo los ciudadanos toman sus decisiones. Con dicho objetivo, los modelos de comportamiento proporcionan prescripciones para una política óptima que puede tener lugar de manera no paternalista. Del mismo modo, los gobiernos pueden dirigir (*nudge*) hacia opciones más deseables desde el punto de vista político reduciendo las prohibiciones o restricciones. Así, por ejemplo, se han realizado experimentos para promover la conservación de energía y recursos, aumentar las tasas de ahorro, aumentar la productividad, mejorar el aprendizaje de los estudiantes, fortalecer las instituciones, y reducir la corrupción y el fraude. La economía del comportamiento ofrece nuevas herramientas para las decisiones políticas, pero también puede mejorar las predicciones sobre los efectos de las políticas existentes contribuyendo con nuevas formas de evaluar la efectividad de las mismas y generando nuevas implicaciones sobre el bienestar de la sociedad.

Abriendo este bloque, **José Vila**, **Yolanda Gómez**, **Juan Martínez** y **José Luis Cervera-Ferri** realizan una comparación entre dos metodologías utilizadas para evaluar el impacto de políticas públicas: el método *baseline* y el método cuasi-experimental. El primer método compara la situación de los beneficiarios antes y después de la aplicación de la política. El segundo método es una variación de la metodología de experimentos aleatorios controlados, donde la aleatorización se realiza *a posteriori* de la implementación de la política. En este último caso, los autores utilizan individuos «espejo» que cuentan con características similares a aquellos sobre los que se ha implementado la política en cuestión. Dichos sujetos «espejo» son escogidos de forma aleatoria entre todos aquellos que cumplen con las características enumeradas. Para llevar a cabo el estudio, utilizan datos sobre las Políticas Activas de Empleo (PAE) de formación, fomento de empleo e inserción laboral aplicadas en la Comunidad Valenciana durante el periodo 2012-2015, cruzando datos provenientes de la base de datos TAURÓ del Servicio Valenciano de Empleo y Formación y del Servicio Público de Empleo Estatal. Los resultados muestran que existen diferencias significativas entre ambos métodos de medición del impacto, lo que puede deberse a que el método *baseline* sobreestima dicho impacto (en algunos casos más de un 60 %), ya que no es capaz de atribuir en qué medida los cambios se deben al efecto de las Políticas Activas de Empleo o a la propia evolución del ciclo económico. Sin embargo, el método cuasi-experimental parece subsanar dicha sobreestimación, resultando una alternativa preferible a los métodos *baseline* ya que, además, no presenta las dificultades logísticas ni éticas que conllevan los métodos experimentales que supondrían privar de las PAE de forma deliberada a un grupo de sujetos que actuaría como grupo de control.

**Nuria Rodríguez-Priego y Ágnes Pintér** estudian la posible relación existente entre el nivel de fraude fiscal y la corrupción, desde el punto de vista de la economía experimental. Para ello, realizan una revisión de la literatura sobre evasión fiscal, desde la teoría económica neoclásica basada en la economía de la criminalidad y la utilidad esperada, y posteriormente revisada y extendida por diversos autores, hasta los enfoques experimentales más recientes, prestando especial atención a los experimentos tanto de laboratorio como de campo que se han llevado a cabo en el área. A este respecto, las autoras indican las bondades e inconvenientes de utilizar el laboratorio en comparación a los estudios de campo, refiriéndose principalmente al mayor nivel de control del experimento cuando se hace referencia al laboratorio, y a la mayor validez externa en el caso de los experimentos de campo. En cuanto a la teoría de la corrupción, los resultados muestran que dicho fenómeno se ha estudiado tanto desde un enfoque individual, siendo el agente el que se implica en una actividad corrupta, como a nivel agregado. Los experimentos de campo realizados en torno a la relación entre corrupción y fraude fiscal se centran principalmente en su estudio a nivel de empresas o agregado, por lo que existe una brecha en la literatura en cuanto a la relación a nivel individual. Por otra parte, existe una dificultad añadida en cuanto al estudio de los determinantes y posibles soluciones, dado el propio carácter reservado de quienes cometen actos corruptos. En este sentido, los experimentos de laboratorio resultan de gran utilidad, ya que permiten crear escenarios que imitan situaciones similares a la vida real, combinando a su vez corrupción y decisiones fiscales individuales. Este trabajo concluye haciendo hincapié en la escasa literatura existente que relacione ambos conceptos, desde el punto de vista de los agentes económicos individuales, y abriendo posibles futuras líneas de investigación al respecto.

A continuación, **Adrián Caballero** y **Raúl López-Pérez** presentan varias predicciones de los modelos económicos sobre agentes optimistas no bayesianos y las contrastan con la literatura existente. Comprender las razones y circunstancias que llevan a dichos agentes a tomar sus decisiones es una cuestión relevante desde el punto de vista de las políticas públicas. Este tipo de sesgos comportamentales pueden influir y conllevar elecciones sub-óptimas en cuanto a nuestras decisiones de consumo, inversión y ahorro, así como en las decisiones que tomamos sobre el cuidado de nuestra salud. Los modelos predicen que si los individuos tienen interés en creer que algún evento es cierto, esto es suficiente para provocar tal creencia, siempre que esta no sea demasiado inexacta. Esto contrasta con la evidencia empírica que muestra que los sujetos son, a veces, sorprendentemente precisos, incluso si tienen interés en mantener alguna creencia sesgada, y si en algunos casos mantienen creencias no Bayesianas pesimistas tales como asignar probabilidades extremadamente bajas a ciertos eventos. Los autores presentan una serie de predicciones y las contrastan con la evidencia empírica para mostrar la necesidad de profundizar en la investigación de diversos aspectos de los modelos optimistas no bayesianos. Finalmente, proponen una serie de cuestiones por resolver, así como oportunidades para futuras investigaciones al respecto.

Finalmente, el trabajo de **Álvaro J. Vargas Villamizar** y **Luis Alejandro Palacio García** investiga si el pliego de condiciones en un contrato público genera incentivos para la entrega de sobornos en la contratación pública. A través de un experimento de laboratorio, los autores analizan cómo la discrecionalidad del burócrata y las asimetrías de información afectan al nivel de entrega de sobornos y la discriminación de concursantes en una subasta para la adjudicación de contratos públicos. Para ello, utilizan cuatro tratamientos en los que los sujetos que actúan como burócratas obtienen una mayor o menor cantidad de información relevante, antes de tomar la decisión de permitir (o no) que los sujetos que actúan como concursantes entren a la subasta. Los resultados muestran que cuanto mayor era la asimetría de información a la que se sometió a los sujetos en el rol de burócratas, menor fue el incentivo por parte de los sujetos concursantes de sobornar a los primeros. Por otra parte, la discriminación llevada a cabo por los burócratas sobre la entrada de concursantes a la subasta se basó exclusivamente en la posibilidad de observar el soborno de forma directa, imperando la discriminación negativa sobre aquellos concursantes que ofrecían menores sobornos. Estos resultados muestran la necesidad de reducir el poder del burócrata sobre el pliego de condiciones, de forma que dichos requisitos sean preestablecidos para evitar la reciprocidad hacia aquellos licitantes que estén predispuestos a ofrecer sobornos.

En el segundo bloque del monográfico se aborda la aplicación de experimentos en teoría de juegos desde diversos enfoques. La teoría de juegos está íntimamente ligada a los experimentos económicos, ya que se han utilizado tradicionalmente para mejorar la comprensión de los principios generales del comportamiento estratégico. Aunque originalmente la teoría de juegos tenía como finalidad predecir el comportamiento de los individuos basándose en nociones de equilibrio, supuestos sobre preferencias de los jugadores, racionalidad, decisiones factibles e información disponible, poco a poco se ha ido ligando a la implementación empírica en múltiples contextos. Los juegos clásicos (dictador, ultimátum, dilema del prisionero, etc.) permiten su aplicación en situaciones muy diversas, lo que los hace tremendamente útiles para simular entornos reales y replicar condiciones muy próximas a aquellas que se producen en el entorno bajo investigación.

El trabajo de **Antonio M. Espín, Filippos Exadaktylos y Teresa García-Muñoz** abre este bloque. Los autores analizan a través de modelos Hurdle los determinantes de la decisión de donar, por un lado, y la de cuánto donar, por otro, en el juego clásico del dictador. Dicho juego está compuesto por dos jugadores, dictador y receptor. La decisión sobre cómo repartir una cantidad de dinero la toma el jugador en el rol de dictador, mientras que el jugador receptor no toma ningún tipo de decisión. La teoría económica predice que, según el principio de maximización del beneficio personal, el dictador se queda con el total de la cantidad, y el receptor no recibiría nada. Sin embargo, diversos experimentos económicos muestran resultados alejados de la anterior predicción. Los resultados de este trabajo muestran que, en cuanto a las variables sociodemográficas, las mujeres, los divorciados, las familias con un menor número de miembros, y aquellas con mayores ingresos tienen una mayor pro-

babilidad de donar. Respecto a las preferencias sociales, la compasión y el altruismo aumentan la probabilidad de donar, mientras que la reciprocidad negativa disminuye dicha probabilidad, a pesar de que en el juego del dictador no existen elementos que permitan una interacción recíproca posterior. Además, en relación al capital social, se observa que los individuos que consideran que el sentido de la justicia está extendido en la sociedad, son más proclives a donar. Si observamos los resultados sobre la cantidad donada, en relación a las variables sociodemográficas, los 49 años es la edad a la que se dona una mayor cantidad; los hombres donan más que las mujeres; los desempleados y trabajadores por cuenta propia donan menos; y los individuos que más contribuyen en cuanto a ingresos del hogar, donan una menor cantidad. En cuanto a las preferencias sociales y capital social, únicamente la compasión está asociada positivamente con la cantidad donada. Estos resultados podrían tener implicaciones para organizaciones de *crowdfunding* o de voluntariado.

El juego del dictador también se testa en un entorno distinto en el trabajo de **Marina Bosque Mercader y Andrés Gago**, donde los autores estudian el comportamiento pro-social entre los miembros de su grupo y la discriminación hacia los miembros de fuera del grupo. Para ello, realizan un experimento en una escuela de Barcelona con alumnos de cuarto y quinto curso de primaria para examinar si la pertenencia a un mismo grupo de clase induce favoritismo hacia miembros del grupo. Los participantes en el experimento realizan tres juegos que miden cómo estos asignan recursos a terceros y otros aspectos relacionados con el comportamiento pro-social: el juego de asignación otro-otro, el juego del dictador, y el juego de respuesta. En el centro donde se realiza el experimento, en cuarto curso los alumnos aún no han sido reasignados a un nuevo grupo, por lo que cualquier par de estudiantes únicamente responde a dos categorías: aquellos que pertenecen al mismo grupo, y aquellos que pertenecen a distintos grupos. En quinto curso, los alumnos ya han sido reasignados, por lo que las categorías se multiplican por dos creando, además de las anteriores, dos categorías nuevas: aquellos que formaban parte de distintos grupos, pero ahora han sido asignados al mismo; y aquellos que pertenecían al mismo grupo, pero ahora han sido asignados a grupos distintos. Los resultados muestran que los estudiantes estás más predispuestos a favorecer a aquellos compañeros que siempre han formado parte del mismo grupo. Además, en cuanto a las actitudes pro-sociales, se observa que están más inclinados a compartir sus recursos, y a recompensar las acciones benévolas de estos, a la vez que son más reacios a castigar los comportamientos hostiles. Por otro lado, tras la reorganización de los grupos en quinto curso, se observa que los estudiantes que comparten nuevo grupo tratan a sus compañeros como miembros del mismo, con las consecuencias positivas que ello conlleva. No obstante, los estudiantes que parecen sufrir discriminación son aquellos que no han pertenecido en ningún momento al mismo grupo. Curiosamente, los estudiantes también tienden a beneficiar a aquellos compañeros que anteriormente habían formado parte de su grupo, aunque tras la reorganización hayan sido asignados a clases diferentes. A pesar de las limitaciones mencionadas por los autores, estos resultados podrían ser útiles para examinar si la reorganización de los grupos escolares,

mezclando alumnos de distintas clases en un nuevo grupo, es una buena estrategia para expandir el comportamiento pro-social y luchar contra la discriminación a nivel escolar.

A continuación, **Aurora García-Gallego, Nikolaos Georgantzís** y **Ainhoa Jaramillo-Gutiérrez** examinan el juego del ultimátum en un entorno en el que los sujetos tienen el rol de empleadores o bien el de empleados. Aquellos participantes en el rol de empleadores tienen una cantidad inicial de dinero, y deben decidir qué parte ofrecer como salario a los jugadores que actúan de empleados, quedándose con la cantidad restante. Los sujetos en el rol de empleados decidirán si aceptan la oferta, en cuyo caso deben llevar a cabo una tarea de esfuerzo real, o si la rechazan y, como consecuencia, ambos participantes no ganan nada en el juego. El objetivo de este trabajo es examinar si existe un efecto de género en las decisiones, y si dicho efecto está determinado por las diferencias culturales. Para ello, comparan muestras de tres grupos de estudiantes universitarios reclutados en Grecia, España y Reino Unido. El principal resultado de este estudio es que las diferencias de género dependen significativamente del país de origen del grupo de sujetos. Las mujeres griegas y españolas son más proclives a rechazar una oferta en relación a los hombres, mientras que las mujeres de Reino Unido muestran la tendencia opuesta. Por otra parte, en Reino Unido, en comparación con los otros dos países mediterráneos, las ofertas por parte de los empleadores son más altas, mientras que la propensión a aceptar por parte de los empleados es más baja. Algunas de estas diferencias están basadas en patrones de género. Por un lado, las empleadoras de sexo femenino en Reino Unido son las más generosas en comparación a estudios previos sobre el juego del ultimátum. Por otro lado, los empleados de sexo masculino en este país son los más propensos a rechazar ofertas.

Cierran el segundo bloque del monográfico, **Luis Alejandro Palacio García** y **Alexandra Cortés Aguilar.** Los autores presentan un protocolo experimental como estrategia pedagógica, que permite modelar diferentes niveles de conflicto a través del uso de teoría de juegos en clase. Para ello distinguen entre el juego simultáneo y el juego secuencial, asignando a los estudiantes el rol de «paloma» en el caso de que decidan cooperar en el juego, o el rol de «halcón» si deciden no cooperar. Los autores controlan el nivel de conflicto asignando a los participantes a uno de cuatro niveles de conflicto posibles entre los que varía el equilibrio de Nash en estrategias puras. En el primer nivel de conflicto, el equilibrio es único y la estrategia «paloma» es dominante. En el segundo nivel, denominado «complementos estratégicos», la mutua cooperación es un equilibrio de Nash, pero no es único en estrategias puras. El tercer nivel se denomina «sustitutos estratégicos», donde la mejor estrategia es hacer lo contrario que el oponente, constituyendo un modelo de anticipación y disuasión. Por último, en el cuarto nivel de conflicto, la estrategia «halcón» es dominante, identificándose con el clásico dilema del prisionero. Los resultados muestran que la proporción de sujetos cooperadores se reduce por niveles de conflicto, siendo estas diferencias significativas. No obstante, no se observan diferencias en cuanto a la probabilidad de cooperar o no en función de si el juego es simultáneo o secuencial.

CUADERNOS ECONÓMICOS ICE
INFORMACIÓN COMERCIAL ESPAÑOLA

# Evaluación de políticas públicas con métodos cuasi-experimentales: una aplicación a las políticas activas de empleo en la Comunidad Valenciana

# *Evaluation of Public Policies Applying Quasi-Experimental Methods: An Application to Active Labour Market Policies in the Valencian Community*

José Vila
Universidad de Valencia, ERI-CES, IDAL y DevStat

Yolanda Gómez
José Luis Cervera-Ferri
DevStat

Juan Martínez
Servicio Valenciano de Empleo y Formación

***Resumen***

*Este artículo presenta un estudio comparativo de dos metodologías de medición de impacto de políticas públicas: los métodos baseline y cuasi-experimental. Ambos métodos son aplicados a la evaluación de las políticas activas de empleo implementadas en la Comunidad Valenciana en el periodo 2012-2015. La aplicación de ambos métodos a un mismo conjunto de políticas permite cuantificar un fenómeno (habitualmente citado en la literatura, pero escasamente validado empíricamente) como es el grado de confusión entre los efectos de una política y los de la evolución de su entorno, asociados a la utilización del método baseline. Los resultados de este análisis muestran que, en algunos casos, más de dos terceras partes del impacto medido a través de indicadores baseline es en realidad atribuible a variaciones del entorno y no a la política evaluada. Como consecuencia, los resultados de este artículo muestran los peligros de afrontar la evaluación de políticas desde un enfoque baseline y la conveniencia de recurrir a métodos alternativos con base experimental.*

***Palabras clave:*** *evaluación de políticas, métodos experimentales, cuasi-experimentos, políticas activas de empleo.*

***Clasificación JEL:*** *C21, C93, J48.*


DOI: https://doi.org/10.32796/cice.2020.99.7006
Recibido: Febrero 2020 · Aceptado: Abril 2020

***Abstract***

*This article presents a comparative study of two methodologies for measuring the impact of public policies: the baseline and quasi-experimental methods. Both methods are applied to the evaluation of Active Labour Market Policies implemented in the Valencian Community during the period 2012-2015. The application of both methods to the same set of policies allows us to quantify a phenomenon (that is usually cited in the literature, but scarcely empirically validated) such as the degree of confusion between the effects of a policy and those of the evolution of its environment associated with the use of the baseline method. The results of this analysis show that, in some cases, more than two thirds of the impact measured through baseline indicators is actually attributable to variations in the environment and not to the evaluated policy. As a consequence, the results of this article show the dangers of approaching policy evaluation from a baseline perspective and the convenience of resorting to alternative methods on an experimental basis.*

***Keywords:*** *policy evaluation, experimental methods, quasi-experiments, active labour market policies.*

***JEL Classification:*** *C21, C93, J48.*

## 1. Introducción

Este artículo presenta un estudio comparativo de dos metodologías de medición de impacto de políticas públicas: *baseline* y experimental. El primer método compara la situación de los beneficiarios de la política antes y después de su aplicación. El segundo método es una variación de la metodología de experimentos aleatorios controlados (EAC), donde la aleatorización se realiza a posteriori de la implementación de la política a evaluar. En este trabajo, ambos métodos son descritos en detalle y aplicados al caso de estudio de la medición del impacto de las políticas activas de empleo (PAE) aplicadas en la Comunidad Valenciana en el periodo 2012-2015. Las PAE vienen definidas por el Real Decreto-ley 3/2011, de 18 de febrero, de medidas urgentes para la mejora de la empleabilidad y la reforma de las políticas activas de empleo (BOE núm. 43, de 19 de febrero de 2011) como:

> El conjunto de acciones y medidas de orientación, empleo y formación dirigidas a mejorar las posibilidades de acceso al empleo, por cuenta ajena o propia, de las personas desempleadas, al mantenimiento del empleo y a la promoción profesional de las personas ocupadas y al fomento del espíritu empresarial y de la economía social (art. 10).

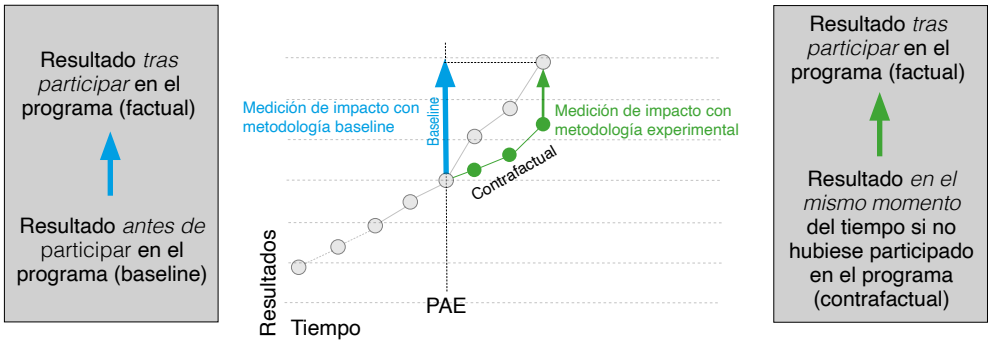
Las PAE juegan así un papel crucial por su contribución tanto a la creación de empleo como a la inserción de los desempleados. Aunque existen abundantes ejemplos de medición de impacto con métodos *baseline* o cuasi-experimentales (White y Sabarwal, 2014; Mueller y Gaus, 2017), la novedad de este artículo es la aplicación de ambos métodos a un mismo conjunto de políticas, lo que permite contrastar empíricamente su efectividad a la hora de evaluar políticas públicas.

Aunque no exentos de crítica, los EAC han sido definidos como el «patrón oro» de la medición de impacto (Bédécarrats *et al*., 2018; Grossman y Mackenzie, 2005), ya que son capaces de cuantificar el impacto realmente atribuible a una política.

Sin embargo, el método *baseline* puede confundir el impacto de la política y el de la evolución del entorno en el que esta se aplica (Athey y Imbes, 2017). A pesar de su creciente utilización en la fase de diseño de políticas públicas, la aplicación de EAC presenta dificultades logísticas y éticas para su evaluación. Esto se debe a que los EAC requieren una aleatorización previa entre los potenciales beneficiarios para determinar sobre quiénes de ellos se van a aplicar. Aunque esta aleatorización puede hacerse fácilmente a pequeña escala durante la fase de diseño, resulta complicada de realizar a mayor escala para evaluar la política una vez implementada. En este marco, surge la necesidad de plantear variaciones de los EAC que, manteniendo al máximo la capacidad de medir el impacto atribuible a una política, sean más sencillos de implementar en la práctica (White y Sabarwal, 2014). Ese es el caso del método cuasi-experimental analizado en este trabajo.

Para poder atribuir un impacto a una PAE, es necesario establecer una relación causa-efecto entre las acciones asociadas a esta política y sus indicadores clave de desempeño (ICD) (Lowndes *et al*., 2017). Sin embargo, la comparación de la situación de los beneficiarios antes de su participación en acciones de una PAE (medición *baseline*) con su situación después de dicha participación (medición factual) no es capaz de aislar el impacto de la PAE del de otras variables y establecer causalidad. Por ejemplo, en un entorno recesivo, un fuerte crecimiento de la proporción de desempleados que siguen sin encontrar trabajo tras su participación en una PAE no indica que esta tenga un efecto negativo: el crecimiento del desempleo podría haber sido aún mayor si no se hubiese aplicado la política para paliar la destrucción de empleo generada por la recesión. Como se muestra en la Figura 1, para establecer atribución es necesario comparar el nivel de empleo de los beneficiarios de la PAE después de participar en una acción (medición factual) con el que hubiesen tenido si no hubiesen participado en la acción asociada a la PAE (medición contrafactual). Sin embargo, es imposible efectuar ambas mediciones ya que un desempleado participa o no participa en la acción, pero no puede hacer ambas cosas a la vez.

**FIGURA 1**
**ENFOQUES BASELINE Y EXPERIMENTAL**

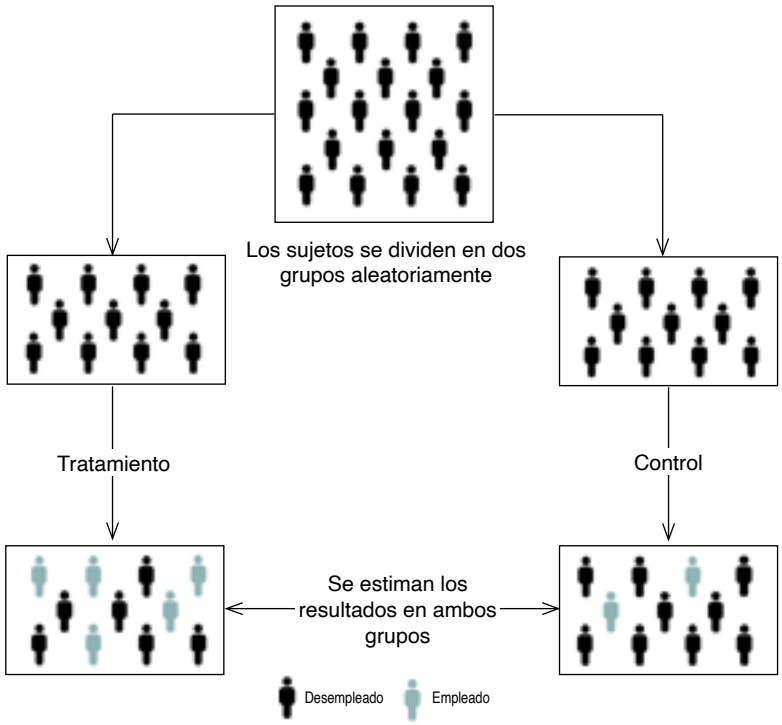



**FUENTE:** Elaboración propia.

Los EAC proporcionan una metodología para estimar la diferencia entre lo que ha pasado y lo que hubiese pasado sin aplicar una política. El punto clave de un EAC (McDermott, 2002) es la asignación aleatoria de los potenciales beneficiarios del programa a uno de dos grupos, control y tratamiento, con antelación a la implementación de las acciones de la PAE. Así, los miembros del grupo control no participan en acciones de la PAE, mientras que los del grupo tratamiento sí lo hacen.

El supuesto fundamental en la aplicación de EAC es que esta aleatorización *a priori* garantiza la comparabilidad entre los sujetos de los diferentes grupos, ya que estos tendrán características similares y habrá sufrido de igual manera los efectos de la evolución del entorno. En un EAC, el impacto atribuible a la PAE se puede estimar como la diferencia entre la medición de empleabilidad factual (proporcionada por el grupo tratamiento) y contrafactual (proporcionada por el grupo control) (Figura 2).

**FIGURA 2**
**ESQUEMA CONCEPTUAL DE LOS MÉTODOS EXPERIMENTALES**




**FUENTE:** Elaboración propia.

La implementación de EAC presenta dificultades, pues supone negar el acceso al programa a ciertos beneficiarios de forma aleatoria (grupo control). Como alternativa, el método cuasi-experimental plantea otra forma de definir los grupos control y tratamiento basada en aleatorizaciones que pueden realizarse después de la implementación de la PAE. En este caso, el grupo tratamiento se establece a partir una muestra aleatoria de beneficiarios que hayan participado en el programa y, para cada uno de ellos, se selecciona aleatoriamente un «espejo» con un perfil semejante, pero que no haya participado en la PAE.

Las PAE son políticas de mercado de trabajo dirigidas a mejorar los niveles de competitividad del sistema productivo (Vooren *et al*., 2019) y se caracterizan por ser anticipadoras y preventivas a la situación de desempleo. Así, las PAE están dirigidas a contrarrestar los efectos negativos que sobre el empleo provocan los cambios en las condiciones económicas y del entorno. Por esta razón, la discriminación entre el efecto de una PAE y el de la evolución del entorno económico es crítica para analizar su efectividad (Caliendo y Schmidl, 2016), convirtiendo a las PAE en candidatas ideales para la comparación de los métodos de medición de impacto *baseline* y cuasi-experimental efectuada en este trabajo.

### *Políticas activas de empleo*

Las PAE están constituidas por tres grupos principales de políticas: formación, fomento de empleo e inserción laboral, cada una de las cuales se compone, a su vez, de una serie de acciones concretas. La definición de cada tipo de PAE se muestra en el Cuadro 1.

**CUADRO 1**
**DEFINICIÓN DE LOS DIFERENTES TIPOS DE PAE**

| **Políticas activas de empleo (PAE)** |
|---|
| **Formación** |
| Políticas de formación, destinadas a fomentar la adquisición, mejora o puesta al día de competencias y cualificaciones profesionales. |
| 1. Acciones de formación para desempleados (o formación profesional para el Empleo, prioritariamente para trabajadores desempleados): están dirigidas prioritariamente a personas en situación de desempleo, a fin de proporcionarles una cualificación profesional para desempeñar una ocupación adecuada a las necesidades del mercado de trabajo |
| 2. Acciones de formación continua: son acciones de formación sectoriales y/o intersectoriales dirigidos tanto a la adquisición de competencias transversales como de competencias específicas dirigidas a la formación de personas trabajadoras. |
| 3. Acciones de formación dual: es el conjunto de acciones e iniciativas formativas, mixtas de empleo y formación, que tienen por objeto la cualificación profesional de los trabajadores en un régimen de alternancia de actividad laboral en una empresa con la actividad formativa recibida en el marco del sistema de formación profesional para el empleo o del sistema educativo. |

**CUADRO 1 *(Continuación)***
**DEFINICIÓN DE LOS DIFERENTES TIPOS DE PAE**

**Políticas activas de empleo (PAE)**

**Fomento de empleo**

Políticas de promoción y creación de empleo.

4. Acciones de empleo público y desarrollo local: a través de este tipo de acciones se promueve la generación de empleo para la realización de obras o servicios de interés general y social, apoyando a las corporaciones locales y entidades que desarrollan este tipo de proyectos en el ámbito de la Comunitat Valenciana.
5. Acciones de empleo estable: conjunto de acciones de formación destinadas a aumentar el empleo y su estabilidad mediante medidas de fomento de empleo dirigidas tanto a la contratación inicial, como a la transformación de contratos temporales en indefinidos, favoreciendo la inserción laboral de trabajadores pertenecientes a los colectivos de desempleados con mayores dificultades de acceso al mercado de trabajo.
6. Acciones para discapacitados: las acciones de fomento del empleo para personas con discapacidad en Centros Especiales de Empleo tienen como finalidad mejorar el nivel de inserción laboral de este colectivo, rompiendo las barreras que dificultan su incorporación al mercado de trabajo, y potenciando la creación de empleo estable adecuado.

**Inserción laboral**

Políticas de orientación, intermediación y colocación que ofrecen información y orientación laboral, al tiempo que se gestionan las colocaciones y posibles contrataciones a partir de los requerimientos de los empleados que precisan las empresas y de sus capacidades para optar a ellos.

7. Acciones de inserción laboral: formada por acciones cuyo objetivo la integración socio laboral de trabajadores en situación o riesgo de exclusión social en empresas de inserción y en empresas de trabajo.

**FUENTE:** Elaboración propia a partir de información de SERVEF.

## 2. Metodología

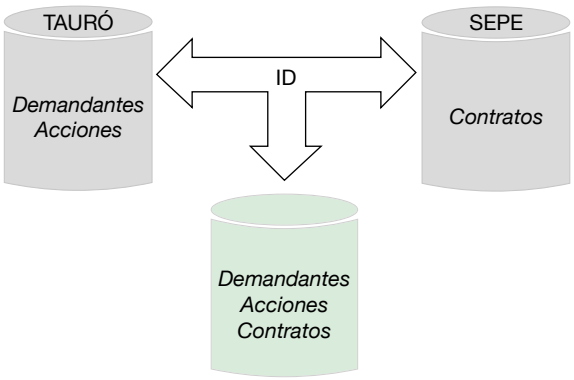
La comparación de métodos de medición de impacto presentada en este artículo cubre las PAE de formación, fomento de empleo e inserción laboral aplicadas en la Comunidad Valenciana en el periodo 2012-2015.

### *2.1. Fuentes de información utilizadas*

La información utilizada para la medición de los indicadores se ha obtenido a partir de la integración de bases de datos TAURÓ del Servicio Valenciano de Empleo y Formación (SERVEF) y por el Servicio Público de Empleo Estatal (SEPE). La base de datos del SEPE contiene información sobre los contratos de trabajo firmados durante el periodo de análisis (2012-2015). En concreto, la base del SEPE contiene casi seis millones de registros que recogen la siguiente información: identificación del trabajador; información sobre la residencia; fecha de inicio y final del contrato; características del contrato (ocupación, tipo de contrato, tipo de jornada y horas de jornada). Por otro lado, la base de datos de SERVEF contiene información sobre los demandantes de empleo de la Comunidad Valenciana, incluyendo datos de

demandantes (más de dos millones de registros con información socio-demográfica de los demandantes de empleo) y de acciones implementadas en el marco de las PAE (tres millones de registros, el código de la acción, fecha de realización, participantes, etc.). Ambas bases de datos contienen información esencial para el cálculo de los indicadores: mientras la de SERVEF nos proporciona información sobre las acciones de formación que han implementado los demandantes de empleo y sus perfiles socio-demográficos, la de SEPE contiene información sobre los contratos de trabajo y su tipología. Para proceder al análisis, ambas bases se integraron por medio de un identificador único anonimizado por usuario. Este indicador permitió cruzar información de las dos bases de datos respetando la confidencialidad de los demandantes y la legislación en protección de datos (Figura 3).

**FIGURA 3**
**RELACIÓN ENTRE LAS BASES DE DATOS**




**FUENTE:** Elaboración propia.

## *2.2. Muestra*

Los ICD has sido calculados a partir de una muestra de 44.236 participantes en las acciones de las diferentes PAE. Se ha utilizado un diseño estratificado no proporcional con 3 estratos definidos por los 3 tipos de acciones evaluadas (Cuadro 2).

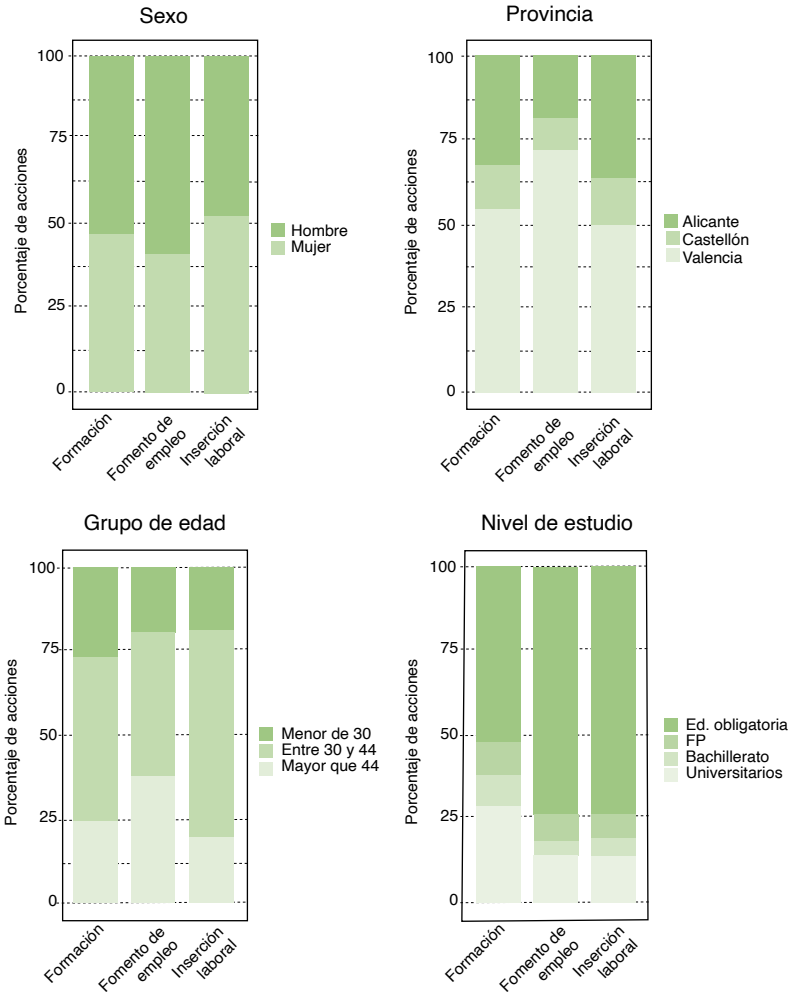
Atendiendo a las características socio-demográficas de los participantes (Figura 4), se observa que la participación de hombres y mujeres en las PAE está equilibrada. Los menores de 30 años participan más en las acciones de formación, los participantes entre 30 y 44 años participan en mayor medida en las acciones de inserción laboral y los mayores de 44 años participan más en las acciones de fomento del empleo. Los participantes que tienen educación obligatoria presentan un mayor porcentaje en las acciones de formación dual. Por el contrario, este grupo participa en menor medida en las acciones de formación para desempleados.

## CUADRO 2
## DISEÑO MUESTRAL

| | Población | Muestra |
|---|---|---|
| Formación | 145.372 | 20.715 |
| Fomento de empleo | 48.144 | 15.467 |
| Inserción laboral | 49.885 | 8.054 |
| **Total** | **243.401** | **44.236** |

**FUENTE:** Elaboración propia a partir de las bases de datos de SERVEF y de SEPE.

## FIGURA 4
## CARACTERÍSTICAS SOCIODEMOGRÁFICAS DE LOS PARTICIPANTES POR TIPO DE ACCIÓN




**FUENTE:** Elaboración propia a partir de las bases de datos de SERVEF y de SEPE.

### *2.3. Cálculo de los ICD: métodos baseline y cuasi-experimental*

La definición de ICD es el paso inicial en la evaluación. La elección de indicadores debe responder a los criterios conocidos como SMART:

- Específicos (S de *Specific*): los indicadores deben poder ser traducidos en términos operativos. Debe quedar claro qué se está midiendo.
- Medibles (M de *Measurable*): los indicadores deben poder ser medidos en términos estadísticos, mediante la observación directa o la estimación siguiendo procedimientos aceptados científicamente. En este caso, deben estar basados en las fuentes de información disponibles que se describen posteriormente.
- Alcanzables (A de *Attainable* y de *Attributable*): los indicadores deben poder ser comparados con un objetivo a alcanzar. Al mismo tiempo, los cambios en los indicadores deben poder atribuirse a las intervenciones realizadas y no a factores ajenos. Este criterio hace necesario un enfoque experimental, que se describe más abajo.
- Realista (R de *Relevant*): los indicadores deben ser una medida válida del impacto de las PAE, es decir, que existe una relación causal entre las PAE y los valores alcanzados por los indicadores.
- En Tiempo (T de *Time*): los indicadores deben poder reflejar cambios en el tiempo, y ser obtenidos con un desfase temporal que los mantenga válidos.

Frente a la elección de un sistema óptimo de indicadores, hay que considerar aspectos pragmáticos de factibilidad y coste. Así, es necesario analizar las fuentes de información existentes antes de decidir una operación de recogida específica de información que sería muy costosa.

Para el presente estudio, se ha definido un sistema de ICD que recogen dos dimensiones diferentes de impacto de una PAE: su capacidad de incorporar al demandante al mercado laboral (*empleabilidad*) y de incrementar el número de días trabajados por este (*duración*). Estos grupos de indicadores son:

- Indicadores de *empleabilidad* que cuantifican el impacto de una acción a través del incremento (o decremento) de la probabilidad de contratación de los participantes en dicha acción (probabilidad que se estimará mediante una proporción).
- Indicadores de *duración* que cuantifican el impacto de una acción a través del incremento (o decremento) del número de días trabajados por los participantes en dicha acción calculado en forma de porcentaje para los diferentes plazos temporales.

Se trata, por tanto, de indicadores que miden diferencias, no niveles. Estos indicadores se han calculado para tres plazos temporales: corto plazo (6 meses), medio plazo (12 meses), y largo plazo (18 meses).

Los ICD se han calculado de dos formas distintas. En primer lugar, se ha aplicado un enfoque *baseline*, comparando la situación laboral del participante (empleado o desempleado) antes y después de su participación en la PAE. Los ICD de empleabilidad *baseline* se han calculado pues como la diferencia entre el porcentaje de contratados en la muestra antes y después de realizar una acción. Análogamente, los ICD de duración *baseline* se han obtenido como la diferencia entre el porcentaje de los días trabajados antes y después de realizar la acción.

La aplicación del método cuasi-experimental requiere identificar para cada uno de 44.236 participantes de la muestra a otra persona que no haya participado en dicha acción y que tenga un perfil comparable, a la que denominaremos «espejo». El espejo de cada demandante en la muestra ha sido seleccionado aleatoriamente entre todos los sujetos de la base de datos integrada que no han realizado ninguna acción y tienen un perfil similar al del demandante original en función de todas las variables disponibles. Así, cada espejo tiene el mismo sexo, provincia de residencia, grupo de edad, nivel de estudios y una diferencia menor de 30 días de experiencia laboral que el demandante de la muestra al que refleja.

Una vez identificados los espejos, los ICD de empleabilidad cuasi-experimentales se han calculado como la diferencia entre el porcentaje de participantes de la muestra contratados tras ser beneficiario de la PAE y el porcentaje de espejos contratados en ese mismo momento. Del mismo modo, los ICD de duración cuasi-experimentales se calculan como la diferencia en los porcentajes de días trabajados por los participantes de la muestra y sus espejos.

## 3. Resultados

La Figura 5 y el Cuadro 3 comparan los resultados de la estimación de los ICD con ambos métodos para los tres tipos de PAE y los tres horizontes temporales considerados. Hay que destacar que, dado los elevados tamaños muestrales en los tres grupos, la potencia estadística de los t-test efectuados para comparar la hipótesis nula de igualad de indicadores obtenidos con ambos métodos de cálculo es muy elevada. Como resultado, los t-test efectuados proporcionan evidencia empírica suficiente para rechazar que no existen diferencias entre los valores de los ICD *baseline* y cuasi-experimentales (el *p*-valor es inferior a 0.00 para los indicadores de empleabilidad y duración de los tres tipos de PAE y en los tres horizontes temporales).

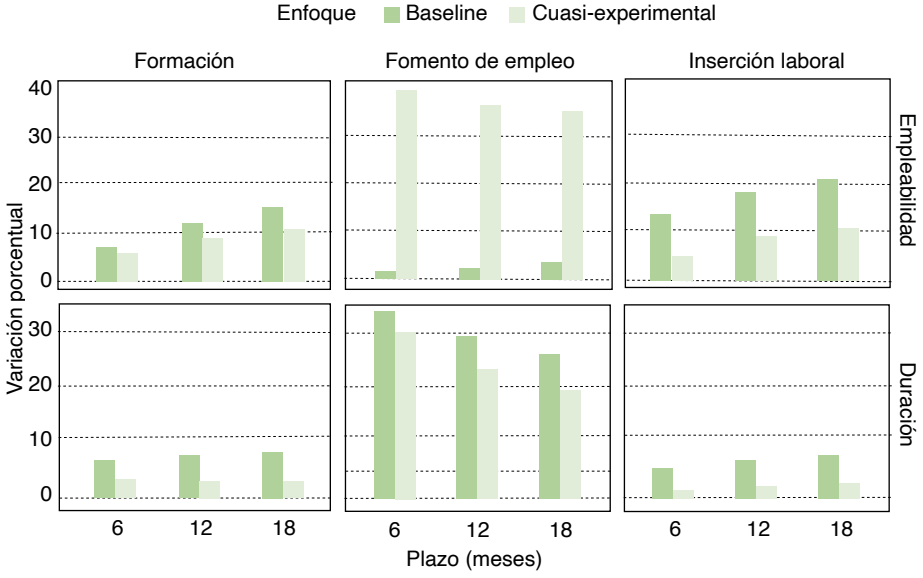
Con la única excepción de los ICD de empleabilidad para las PAE de fomento de empleo, la comparación de los ICD *baseline* y cuasi-experimentales exhiben un mismo patrón. Si bien ambos métodos de medición establecen que las PAE han tenido un impacto positivo tanto sobre empleabilidad y como duración del empleo, la estimación *baseline* es siempre superior a la cuasi-experimental. Esta diferencia sistemática puede explicarse por el hecho de que las estimaciones *baseline* confunden el efecto de las PAE con el de la evolución del ciclo económico. Dado que las políticas se han implementado en una fase expansiva del ciclo (2015-2017), la

**CUADRO 3**
**ICD DE EMPLEABILIDAD Y DURACIÓN ESTIMADOS CON LOS MÉTODOS BASELINE Y CUASI-EXPERIMENTAL**

| Indicador y grupo | Baseline | | | Cuasi-experimental | | |
|---|---|---|---|---|---|---|
| | 6 meses | 12 meses | 18 meses | 6 meses | 12 meses | 18 meses |
| Empleabilidad para formación | 6,9 | 12,3 | 14,8 | 5,6 | 9,0 | 10,5 |
| Empleabilidad para fomento de empleo | 1,1 | 2,3 | 3,2 | 38,6 | 35,6 | 34,7 |
| Empleabilidad para inserción laboral | 13,2 | 17,4 | 20,2 | 4,9 | 8,9 | 10,9 |
| Duración para formación | 6,3 | 7,7 | 8,2 | 3,1 | 3,2 | 3,0 |
| Duración para fomento de empleo | 33,6 | 28,6 | 25,9 | 29,5 | 22,8 | 19,0 |
| Duración para inserción laboral | 5,4 | 6,8 | 7,5 | 1,1 | 2,2 | 2,5 |

**FUENTE:** Elaboración propia.

**FIGURA 5**
**COMPARACIÓN DE LOS ICD DE EMPLEABILIDAD Y DURACIÓN ESTIMADOS CON LOS MÉTODOS BASELINE Y CUASI-EXPERIMENTAL**



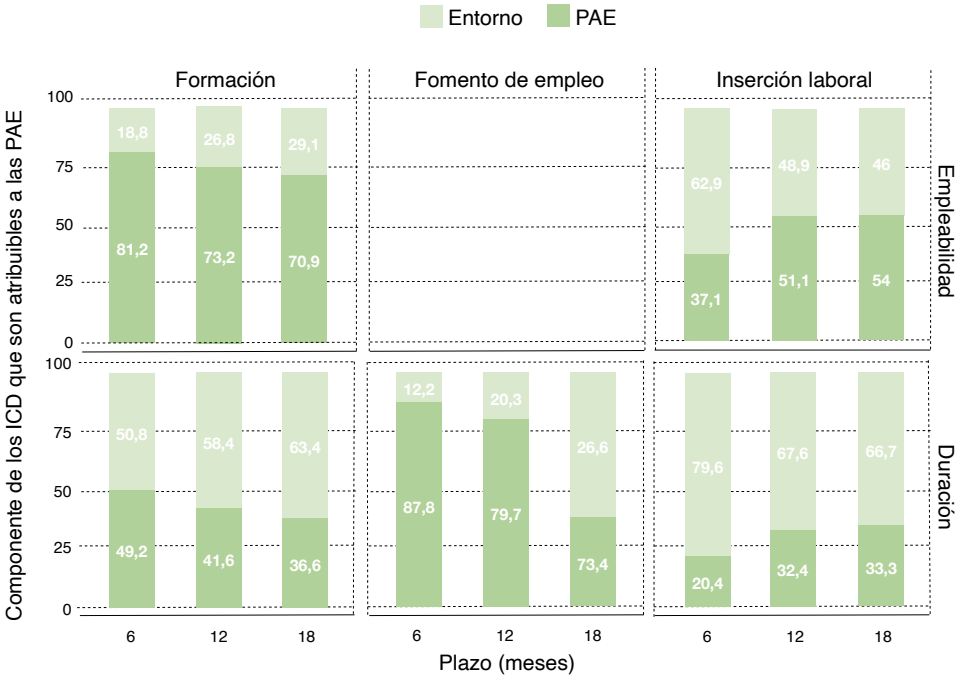

**FUENTE:** Elaboración propia

medición *baseline* cuantifica tanto el impacto positivo de la PAE en el incremento de empleabilidad y duración como el que se produce como consecuencia de la mejora de la situación económica (y que se hubiese producido de todas formas, aunque las PAE no hubiesen sido implementadas). Sin embargo, las estimaciones

cuasi-experimentales aíslan el impacto que puede ser atribuido estrictamente a la implementación de la política. La Figura 6 presenta cómo el impacto total de política y entorno (recogido por el valor de los indicadores *baseline*) se distribuye entre el efecto que puede ser atribuido a las PAE y a la evolución de la economía, respectivamente. Por ejemplo, para las PAE de formación, un 18,8 %, 26,8 % y un 29,1 % del valor estimado para los ICD *baseline* de empleabilidad a 6, 12 y 18 meses no se puede atribuir a la política sino al entorno. Los errores de atribución son incluso mayores en otros casos, suponiendo en algunos casos más de dos terceras partes del impacto estimado con los ICD *baseline* (indicadores de duración para inserción laboral).

Conviene señalar que los ICD de empleabilidad para PAE de fomento de empleo son los únicos que muestran un patrón diferente. En este caso, la estimación por el método cuasi-experimental proporciona valores mucho más elevados que los de la estimación *baseline* para los tres horizontes temporales considerados. Este resultado se debe a una particularidad de este tipo de PAE, que suelen incluir la firma de un contrato o la mejora de las condiciones de contrato ya existente en el momento concreto en que comienza la acción.

**FIGURA 6**
**COMPONENTE DE LOS ICD QUE SON ATRIBUIBLES A LAS PAE**
**(En %)**




**FUENTE:** Elaboración propia.

## 4. Conclusiones

Este artículo presenta dos evaluaciones alternativas de las políticas activas de empleo (PAE) implementadas en la Comunidad Valenciana durante el periodo 2015-2017, aplicando un enfoque *baseline* y cuasi-experimental respectivamente. Dicho análisis ofrece evidencia empírica de cómo la estimación de ICD con el método *baseline* puede llegar a proporcionar información imprecisa, la cual puede dificultar la evaluación y la optimización de políticas públicas. Este método de cálculo, basado en la comparación de la situación de los beneficiarios de la política antes y después de participar en ella, confunde el resultado atribuible a dicha política con los efectos de la variación del entorno. Así, los ICD de las PAE evaluadas en un periodo expansivo (como sucede en el caso de estudio presentado) aparecen sobrestimados si se aplica la metodología *baseline*. Esta sobrestimación es en algunos casos crítica (Figura 6), siendo, en algunos, atribuible al entorno y no a la implementación de la política más de un 60 % del valor de los ICD *baseline*.

Sin embargo, la contribución de este artículo no es señalar una debilidad del método *baseline*, ampliamente discutida en la literatura, ni de los peligros que su utilización puede conllevar en la evaluación y diseño de políticas públicas. Más allá de esto, el artículo propone un método de evaluación de fácil implementación que no requiere una privación aleatoria de los efectos de la política a una parte de la población (como es el método cuasi-experimental descrito en la sección 2) y permite obtener mediciones de los ICD más fiables y con capacidad de discriminar el nivel de impacto que puede ser efectivamente atribuible a la política misma.

En términos generales, las PAE analizadas han tenido un impacto positivo, tanto en empleabilidad, como en duración. Descontado el efecto del contexto económico (enfoque pseudo-experimental) el impacto sobre ambos tipos de indicadores sigue siendo positivo. Por tanto, podemos concluir que las PAE aplicadas están siendo efectivas. Como es de esperar dado su diferencia de diseño, los diferentes grupos de acciones tienen impacto en dimensiones diferentes del mercado laboral. Así las acciones de formación e inserción laboral incrementan en mayor medida la empleabilidad, mientras que las acciones de fomento de empleo son más efectivas en cuanto a duración.

Como conclusión, los métodos de evaluación de políticas basados en variaciones de la metodología de EAC puede proporcionar una alternativa a los habituales métodos *baseline*. Esta alternativa no presenta las dificultades logísticas ni éticas que conllevan los EAC y son capaces de discriminar cuál es el impacto que se puede atribuir a una política. Sería deseable que los métodos con base cuasi-experimental vayan sustituyendo pronto a los habituales métodos *baseline* para optimizar la efectividad de las políticas y el uso de los recursos públicos.

## Referencias bibliográficas

Athey, S., & Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, *31*(2), 3-32.

Bédécarrats, F., Guérin, I., & Roubaud, F. (2018). *The gold standard for randomised evaluations: from discussion of method to political economy* (Working Papers CEB 15-009). Universite Libre de Bruxelles.

Real Decreto-ley 3/2011, de medidas urgentes para la mejora de la empleabilidad y la reforma de las políticas activas de empleo. *Boletin Oficial del Estado,* núm. 43, de 19 de diciembre de 2011, pp. 19240 a 19260.

Caliendo, M., & Schmidl, R. (2016). Youth unemployment and active labor market policies in Europe. *IZA Journal of Labor Policy*, *5*(1), 1.

Grossman, J., & Mackenzie, F. J. (2005). The randomized controlled trial: gold standard, or merely standard? *Perspectives in Biology and Medicine*, *48*(4), 516-534.

Lowndes, V., Marsh, D., & Stoker, G. (Eds.). (2017). *Theory and methods in political science*. Macmillan International Higher Education.

McDermott, R. (2002). Experimental methodology in political science. *Political Analysis*, *10*(4), 325-342.

Mueller, C. E., & Gaus, H. (2017). Quasi-experimental comparison group designs for social policy evaluation. In *Handbook of Social Policy Evaluation* (pp 38-56). Edward Elgar Publishing.

Vooren, M., Haelermans, C., Groot, W., & van den Brink, H. M. (2019). *The Effectiveness of Active Labour Market Policies: A Systematic Meta-Analysis*. https://doi.org/10.1111/joes.12269

White, H., & Sabarwal, S. (2014). *Quasi-Experimental Design and Methods: Methodological Briefs-Impact Evaluation No. 8*. https://www.unicef-irc.org/publications/753-quasi-experimental-design-and-methods-methodological-briefs-impact-evaluation-no.html

CUADERNOS ECONÓMICOS ICE INFORMACIÓN COMERCIAL ESPAÑOLA

# Experimental Trends in Corruption and Tax Evasion

## *Tendencias experimentales sobre corrupción y evasión fiscal*

Nuria Rodríguez-Priego
Ágnes Pintér
Universidad Autónoma de Madrid

***Abstract***

*Nowadays, among the main social concerns of the Spanish society we find corruption and fraud, which occupy the second place in terms of relevance according to recent data. The traditional neoclassical economic approaches concentrate on the construction of economic models that aim to explain the causes of both social phenomena. However, in the last decade several studies have emerged focusing on a behavioral approach to improve the explanatory capacity of such models. In this sense, economic experiments have been implemented to examine, from a different perspective, the determinants of corruption, on the one hand, and fiscal fraud, on the other. Few empirical studies relate both social problems, especially when we refer to the tax evasion of individual economic agents. Therefore, this paper aims to review and deepen the understanding of the factors that determine and relate tax evasion and corruption, from an experimental approach.*

***Keywords:*** *tax evasion, fraud, corruption, behavioral economics, economic experiments.*
***JEL Classification:*** *H20, H26.*

***Resumen***

*Entre las principales preocupaciones de la sociedad española, encontramos la corrupción y el fraude, que ocupan el segundo lugar en términos de relevancia según datos recientes. Los enfoques neoclásicos tradicionales se fundamentan en la construcción de modelos económicos que intentan explicar las causas de ambos fenómenos. Sin embargo, en la última década han surgido varios estudios centrados en un enfoque comportamental para mejorar la capacidad explicativa de dichos modelos. En este sentido, se han implementado experimentos económicos para estudiar, desde una perspectiva diferente, los determinantes de la corrupción, por un lado, y del fraude fiscal, por el otro. No obstante, pocos estudios empíricos relacionan ambos problemas sociales, especialmente cuando nos referimos a la evasión fiscal de agentes económicos individuales. Por ello, el objetivo de este artículo es revisar y profundizar en la comprensión de los factores que determinan y relacionan la evasión fiscal y la corrupción, desde el enfoque experimental.*

***Palabras clave:*** *evasión fiscal, fraude, corrupción, economía del comportamiento, experimentos económicos.*
***Clasificación JEL:*** *H20, H26.*


DOI: https://doi.org/10.32796/cice.2020.99.7007
Recibido: Febrero 2020 · Aceptado: Abril 2020

## 1. Introduction

Fraud and corruption are the second cause of concern among the citizens in the Spanish context, being only overcome by unemployment. These two problems have been escalating positions from 2011 to March 2013, when already 44.5 % of Spaniards placed them among one of the three main concerns. The latest data indicates that 26.2 % of the citizens are still worried about these two issues, keeping them on the second place in the last official barometer.[1]

Tax fraud is a reality of any society where there is a tax system. The submerged economy or shadow economy includes all economic activities that have not been officially accounted for –so they avoid administrative regulation– but would yield an increase in the Gross Domestic Product if they were registered (Feige, 1994; Dell'Anno & Schneider, 2004; Schneider, 2015). Tax fraud is a detriment to public revenues, which reduces the available resources to be allocated on public services and infrastructure, such as education and health, among others, that in turn translates to lower living standards for the whole society. More specifically, tax evasion refers to illegal acts in which fiscal responsibility is hidden or ignored, either intentionally or due to ignorance or confusion.[2] Regardless of intentionality, tax evasion is considered an illegal act and, consequently, is penalized by the government if detected.

Tax evasion was estimated at €46 billion in 2016 for the EU-28, which counts for 0.32 % of the GDP, including three components of tax evasion: capital income, stock of wealth, and evasion on personal income tax. In Spain, the estimation counted for €3.19 billions. This is approximately 7 % of the total estimated revenue loss in EU-28 and 0.3 % of Spain's GDP (Vellutini et al., 2019).

On the other hand, although corruption and tax evasion are two separate social problems, tax contract theories raise the hypothesis that government performance can affect tax collection (Timmons & Garfias, 2015). The data obtained in previous studies show that, when we refer to the decision to declare or evade taxes in the company environment, corruption and the possibilities of bribery of public authorities have a positive effect on the levels of tax evasion (Alm et al., 2016). As for other types of taxes, the evidence shows similar results. In the case of property taxes, the compliance levels decrease as the revealed corruption increases, and increase as audits reveal low (or zero) corruption levels (Timmons & Garfias, 2015).

The concept of corruption is defined as the abuse of public power for private gain (Tanzi, 1998). For companies, the abuse of power could be suffered through the payment of bribes to start some type of business (Choi & Thum, 2005), but other authors have indicated that such abuse could take place more generically. The

[1] Centro de Investigaciones Sociológicas (2019). *Percepción de los principales problemas de España*. http://www.cis.es/cis/opencms/ES/11_barometros/indicadores.html. http://www.cis.es/cis/export/sites/default/-Archivos/Indicadores/documentos_html/TresProblemas.html.

[2] OCDE. *Glossary of Tax Terms (2019)*. Centre for Tax Policy and Administration. http://www.oecd.org/ctp/glossaryoftaxterms.htm.

literature has identified several aspects, such as the diversion of funds, excessive billing of goods and services, and irregularities in the contracting process, including fraud and contracts without tender (Timmons & Garfias, 2015). Although the latter do not have a direct impact on the individual economic agent, it affects his perception of how the resources –collected through the payment of taxes– are used.

Corruption is considered a serious social threat in both developed and developing countries as it decreases productivity at the enterprise level due to inefficient assignments of contracts, distorts economic growth and reduces foreign investment (Habib & Zurawicki, 2002; Aidt, 2011; Faruq et al., 2013). In addition, corruption reduces voluntary contributions to public goods and increases inequality (Beekman et al., 2014; Justesen & Bjørnskov, 2014).

Despite the extensive literature on tax evasion and corruption, research that focuses on combining both problems and examining them from the point of view of behavioral economics is scarce. On the one hand, there are theoretical studies that relate corruption and evasion (Chander & Wilde, 1992; Besley & McLaren, 1993; Hindriks et al., 1999; Sanyal et al., 2000; Acconcia et al., 2003), but these studies are based on economic models that are not very flexible and do not consider the possibility of behavioral biases. On the other hand, as far as our knowledge goes, studies that adopt an experimental approach mainly focus on examining the relationship between corruption and tax evasion from the point of view of companies (Alm et al., 2016), or on the property tax (Timmons & Garfias, 2015).

The objective of our paper is to resume and deepen the knowledge of the factors that determine tax evasion, on the one hand, and corruption, on the other, from the perspective of the experimental economy. To do this, we will take into account that individuals do not always follow the assumptions on which traditional theories are based, hence sometimes they do not behave in a purely selfish way, but rather care about concepts such as justice and fairness, social welfare, and others (Andreoni & Miller, 2002; Kahneman et al., 1986). This study would thus contribute to the empirical literature by providing information on the incentives that lead individuals to evade taxes, focusing specifically on the income tax of natural persons, since, as we have mentioned before, tax evasion related to individual economic agents is more difficult to detect by the administration. Moreover, we will explore the possible relationship between tax evasion and corruption, as separately both issues are widely studied, but few experimental studies tried to analyze both phenomena together.

## 2. Tax evasion

### *2.1. From traditional to modern economic theories*

Tax *fraud* and tax *evasion* are terms used as synonymous. Both refer to the behavior of an economic agent who hides, disguises or manipulates his economic

capacity by not paying the amount of taxes that he is obliged to pay according to the laws in force. On the other hand, tax *avoidance* is the behavior of taking advantage of a tax rule that was intended for a different purpose by the legislator, with the objective of paying fewer taxes. The importance of tax evasion lies in its relationship with public services, meaning that the more taxpayers evade, the less public services are provided. In addition, it also affects the precision of macroeconomic data and makes difficult to measure the real impact of taxation (Alm, 2012).

Traditional economic theory has tried to explain and measure tax evasion in the last decades. Allingham and Sandmo (1972) analyzed taxpayer's intentional decision to evade taxes, building their economic model on previous literature related with the economics of criminality from Becker (1968) and Tulkens and Jacquemin (1971), and the economics of uncertainty (Mossin, 1968). Their model is based on individual personal income and assumes that the taxpayer is risk averse. When deciding how much income to declare, the taxpayer can choose to declare his real income, or a lower quantity. In any case, the authorities may decide to investigate him. If the economic agent decides to evade, he will have to pay a penalty based on the quantity he has undeclared, plus the corresponding undeclared taxes. However, if the authorities do not audit him, he will be in a better situation compared to absolute tax compliance. The economic agent's decision responds to his expected utility maximization. The optimum is subject to the probability of being investigated, the size of the penalty, and the degree of risk aversion of the taxpayer. As a result, individuals pay taxes only if they think they will be caught and punished.

In this model, actual income is exogenously given and noted as $W$. This quantity is known by the taxpayer but not by the public authorities. The tax rate is set at $\theta$ on declared income, $X$. There is a probability of being audited $p$. Hence, the taxpayer pays $\theta \cdot X$, unless he is audited after deciding to underreport. In this case, he will have to pay the underreported quantity $W - X$ plus a penalty rate, $\pi$ ($\pi > \theta$). The taxpayer will choose $X$ so that he maximizes his expected utility:

$$E[U] = (1 - p) \cdot U(W - \theta \cdot X) + p \cdot U(W - \theta \cdot X - \pi(W - X)) \quad [1]$$

If we note:

$$Y = W - \theta \cdot X$$
$$Z = W - \theta \cdot X - \pi(W - X) \quad [2]$$

The first order conditions equals $E'[U]$ to zero:

$$-\theta(1 - p) \cdot U'(Y) - (\theta - \pi)p \cdot U'(Z) = 0 \quad [3]$$

The second order condition

$$D = \theta^2(1-p)U''(Y) + (\theta-\pi)^2 p \cdot U''(Z) \quad [4]$$

is negative because we assume concavity of the utility function.

If we set the expected utility at $X = 0$ and $X = W$, we can evaluate whether it exists an interior solution. It would require that:

$$\frac{\partial E[U]}{\partial X}_{X=0} = -\theta(1-p) \cdot U'(W) - (\theta-\pi)p \cdot U(W(1-\pi)) > 0 \quad [5]$$

and

$$\frac{\partial E[U]}{\partial X}_{X=W} = -\theta(1-p) \cdot U'(W(1-\theta)) - (\theta-\pi)p \cdot U(W'(1-\theta)) < 0 \quad [6]$$

We can rewrite these conditions as:

$$p\pi > \theta\left[p + (1-p)\frac{U'(W)}{U'(W(1-\theta))}\right] \quad [7]$$

$$p\pi < \theta \quad [8]$$

Expression [8] means that the taxpayer will evade if the probability of being audited multiplied by the penalty is lower than the tax rate. Both [7] and [8] guarantee an interior solution.

Other authors have introduced new elements to the basic Allingham and Sandmo (A-S) model. For example, Yitzhaki (1974) stated that the degree of tax evasion should decrease when the penalty rate ($\pi$) or the probability of being audited ($p$) increases. These results follow intuitive expectations and empirical evidence. This model also predicts a counterintuitive effect, which refers to the fact that the degree of tax evasion ($W - X$) should decrease as the tax rate ($\theta$) increases, although empirical results go in the opposite way and find that higher tax rates decrease the levels of compliance (Clotfelter, 1983; Hashimzade et al., 2013).

A common criticism to the A-S model and its extension is that they do not account for tax morale. It means that when the probability of being detected ($p$) or the penalty rate ($\pi$) is low enough, the agent's decision should be to underreport or evade taxes. The decision to evade should be even more obvious when there is no other source of information than the declared income of the individual itself. However, there is always a percentage of willful payment, which cannot be explained by the basic model (Alm, 2012). In order to correct some of these potential weaknesses, Traxler (2010) included in the A-S model the effect of social norms, which depends on the tax rate sensitivity of the taxpayer.

Dhami and al-Nowaihi (2007) examine this particular aspect of the A-S model, namely, the surprisingly high levels of compliance that are not explained by the expected utility theory. According to their research, the levels of tax evasion should be greater if we consider the expected utility theory, given the relatively low audit probabilities and penalties. They propose an alternative theoretical model based on cumulative prospect theory. This model considers loss aversion with respect to a referred income, and it suggests that individuals tend to overweight small probabilities and to underweight large ones. Their theory seems to predict more precisely the extent of tax evasion when the audit probability and the penalty are low. Traxler (2010) incorporates the concept of tax morale as a social norm into the traditional model of tax evasion with the purpose of determining why taxpayers comply when the audit probability and fine are low. He claims that behavior is influenced by social norms related to compliance with a reference group. Additionally, Lago-Peñas and Lago-Peñas (2010) have found that sociodemographic characteristics, personal financial experiences, political attitudes and other variables also have an influence on tax morale.

Further theoretical extensions to the basic model are still grounded on its approach to crime and the expected utility theory. The variations include alternative functions to measure the effect of the imposed penalty, or incorporate more complexity and uncertainty on the fiscal parameters, make the labor supply endogenous, increasing the complexity of the income tax, and other. However, these models make more complex the statistics of fiscal compliance (Alm, 2012; Hashimzade et al., 2013).

Andreoni et al. (1998) review the literature changing the focal point to compliance instead of tax evasion. They also criticize the A-S model, as it assumes that the audit probability is constant and fixed ex-ante. On contrast, theoretical models predict the probability of being audited as a function of income, differentiating between two groups. The first group include principal-agent models where the audit probability is determined ex-ante by the public institution, and its decision has the objective of maximizing audit revenues. The second group refers to models where the audit probability is determined ex-post, after receiving the economic agents' declarations, and the decision is solved as a sequential game between the taxpayer and the government. Both types of models assume that taxpayers only provide information to the tax authorities on how much income they declare, though they also provide many other sources that can signal their level of tax compliance. Furthermore, in this study, the authors review the literature related to fiscal compliance and discuss three variables that seem to have an effect on it: moral rules, feelings of fairness toward the taxes or its prosecution, and taxpayer evaluation of government expenditure and corruption.

In this paper, we will pay special attention to the effect that this last variable, corruption, may have on tax evasion and compliance.

### *2.2. Mind the tax gap*

The *tax gap* is the difference between the quantity that taxpayers are obliged to pay by law, and the quantity they finally report and pay (Mazur et al., 2007). We can differentiate between tax gap coming from individual economic agents' reported taxes, like the income tax, and taxes related to firms, such as corporate taxes, sales taxes, value added taxes, property taxes and other. When estimating this gap, the literature shows different methods. Alm (2012) distinguishes between *traditional* and *modern* approaches. Regarding the traditional approaches, this author refers to *direct* vs. *indirect* methods.

The first direct method is based on random audits with the aim of detecting tax evaders. The main inconveniences of this method are that it is costly for the government; hence, the proportion of audits over the total sample of declarations cannot be very high. On the other hand, taxpayers may not know the probability of being audited by the authorities, which makes their decision rely on their risk aversion. In addition, when using this method, the computed level of tax evasion results from an estimation assuming that the rest of the tax filer population behaves in the same way as the audited sample does.

A second direct method to measure the tax gap is through surveys where subjects report if they evade or not. Moreover, this method provides information regarding sociodemographic and attitudinal variables (Andreoni et al., 1998). However, subjects participating in a survey where they have to self-report tax avoidance, have incentives to cheat if they think that their answers are potentially incriminating (Baumeister, 1982).

The last direct information on tax evasion may come from data originated from tax amnesties. Although the data obtained through this method detects real tax evasion, it only accounts for a small and unknown proportion of the evaded quantity in the whole economy. However, it can become a useful mean to recover part of the public revenues, and to make emerge black money from the shadow economy (Andreoni et al., 1998).

Regarding indirect methods, the literature has flourished in the last decades. One of the most common indirect approaches estimates tax evasion using the *shadow economy* as a proxy. The shadow economy has been defined as the set of economic activities that are not registered when computing the GDP of a country, but that would contribute to increase it if they were accounted for (Feige, 1994). It can be inferred through measurable data such as the electricity consumption, the currency demand, or national income accounts (Slemrod, 2007), and many other methods (e.g. MIMIC model, monetary approach, etc.).

Although the shadow economy and the levels of tax evasion in a country go hand in hand, these are not synonymous. There can be not declared economic activities that should count as part of the shadow economy, but it does not always imply that these activities would generate taxes, hence individuals in these cases could not be considered as tax evaders. In addition, there can be economic agents that declare

their activity, but do not declare all their income. Hence, they would be tax evaders, but their activity should not count for the shadow economy estimations (Fernández-Leiceaga et al., 2018).

Other indirect methods come from a *behavioral economics* approach. Pickhardt and Prinx (2014) review the behavioral dynamics of tax evasion and differentiate between four types of subgames depending on the economic agents that interact: taxpayers vs. taxpayers, taxpayers vs. authority, taxpayers with practitioners vs. authority, and tax lawmakers vs. taxpayers. For the present paper, we will focus on the second type of subgames, where relationships can be dominated by *trust* in the authorities from taxpayers' side (or vice versa), or by the *power* exerted from the authorities' side to detect and punish tax evasion (Kastlunger et al., 2013; Kogler et al., 2013).

This last approach includes both lab and field experiments. On one hand, lab experiments offer a controlled environment where the researcher can decide the audit probability, the quantity of the penalty, and many other variables. However, the results must be interpreted cautiously (Andreoni et al., 1998). On the other hand, some researchers have conducted controlled field experiments where the tax authorities send letters to taxpayers testing whether different pieces of information have an effect on the quantity of income reported (Slemrod et al., 2001; Iyer et al., 2010; Biddle et al., 2018). As the purpose of this paper is to focus on *experimental trends*, the next section provides some examples of the behavioral economics literature on tax compliance and evasion.

## *2.3. Experiments on taxation*

### 2.3.1. Lab experiments

In the last decades, lab experiments in tax evasion have flourished with the purpose to combine economic and psychological theory with evidence (Pickhardt & Prinz, 2014). The main advantage of conducting experiments in a lab is the possibility to control the process, as well as deciding ex-ante the audit probability, the quantity of the penalty, and the tax rate, among other variables. In addition, we can also decide the incentives that subjects will receive and establish quotas for the sample under study.

The usual process begins with the experimental design in which the researcher decides the number of treatments to be tested, and prepare the instructions for the experiment, usually using a neutral wording in order to avoid framing effects. Subsequently, recruited subjects are assigned to the various treatments. Participation can be incentivized with a fixed amount of money, providing subjects with an initial endowment, or with a variable amount, assigning the final payment based on the number of tasks and the results that participants perform during the experiment, or a mixture of both. Usually, a show-up fee is added to these incentives. The experimenter

can also choose the matching protocol (i.e., one-shot or repeated) that fits better the studied (real-life) scenario. Repeated protocols simulate decisions that take place on regular basis. In this case, the experimenter can also decide whether to repeat the game using partner matching (meaning that the same partners or opponents play the game repeatedly over several repetitions) or stranger matching (the game is played repeatedly but matched to a different partner every round). In some experiments, the amount collected through taxes is assigned to a public good, what participants may benefit from at the end of the experiment.

Andreoni et al. (1998) review the main outcomes resulting from lab experiments since 1978. Alm (2012) also provides an extensive review on experiments conducted on tax compliance. Both papers find that lab experiments relate high tax rates with greater evasion, which contradicts the standard theory. Regarding the penalty rate and probability of being audited, both variables have a positive influence on compliance. This last result is in line with a meta-analysis conducted for 20 lab experiments (see Blackwell, 2007). In addition, the probability of being detected evading will raise the perception of being audited in the future, which will influence future tax compliance. This is a consequence of individuals making poor predictions on the probability of audit and penalty rates for evasion, tending to overestimate both of them.

One of the aspects that deserves further attention of the experiments conducted in labs is the fact that taxpayers do not always know the objective probability of an audit, but they have a subjective view of it, which is usually overestimated. It means that in lab experiments there is a higher compliance rate than what the expected utility theory predicts (Alm, 2012).

When comparing with empirical methods, lab experiments let the researcher introduce more variables in a controlled way and generate precise data on taxpayers' decisions. However, external validity of results can be questioned. Hence, results must be interpreted cautiously, and replications are needed to establish sufficiently grounded conclusions (Alm, 2012).

### 2.3.2. Field experiments

Empirical studies on tax evasion can circumvent the question of external validity of experimental methods. However, as mentioned before, this method does not let the researcher the freedom to introduce or omit variables, and reduces the level of control exerted by the experimenter.

Most field studies test the effects of potential determinants of compliance by using a physical letter. The common procedure is to send a "deterrence letter", carefully designed according to the objective of the study, to a (randomly) selected set of taxpayers from behalf of the tax authority. Then, the results of tax compliance are compared with a control group or with the behavior of the same subjects in a previous period. The novelty of this method was that the letters were sent directly from the revenue administration, therefore the experiment captures real behavior. Secondly,

it does not change the rules (the legislation), the letters only mean to increase the perceived probability of audit or to make taxpayers aware of their obligations and the consequences of their compliance decisions.

Blumenthal et al. (2001) conducted a first field experiment in Minnesota assessing the effects of two letters with normative appeals compared to a control group. Though, they found no treatment effect, letters seemed to impact tax compliance for some groups of taxpayers.

In a second experiment in Minnesota, the tax authority sent letters to a group of taxpayers that was selected along two dimensions: income and opportunity to evade. The letters notified these individuals that they had been selected to be part of a study and that their tax declarations would be "closely examined". The letters had a significant impact on reported income, although the effect varies across categories: tax revenues increased for low and middle-income taxpayers (combined with high evasion opportunity the effect is stronger than for those with low-opportunity), but the result was the opposite for high-income taxpayers (Slemrod et al., 2001). Following this point, Kleven et al. (2011), in a field experiment conducted in Denmark, found that letters warning about the possibility of being audited had a significant effect on self-reporting taxpayers increasing their tax compliance, but it had no effect on third-party workers.

Carrillo et al. (2013) evaluate the effectiveness of deterrence letters in Argentina, combining the usual factors (probability of audit and penalties) with equity and fairness. In their treatments the deterrence message is much softer than in the previously mentioned articles. They find that except of the letters influencing the perceived probability of audit and the penalties for evasion, the studied factors do not have a significant effect on compliance.

In another study conducted in Washington State, researchers sent letters to firms with the purpose of testing the effects of an increased perceived detection risk and/or the awareness of penalty ($2 \times 2$ design). They found that penalty-related letters lead to an increase in the declared income (used as the base for calculating tax obligations). The detection and penalty manipulations showed significant effects for taxpayers with decreasing revenues, but not for taxpayers with increasing revenues (Iyer et al., 2010).

Torgler (2004, 2013) conducted two studies in Switzerland to test the effect of moral persuasion on tax compliance. The researcher found hardly any effect on taxpayers' compliance behavior in any of the field experiments.

In Australia, the tax authorities conducted two field experiments with the purpose of nudging businesses to comply with their tax obligations. In the first study, they found no significant effect after changing internal guidelines used by field auditors to raise awareness of the relevance of tax debt payments. In the second study, however, they found that simplifying a follow-up letter sent to the firms and changing the phone script of desk auditors reduced the proportion of default assessment. While the previous studies were concerned about finding the most effective content of the letters sent to the taxpayers, some recent studies ask what the most effective

way to deliver the message is. Nowadays most tax administrations have a great variety of ways to communicate with their clients besides physical letters: e-mails, SMS, phone-calls, messages, or personal visits. Doerrenberg and Schmitz (2015) carried out a field experiment, sending the same message (combining deterrence and moral suasion) to small firms by post or through the personal visit of a tax official. Compared to a non-treated control group, both treated groups report a higher taxable income but the differences (probably due to the relatively small sample size) are not significant. In another experiment, Ortega and Scartascini (2016) circumvent the problem of sample size and test the effectiveness of three delivery methods: letters, e-mails and personal visits, sharing exactly the same message with taxpayers, in any of the three cases. They find that all three methods are effective in encouraging the payment of outstanding tax liabilities, but personal visits are extraordinarily successful. Given that most tax administration do not have the capacity to employ this highly expensive but very effective method to increase tax compliance, the authors recommend a combined use of the existing delivery methods in order to achieve large revenue gains.

Although the literature on field experiments is large, given the differences in design and methodology, it seems to be difficult to find robust and meaningful effects as to compare to those of laboratory experiments. In addition, the method and the context in which field experiments are conducted makes it more difficult to obtain reliable information from the participants on the reasons behind the found effects. Though, both lab and field experiments suggest that raising either the fine or the (perceived) probability of audit will usually increase tax compliance.

On the other hand, most field and lab experiments assess the immediate effects of the treatments, only a few papers (Carrillo et al., 2016; Dunning et al., 2016; Manoli & Turner, 2014) raise the question of longer-term effects and sustainability. In general, they suggest that the learning effects of the treatments are short-term (no longer than 3 years), therefore, to maintain the achieved compliance levels, regular communication between the tax administration and the taxpayers would be advisable. Importantly, learning in real life situations can happen through a different channel as well. As some recent articles highlight it, learning can not only happen through the official channels, but also on one's social network. Alstadsaeter et al. (2014) and Drago et al. (2015) show evidence on this kind of learning, and emphasize the importance of the found large spillover effects: non-treated taxpayers who form part of the same social network as some treated one, were more likely to increase tax compliance. Surprisingly, these spillover effects were as large as the direct treatment effects. These results have important policy implications and open new lines of research in order to make better use of the resources.

## 3. Fraud/Corruption

Corruption and tax fraud have been studied separately in the literature for a long time, and already their effects on the economy are considerable. Although earlier research claimed that corruption may increase efficiency (Leff, 1964; Huttington, 1968), a broad set of empirical evidence shows the dramatic effects that corruption and fraud have on a country's social and economic development. Corruption and fraud hinder economic growth (for example, Fisman & Svensson, 2007; Meon & Sekkat, 2005) and increase poverty and inequality (Gupta et al., 1998; Olken, 2006). But there is no doubt that –as fiscal contract theories hypothesize– tax compliance is deeply intertwined with the political environment that influences the social contract between the citizens and the government. Clearly, taxpaying is the result of both a quasi-voluntary compliance and (costly) law enforcement, and both components can be affected by the degree of perceived or revealed corruption. Low corruption and government accountability have been emphasized as the pre-requisites for a country's successful tax performance (Bird et al., 2008).

Economic theory of corruption has focused on the design of institutions that can prevent public officials from abusing their position. Theoretical models suggest that when private returns to corruption are high, or when the probability or consequences of detection are limited (for example due to weak institutions or capacity constraints), individuals are more bound to act corruptly. Therefore, such institutions should set up a system of rewards and punishments that render the expected return of corruption lower than the returns to truthful behavior. Hence, the (early) literature on corruption theory focuses on the wages of public officials, the principal's monitoring capacity (that will determine the detection probability), and other possible penalties. More recent studies show, that as the proportion of corrupt individuals in a society increases, it gets easier (and cheaper) to find a partner with whom to engage in a corrupt transaction and escape detection and/or punishment. This, in turn, results in multiple equilibria involving different levels of corruption, and a country can find itself in a *corruption trap*, from which it is extremely difficult to escape. Moreover, as Hauk and Saez-Marti (2002) show, the level of corruption in a society might influence younger generations' values and attitude towards corruption. Of course, monitoring and incentive systems, together with the government's commitment keep being important pillars in the fight against corruption. Recent theories suggest that well-designed leniency programs can decrease the cost of monitoring and prosecuting corruption. On the other hand, uncertainty regarding the corruptibility of one's partner can also prevent agents from corrupt decisions (Ryvkin & Serra, 2012), therefore solutions like staff-rotation gained attention in recent (tax) administration reforms.

Corruption, besides the economic perspective, has a second dimension, studied by psychologists and cultural anthropologists. They argue that besides the external incentives, intrinsic motivations (like reciprocity, shame, guilt for breaking some social norms, culture, etc.) also play an important role in an individual's choice between

honesty and corruption. Intrinsic motivations are the results of the internalization of social norms, but social norms vary across societies, therefore different policies might be needed in different societies. Moreover, some results suggest that in some cases the external incentives (namely, sanctions) aimed at deterring corruption are counterproductive, as they crowd out intrinsic motivations. What before was a question of right or wrong, now becomes a simple decision problem where the calculus of risks and benefits determines the choice.

In summary, we can see that the theory of corruption studies the phenomena in two interrelated dimensions: on one hand, on individual level, studying the incentives of an agent to engage in a corrupt activity –usually modelled with some kind of bribery game– and on aggregate level, assessing the role of the society. Although the theoretical results are derived in a general setting, they have important implications when applied for the tax administration.

### *3.1. Empirical results on corruption and tax compliance*

There is now a vast and growing empirical literature analyzing the determinants of tax compliance, suggested by the theory, but only a small part of this literature accounts for its relation to corruption, generally on aggregate, or firm level. Given the secretive nature of corrupt interactions, it is difficult to study in the field the determinants and possible solutions to it in a reliable way. Several important drawbacks of the empirical studies should make us careful when evaluating their findings. First, these are usually cross-country econometric studies, that use subjective measures of corruption and might carry important measurement errors. Moreover, the results of different authors often contradict each other, possibly due to different measures, empirical specifications, or some problems with the data set (multicollinearity, omitted variable bias, etc.). Some factors that might be decisive in measuring the degree of corruption of a country (like democratic institutions, level of development, etc.) are endogenous to the existing level of corruption. Finally, from the cross-country econometric studies we do not get information on individual decisions on corrupt behavior. Naturally, collect information in the field about corruption on individual level is complicated as it involves illegal acts and the agents engaged in such activities would rather prefer to hide their participation.

Recent literature uses alternative methods to measure corruption in the field in a more accurate and reliable way. The idea is to measure corruption as the "missing money". These methods include for example the comparison of administrative data on public funds obtained at different levels of the government (Reinnika & Svensson, 2004), records of exported and imported goods produced by the exporting and importing country (Fisman & Wei, 2004), or data of corruption audits of different years (Ferranz & Finan, 2011). Although these field studies increased the understanding of the factors and costs of corruption in some settings, only a few rely on direct corruption measures (Sequeira, 2012; Olken & Barron, 2009).

As we mentioned above, field studies on corruption in the domain of tax compliance are scarce and concentrate on aggregate data. Despite all the difficulties that we discussed before regarding the study of corrupt behavior, the advantage of these studies is that the analyzed data comes from real taxpaying decisions. Most empirical papers study the determinants and consequences of corruption (in the country or in the tax administration) on the tax revenues, usually in developing countries.

Fjeldstad (2005) studies the experiences in the fight against corruption in the Uganda Revenue Authority (URA). Similarly to a number of countries in Latin America, the reform in the tax administration appeared to be successful in the first years, meaning a significant increase in the tax revenues and decline in corruption. However, these positive changes vanished in the following years. The main explanations for this change are the public perception of a high level of corruption, some external factors, like general economic trends, import restrictions, or the decreased staff productivity due to lack of motivation. However, there are many indications that an increase in fiscal corruption has contributed to the fall in the reported revenues. The study points out that corruption and tax evasion seem to increase in all levels of the URA, and even the number of tax officials openly demanding bribes has increased. The article identifies several factors that contributed to this situation: declining wages; bonus systems and revenue targets; the used policy to hire and (not) fire workers at operative levels; human resource management and job insecurity that affected primarily the managers, and turned out to increase their corruptibility; political interference; patronage and taxpayers' low compliance often due to their eroded confidence in the fairness and impartiality of the tax administration. The paper concludes that in an environment where there is a huge demand for corrupt services and monitoring –due to different factors– is ineffective, higher wages rather than restraining corruption might work as a bonus on top of the bribes taken by corrupt officers. One of the reasons behind this is the impact of social and family networks over which the URA has no influence. Therefore, one of the challenges to increase the efficiency is to change the impact of traditional values and social obligations on the behavior of public officials. On the other hand, the complicated and non-transparent tax structure, together with the lack of taxpaying culture in the country also represent an obstacle to building a stable, long-term revenue base in Uganda.

Timmons and Garfias (2015) use randomized auditing reports from the Brazilian federal government to study the effects of information about corruption on firms' property tax reports. Besides analyzing the relation between revealed corruption and tax revenues, they also relate these issues with fiscal accountability and the likelihood of adopting Participatory Budgeting (PB). Their main objective is to understand the role of information on tax collection and government performance. The authors test the hypothesis that the use of PB (i.e., increased and direct citizens' control over resource allocation) would decrease government corruption, improve government performance and increase the satisfaction of the society, therefore raise tax revenues.

In other words, on one hand, revealed corruption should incentivize the adoption of PB and other fiscally accountable institutions, and on the other hand, tax revenue should increase with good reports on corruption (and decrease with bad reports). These hypotheses are confirmed by the data, but importantly, the changes in property tax revenue are only experienced on the short run. Therefore, repeated access to (reliable) audit reports should be granted to citizens in order to maintain the achieved results. The data also confirms that revealed corruption increases the probability of adoption of PB.

Uslaner (2010) studies empirically the relation between firm tax evasion and corruption in some transition countries during 2002-2005, based on surveys of businesspeople (not ordinary citizens) in 26 formerly communist countries. These countries are, most probably partially due to communist heritage, different from the West: tax morale is rather low, the shadow economy is much greater, the corruption is far worse, the quality of basic (public) services is poor, the legal system is generally seen as unfair, and economic inequality is –although already higher than in the West– growing. Therefore, the people in transition countries see corruption as long-term, insoluble problem –although in many cases, corruption in the course of one's duties (like gift payments, gratitude money) is not even perceived as immoral. The results of the paper confirm that the perception of the quality of government services and the degree of corruption play a key role in tax compliance. Other factors, like the size and ownership of the firm, the fairness of the legal system, the tax rate, or the perceived probability of audits also matter but to a smaller extent.

In a recent paper with wider focus, Alm et al. (2016) use firm-level data on tax reporting to study how the potential for bribery of tax officials affects the tax report of a firm. Their empirical framework assumes that the firm decides how much to report, taking into consideration that bribing a corrupt official is also possible. Their results confirm the argument that the quality of the tax enforcement is one of the prime determinants of tax compliance. They find that corruption of tax officials is a significant determinant of tax evasion, as with corrupt tax officials the reported sales (for taxes) reduce significantly. Moreover, the higher the bribes required, higher the level of tax evasion. Therefore, they conclude that governments seeking to fight tax evasion, and increase the tax revenues should first ensure the honesty of the tax administration.

The issue of the relation between public service delivery and tax compliance has been studied in field experiments as well, although typically with a different focus and inconclusive results. They generally find little support for its effectiveness in increasing tax compliance, revealing either no effect (Ariel, 2012) or much smaller effect than other moral appeals (Bott et al., 2014; Hallsworth et al., 2017). However, others (Castro & Scartascini, 2013; Carrillo et al., 2016) find some evidence that suggest that a higher level of public goods provision is associated with a higher tax compliance.

To our best knowledge, there are no empirical papers that relate directly corruption and tax compliance on individual level. In this aspect, laboratory experiments can

contribute effectively to our understanding of the factors that affect individual decisions related to tax payments in the presence of corruption.

### *3.2. Laboratory experiments*

Laboratory experiments are able to create scenarios that mimic real-life situations combining corruption and individual tax decisions. The use of experimental methods and data gathering can be parallel to the collection of field data, as the advantages of one method are the drawbacks of the other, and vice versa. In laboratory experiments the data is obtained in a controlled, though highly stylized and artificial, environment and the experimenter has control over many individual factors that can be tested while keeping all others constant. Taking into account that individual tax decisions including dishonest behavior (like engaging in corruption) are difficult to observe in real-life data, laboratory experiments can result very useful. Although there are very few experimental studies that relate explicitly corruption with individual tax decisions and tax compliance, the results of the existing experimental literature on corruption in general can be relevant in this area as well. Since corruption has many facets, experimentalists modelled it in many different ways, according to the studied situation and the objectives of the study. Hence, although this results in a higher diversity, it also means that there is no consensus either on the design or on the policy implications of the results.

A bunch of experimental studies on corruption test the institutional factors suggested by the theory. Bribery (i.e., private payment to a public official in exchange of a benefit or the avoidance of some cost) in most countries is considered an illegal act and is persecuted by the governments, although the probability of detection and the consequences vary significantly across the world. Bribery can be seen as a trust game between the briber and the bribe, as there is no legal contract between the involved agents, due to the illegality of the act. The effects of being detected and the consequent punishment on the individual decisions can be neatly studied in the lab in a controlled environment.

Abbink et al. (2002) model bribery as a situation of undesirable reciprocity. They implement in the lab a simple two-player, three-stage bribery game with three treatments, in order to isolate the effect of (i) the reciprocal relation between agents, (ii) the negative externality on the public, and (iii) the severe punishment when being caught. The reciprocity is identified through the extent of exchange of favors, that are individually beneficial but harmful to the public thus prohibited. To isolate the negative externality caused to the others, in the second treatment the reciprocal answer to the bribe inflicts costs on others. In the third treatment, authors add a severe punishment that the corrupt agents suffer when caught. Their results show that the costs inflicted on others do not affect neither the average bribe, nor the average number of reciprocal choices. However, the possibility of a severe punishment acts as a strong deterrence factor against corruption.

In a follow-up study Abbink (2004) studies the effect of staff rotation on corrupt behavior, simply re-matching partners in each round in the above described bribery game, repeated over 30 rounds. He finds that both average bribes and the average number of reciprocal choices decreased, the effect on the latter being stronger. On the other hand, the effects over time can not be detected. This suggests that staff rotation can be a powerful tool in reducing bribery, but the effects are short run and the frequent reallocation of public officials might cause extra costs and efficiency losses (both on individual and institutional level).

Schulze and Frank (2003) conduct a lab-in-field kind experimental bribery game without strategic interaction. Participants had to make an individual decision in a hypothetical situation, choosing the size of the bribe that came included in the decision. In their setting, the probability of detection of the corruption increases with the size of the bribe, and a penalty is applied when corruption detected. Compared to the case without penalty, the detection mechanism does not reduce the frequency of bribery, it only shifts the distribution of bribes to the moderate levels, eliminating the very high bribes and the honest decisions as well. Hence, the authors find evidence on a crowding out of intrinsic motivation.

Berninghaus et al. (2013) study the corrupt behavior of public officials in a game, where corruption is modelled as a coordination game. In the game public officials have to choose between a safe (honest) option with a fixed return and a risky (corrupt) one that yields zero payoff if the corrupt decision is discovered and a high payoff if not. The probability of getting caught depends negatively on the number of corrupt officials, mimicking the empirical observation that in a corrupt society, dishonest officials have less to fear than in one where corruption is seen immoral. In this setting, two pure-strategy equilibria exist in which either no-one or everyone is corrupt. The authors find that beliefs are a stronger factor than risk attitudes in predicting a subject's propensity of becoming corrupt.

Serra (2011) studies a two-stage bribery game without reciprocity: in the first stage the public official can ask for a bribe, and in the second stage the client decides whether or not to pay it. In the first case the decision that favors the client is automatically implemented, which in turn inflicts negative externalities on some passive players (citizens). The paper analyzes two treatments, representing two approaches to combat corruption: in the first treatment the probability of getting punished when asking for a bribe is fixed; while in the second treatment, the probability of getting caught is still 4 %, but this lottery is only played out if a client reports the official's request for bribe to the authorities. The author finds that bribe demands are significantly lower in the second treatment despite the lower risk of suffering a penalty, but unfortunately the data –due to the design of the experiment– is not able to explain the reasons of this observation.

A partially related situation is studied in Schikora (2011), who studies the effects of whistle blowing with and without leniency in the lab. In their experiment, fixed pairs of players interact repeatedly for the provision of a service in a neutrally described situation. Any of the two agents can initiate a corrupt transaction by

either offering or demanding a bribe in the different stages of the game. If a bribe is paid, the bribe can (but is not obliged to) reciprocate it. This reciprocity represents a corrupt decision, causing negative externalities on other participants, which are also involved in the same experiment on decision-making. This baseline treatment is compared with symmetric and asymmetric whistle-blowing treatments. In the first treatment any of both agents can blow the whistle when a corrupt act is observed, and the consequences are the same for both agents (independently of who blew the whistle). With asymmetric whistle-blowing the bribee is granted leniency and only the briber suffers the consequences. The author finds that while giving the briber the possibility to blow the whistle against a corrupt official that demanded a bribe is effective in reducing corruption, the briber can also use whistle-blowing as a threat to force the official to reciprocate. On the other hand, whistle-blowing with asymmetric leniency reduces corruption significantly.

Buckenmaier et al. (2018) study collusive corruption within a tax evasion framework in the bribery game introduced in Abbink et al. (2002). In the basic game without leniency, each taxpayer receives a fixed income and is taxed at a fixed rate that has to be reported to a tax official. In this design, underreporting requires the cooperation of the tax official who can be offered a bribe for the assistance in evading taxes. The reports are audited with a fixed probability and in caught, a penalty is paid by the taxpayer but there are no consequences for the tax official. Leniency program is added in an additional stage, where taxpayer can report the corrupted tax official during an audit and avoid the pending penalty, while the official has to pay a fine. With these rules, the risk of being caught and fined is borne exclusively by the tax official. The authors find that in the presence of leniency, the successful collusion between the agents is less frequent than in the baseline game. This result is driven by the low willingness of tax officials to accept bribes, which in turn translates in less evasion and higher tax revenues. On the other hand, the removal of the leniency program does not cause the opposite effect, suggesting that the positive spillover effects of leniency remain even after the mechanism has been removed.

The relation between public officials' wages and corruption (namely, that higher salaries lead to lower corruption) has also been studied by several laboratory experiments. Van Veldhuizen (2011) studies the bribery game of Abbink et al. (2002) with modified relative payoffs between bribers and officials. Moreover, in this case the negative externalities are deducted from a donation to a charity of the subject's choice that the experimenter makes. The results show that corrupt decisions made by officials with high wages are significantly lower than that of their low-wage counterparts. Similar wage effects were found in experiments in the context of embezzlement (Azfar & Nelson, 2003; Barr et al., 2009) and also in the field (Armantier & Boly, 2011).

The role of non-monetary incentives and intrinsic motivations in deterring corruption has also attracted attention in the experimental literature, although not in specific tax-related issues. Different experimental studies that assess the possible effects of framing in corrupt situations (i.e., enhancing participants' awareness of the

corrupt nature of the actions) arrive to contradictory conclusions (Abbink & Hennig-Schmidt, 2006; Barr & Serra, 2009), thus suggesting that the effects of framing depend on the particular experimental setting. On the other hand, as we already mentioned it in some of the previous papers, (negative) externality caused by corrupt decisions seems to have no importance in the decision-making.

## 4. Conclusions

Taxation policies may reveal much about a country's political context. Taxation not only raises revenue for public spending to fight poverty. It can also redistribute wealth and opportunities, diminishing inequalities and strengthening state-society relations by boosting citizens' ability to demand greater accountability.

In this paper we summarize the recent advances in the experimental literature on tax compliance and corruption, presenting both laboratory and field studies. The experimental study of tax compliance in the presence of corruption is an interesting and important question, given the advantages that the controlled environment offers when analyzing activities that in real life are practically unobservable due to its highly secretive nature. The objective of these laboratory experiments can be to test the predictions of the vast theoretical literature on the determinants of tax compliance (with or without corruption), to collect and analyze direct, micro-level data that in the field would be impossible to get, and compare with the available (real-data) measures in order to give policy recommendations. In this line, future experimental research would be needed to produce controlled data to understand the relation between tax compliance on individual level and corruption, where the decision task is clearly and directly related to different levels or forms of corruption. This would help to evaluate the importance of the factors suggested by theory and understand their interaction with the individual tax decisions. Based on robust results backed by theory, field and laboratory data, useful and reliable policy recommendations can then be formulated in order to achieve (and maintain) a higher tax compliance, lower corruption and hence, higher tax revenues.

## References

Abbink, K., Irlenbusch, B., & Renner, E. (2002). An experimental bribery game. *Journal of Law, Economics, and Organization*, *18*(2), 428-454.

Abbink, K. (2004). Staff rotation as an anti-corruption policy: an experimental study. *European Journal of Political Economy*, *20*(4), 887-906.

Abbink, K., & Hennig-Schmidt, H. (2006). Neutral versus loaded instructions in a bribery experiment. *Experimental Economics*, *9*(2), 103-121.

Acconcia, A., D'Amato, M., & Martina, R. (2003). *Corruption and tax evasion with competitive bribes* (Working Paper, No. 112). Centre for Studies in Economics and Finance.

Aidt, T. S. (2011). Corruption and sustainable development. *International handbook on the economics of corruption*, *2*, 3-51.

Allingham M., & Sandmo A. (1972). Income tax evasion: a theoretical analysis, *Journal of Public Economics*, *1*(3-4), 323-338.

Alm, J. (2012). Measuring, explaining, and controlling tax evasion: lessons from theory, experiments, and field studies. *International Tax and Public Finance*, *19*(1), 54-77.

Alm, J., Martinez-Vazquez, J., & McClellan, C. (2016). Corruption and firm tax evasion. *Journal of Economic Behavior & Organization*, *124*, 146-163.

Alstadsaeter, A., Kopczuk, W., & Telle, K. (2014) *Social networks and tax avoidance: evidence from a well-defined Norwegian tax shelter*. Mimeo. University of Oslo, Columbia University, and Statistics Norway.

Andreoni, J., & Miller, J. (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica*, *70*(2), 737-753.

Andreoni, J., Erard, B. & Feinstein, J. (1998). Tax Compliance. *Journal of Economic Literature*, *36*(2), 818-860.

Ariel, B. (2012). Deterrence and moral persuasion effects on corporate tax compliance: findings from a randomized controlled trial. *Criminology, 50*(1), 27–69.

Armantier, O., & A. Boly (2011). *Can Corruption be Studied in the Lab?* (Working paper). University of Montreal.

Azfar, O., & Nelson, W. R. (2007). Transparency, wages, and the separation of powers: An experimental analysis of corruption. *Public Choice*, *130*(3-4), 471-493.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological bulletin*, *91*(1), 3.

Barr, A., Lindelow, M., & Serneels, P. (2009). Corruption in public service delivery: An experimental analysis. *Journal of Economic Behavior & Organization*, *72*(1), 225-239.

Barr, A., & Serra, D. (2009). The effects of externalities and framing on bribery in a petty corruption experiment. *Experimental Economics*, *12*(4), 488-503.

Barr, A., & Serra, D. (2010). Corruption and culture: An experimental analysis. *Journal of Public Economics*, *94*(11-12), 862-869.

Becker, G. S. (1968). Crime and punishment: An economic approach. In *The economic dimensions of crime* (pp. 13-68). Palgrave Macmillan.

Beekman, G., Bulte, E., & Nillesen, E. (2014). Corruption, investments and contributions to public goods: Experimental evidence from rural Liberia. *Journal of Public Economics*, *115*, 37-47.

Berninghaus, S. K., Haller, S., Krüger, T., Neumann, T., Schosser, S., & Vogt, B. (2013). Risk attitude, beliefs, and information in a Corruption Game–An experimental analysis. *Journal of Economic Psychology*, *34*, 46-60.

Besley, T., & McLaren, J. (1993). Taxes and bribery: the role of wage incentives. *The Economic Journal*, *103*(416), 119-141.

Biddle, N., Fels, K. M., & Sinning, M. (2018). Behavioral insights on business taxation: Evidence from two natural field experiments. *Journal of Behavioral and Experimental Finance*, *18*, 30-49.

Bird, R. M., Martinez-Vazquez, J., & Torgler, B. (2008). Tax effort in developing countries and high income countries: the impact of corruption, voice and accountability. *Economic Analysis and Policy, 38*(1), 55-71.

Blackwell, C. (2007). *A meta-analysis of tax compliance experiments* (Working Paper No. 0724). International Center for Public Policy, Andrew Young School of Policy Studies, Georgia State University.

Blumenthal, M., Christian, C., Slemrod, J., & Smith, M. G. (2001). Do normative appeals affect tax compliance? Evidence from a controlled experiment in Minnesota. *National Tax Journal*, 125-138.

Bott, K., Cappelen, A. W., & Sørensen, E. Ø. (2014). *You've got mail: a randomised field experiment on tax evasion* (Discussion Paper 26/2014). Department of Economics, Norwegian School of Economics.

Buckenmaier, J., Dimant, E., & Mittone, L (2018). Effects of institutional history and leniency on collusive corruption and tax evasion. *Journal of Economic Behavior and Organization*. https://doi.org/10.1016/j.jebo.2018.04.004

Carrillo, P., Castro, E., & Scartascini, C. (2016). *Do rewards work to maintain and increase tax compliance? Evidence from the randomization of public goods*. World Bank ABCDE 2016 Conference, World Bank, Washington, DC.

Castro, L., & Scartascini, C. (2013). Tax compliance and enforcement in the Pampas: evidence from a field experiment. *Journal of Economic Behavior and Organization, 116*, 65-82.

Chander, P., & Wilde, L. (1992). Corruption in tax administration. *Journal of Public Economics*, *49*(3), 333-349.

Choi, J. P., & Thum, M. (2005). Corruption and the shadow economy. *International Economic Review*, *46*(3), 817-836.

Clotfelter, C. T. (1983). Tax evasion and tax rates: An analysis of individual returns. *The Review of Economics and Statistics*, 363-373.

Dell'Anno, R. & Schneider, F. (2004). *The Shadow Economy of Italy and other OECD Countries: What Do We Know?* (Discussion Paper). Department of Economics, University of Linz.

Dhami, S., & Al-Nowaihi, A. (2007). Why do people pay taxes? Prospect theory versus expected utility theory. *Journal of Economic Behavior & Organization*, *64*(1), 171-192.

Doerrenberg, P., & Schmitz, J. (2015). *Tax compliance and information provision: a field experiment with small firms* (IZA Discussion Papers 9013). Institute for the Study of Labor (IZA).

Drago, F., Mengel, F., & Traxler, C. (2015). *Compliance behavior in networks: evidence from a field experiment* (IZA Discussion Paper 9443). Institute for the Study of Labor (IZA).

Dunning, T., Monestier, F., Pi neiro, R., Rosenblatt, F., & Tu n'on, G. (2016). *Is paying taxes habit forming? Experimental evidence from Uruguay*. Mimeo.

Faruq, H., Webb, M., & Yi, D. (2013). Corruption, Bureaucracy and Firm Productivity in Africa. *Review of Development Economics*, *17*(1), 117-129.

Feige, E. L. (1994). The underground economy and the currency enigma. *Public Finance= Finances publiques*, *49* (Supplement), 119-136.

Fernández-Leiceaga, X., Lago-Peñas, S., Martínez-Vázquez, J., & Vaquero-García, A. (2018). Economía sumergida y fraude fiscal en España: ¿Qué sabemos? Una panorámica de la literatura. En S. Lago Peñas (Ed.), *Economía sumergida y fraude fiscal en España: ¿Qué sabemos? ¿Qué podemos hacer?* (pp. 17-41). Funcas-Estudios de la Fundación.

Ferraz, C., & F. Finan (2011). Electoral Accountability and Corruption: Evidence from the Audits of Local Governments. *American Economic Review*, *101*(4), 1274-1311.

Fisman, R., & J. Svensson (2007). Are corruption and taxation really harmful to growth? Firm level evidence. *Journal of Development Economics*, *83*(1), 63-75.

Fisman, R., & S. J. Wei (2004). Tax Rates and Tax Evasion: Evidence from Missing Imports in China. *Journal of Political Economy*, *112*(2), 471-500.

Fjeldstad, O. H. (2005). Corruption in tax administration: lessons from institutional reforms in Uganda. In S. Rose-Ackerman (Ed.), *International handbook on the economics of corruption* (pp. 484-511). Edward Elgar Publishing.

Gupta, S., et al. (1998). *Does Corruption Affect Income Inequality and Poverty* (Working Paper No. 98/76). International Monetary Fund.

Habib, M., & Zurawicki, L. (2002). Corruption and foreign direct investment. *Journal of International Business Studies*, *33*(2), 291-307.

Hallsworth, M. et al. (2017) The behavioralist as tax collector: using natural field experiments to enhance tax compliance. *Journal of Public Economics, 148*, 14–31.

Hashimzade, N., Myles, G. D., & Tran-Nam, B. (2013). Applications of behavioural economics to tax evasion. *Journal of Economic Surveys*, *27*(5), 941-977.

Hauk, E., & Saez-Marti, M. (2002). On the cultural transmission of corruption. *Journal of Economic theory*, *107*(2), 311-335.

Hindriks, J., Keen, M., & Muthoo, A. (1999). Corruption, extortion and evasion. *Journal of Public Economics*, *74*(3), 395-430.

Huttington, S. P. (1968). *Political order in changing societies*. Yale University Press.

Iyer, G. S., Reckers, P. M., & Sanders, D. L. (2010). Increasing tax compliance in Washington state: A field experiment. *National Tax Journal*, *63*(1), 7.

Justesen, M. K., & Bjørnskov, C. (2014). Exploiting the poor: Bureaucratic corruption and poverty in Africa. *World Development*, *58*, 106-115.

Kahneman, D., Knetsch, J. L., & Thaler, R. (1986). Fairness as a constraint on profit seeking: Entitlements in the market. *The American Economic Review*, 728-741.

Kastlunger, B., Lozza, E., Kirchler, E., & Schabmann, A. (2013). Powerful authorities and trusting citizens: The Slippery Slope Framework and tax compliance in Italy. *Journal of Economic Psychology*, *34*, 36-45.

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.

Kogler, C., Batrancea, L., Nichita, A., Pantya, J., Belianin, A., & Kirchler, E. (2013). Trust and power as determinants of tax compliance: Testing the assumptions of the slippery slope framework in Austria, Hungary, Romania and Russia. *Journal of Economic Psychology*, *34*, 169-180.

Lago-Peñas, I., & Lago-Peñas, S. (2010). The determinants of tax morale in comparative perspective: Evidence from European countries. *European Journal of Political economy*, *26*(4), 441-453.

Leff N. H. (1964). Economic development through bureaucratic corruption. *American Behavioral Scientist*, *8*, 8–14.

Manoli, D. S. & Turner, N. (2014). *Nudges and learning: evidence from informational interventions for low income taxpayers* (Working Paper No. 20718). National Bureau of Economic Research.

Mazur, M. J., Plumley, A. H., & Plumpley, A. H. (2007). Understanding the tax gap. *National Tax Journal*, 569-576.

Meon, P., & G. Sekkat (2005). Does corruption grease or sand the wheels of growth? *Public Choice,* 122, 69-97

Mossin, J. (1968). Taxation and risk-taking: an expected utility approach. *Economica*, *35*(137), 74-82.

Olken, B. (2006). Corruption and the cost of redistribution. Micro evidence from Indonesia. *Journal of Public Economics*, *90*, 853-870.

Olken, B. A., & Barron, P. (2009). The simple economics of extortion: evidence from trucking in Aceh. *Journal of Political Economy*, *117*(3), 417-452.

Ortega, D., & Scartascini, C. (2016). *Don't blame the Messenger: a field experiment on delivery methods for increasing tax compliance*. Mimeo. Inter-American Development Bank

Pickhardt, M., & Prinz, A. (2014). Behavioral dynamics of tax evasion–A survey. *Journal of Economic Psychology*, *40*, 1-19.

Reinikka, R., & Svensson, J. (2004). Local capture: evidence from a central government transfer program in Uganda. *The Quarterly Journal of Economics*, *119*(2), 679-705.

Ryvkin, D., & D. Serra (2012). How corruptible are you? Bribery under uncertainty. *Journal of Economic Behavior and Organization, 81,* 466-477.

Sanyal, A., Gang, I. N., & Goswami, O. (2000). Corruption, tax evasion and the Laffer curve. *Public Choice*, *105*(1-2), 61-78.

Schickora, J. T. (2011). *Bringing good and bad whistle-blowers to the lab* (No. 2011-4). Munich discussion paper.

Schneider, F. (2015). Schattenwirtschaft und Schattenarbeitsmarkt: Die Entwicklungen der vergangenen 20 Jahre, *Perspektiven der Wirtschaftspolitik 16*(1), 3-25.

Schulze, G., & B. Frank (2003). Deterrence versus Intrinsic motivation: Experimental evidence on the determinants of corruptibility. *Economics of Governance, 4,* 143-160.

Sequeira, S. (2012). Advances in Measuring Corruption in the Field. In D. Serra and L. Wantchekon (Eds.), *New Advances in Experimental Research on Corruption. Research in Experimental Economics Volume 15*. Emerald Group Publishing.

Serra, D. (2011). Combining top-down and bottom-up accountability: Evidence from a bribery game. *Journal of Law, Economics and Organization*, *28*, 569-587.

Slemrod, J. (2007). Cheating ourselves: The economics of tax evasion. *Journal of Economic Perspectives*, *21*(1), 25-48.

Slemrod, J., Blumenthal, M., & Christian, C. (2001). Taxpayer response to an increased probability of audit: evidence from a controlled experiment in Minnesota. *Journal of Public Economics*, *79*(3), 455-483.

Tanzi, V. (1998). Corruption around the world: Causes, consequences, scope, and cures. *Staff Papers*, *45*(4), 559-594.

Timmons, J. F., & Garfias, F. (2015). Revealed corruption, taxation, and fiscal accountability: Evidence from Brazil. *World Development*, *70*, 13-27.

Torgler, B. (2004). Moral suasion: An alternative tax policy strategy? Evidence from a controlled field experiment in Switzerland. *Economics of Governance*, *5*(3), 235-253.

Torgler, B. (2013). A field experiment in moral suasion and tax compliance focusing on underdeclaration and overdeduction. *FinanzArchiv: Public Finance Analysis*, *69*(4), 393-411.

Traxler, C. (2010). Social norms and conditional cooperative taxpayers. *European Journal of Political Economy*, *26*(1), 89-103.

Tulkens, H., & Jacquemin, A. (1971). *The cost of delinquency: a problem of optimal allocation of private and public expenditure*. Center for Operations Research & Econometrics.

Uslaner, E. M. (2010). Tax evasion, corruption, and the social contract transition. In J. Alm, J. Martinez-Vazquez, J. and B. Torgler (Eds.), *Tax compliance and tax evasion*. Routledge.

van Veldhuizen, R. (2011). *Bribery and the Fair Salary Hypothesis in the Lab*. Mimeo. University of Amsterdam

Vellutini, C., Casamatta, G., Bousquet, L. & Poniatowski, G. (2019). *Estimating International Tax Evasion by Individuals* (Working Paper No. 76-2019. Taxation papers, European Commission. https://ec.europa.eu/taxation_customs/sites/taxation/files/2019-taxation-papers-76.pdf

Yitzhaki S. (1974). A note on Income tax evasion: A theoretical analysis. *Journal of Public Economics*, *3*(2), 201-202.

CUADERNOS ECONÓMICOS ICE INFORMACIÓN COMERCIAL ESPAÑOLA

# Economic Models of Optimism: What Does the Evidence Say?*

# *Modelos económicos del optimismo: ¿qué muestra la evidencia empírica?*

Adrián Caballero
Universidad Autónoma de Madrid

Raúl López-Pérez
Instituto de Políticas y Bienes Públicos, CSIC**

***Abstract***

*Recent economic models define a (non-Bayesian) optimist as someone whose beliefs and expectations are typically "too rosy", in the sense that her priors and the evidence objectively available statistically warrant a more negative outlook. This paper reviews the existing empirical literature on optimism to assess the empirical relevance of those models. While there exists abundant and compelling evidence in favor of motivated inference, i.e., that preferences shape beliefs, the support for the most specific predictions of the economic models seems mixed, if not negative. We discuss open questions and opportunities for future research.*

***Keywords:*** *biases, inference, motivated beliefs, optimism, overconfidence, self-serving bias.*
***JEL Classification:*** *D01, D80, D83, D84, D91.*

***Resumen***

*En recientes modelos económicos, un optimista (no bayesiano) es aquel agente cuyas creencias y expectativas están frecuentemente «tintadas de rosa», dado que sus a prioris y la evidencia objetivamente disponible implicarían una predicción estadística menos positiva. Este artículo repasa la literatura empírica sobre optimismo con vistas a evaluar la relevancia empírica de tales modelos. Si bien hay evidencia abundante y convincente de que la inferencia motivada existe, esto es, que las preferencias moldean las creencias, las predicciones más específicas de los modelos parecen tener apoyo limitado, si no nulo. Concluimos con una discusión de algunas cuestiones por resolver, así como de oportunidades de cara a futuras investigaciones.*

***Palabras clave:*** *sesgos, inferencia, creencias motivadas, optimismo, exceso de confianza, sesgo por interés personal.*
***Clasificación JEL:*** *D01, D80, D83, D84, D91.*

* The authors gratefully acknowledge financial support from the Spanish Ministry of Economics, Industry and Competitiveness through the research project ECO2017-82449-P, and helpful research assistance by Roberto Lucas and Sergio Rubio.

** Institute of Public Goods and Policies (IPP), Spanish National Research Council (CSIC).


DOI: https://doi.org/10.32796/cice.2020.99.7008
Recibido: Enero 2020 · Aceptado: Marzo 2020

## 1. Introduction

People often have beliefs or expectations about many things: their material future, their abilities or personal characteristics, how well they perform relative to others, the future price of some asset, the world outlook, social causality, the lethality of some virus, etc. In this respect, abundant evidence from Psychology and Economics hints the existence of a positivity bias in people's inferences, a bias also called motivated inference, optimism, self-deception, wishful thinking, or asymmetric updating (Wicklund & Brehm, 1976; Kunda, 1990; Baumeister, 1998; Bénabou & Tirole, 2016; Epley & Gilovich, 2016). When people have some motivation to hold a certain belief, that is, updating often proceeds as if the good, confirmatory evidence received a relatively larger weight than the bad, negative news. This is at odds with Bayesian learning, e.g., Samuelson, 2004, the standard approach to inference in rational choice theory, which requires not to treat a piece of information differently conditional on whether we like it or not.[1]

Understanding how and under what circumstances people infer in a motivated manner is an obviously important question for policy reasons. In effect, this bias has a potential negative side: many decisions like buying some good, exerting effort to get some future reward, investing on some asset, saving, or taking care of our health can be made sub-optimally if people have inflated beliefs or expectations. Kindleberger and Aliber (1978), for example, argue how irrationally optimistic expectations lead to overinvestment during economic booms, and Shiller (2000) includes optimism as one of the biases that can boost the confidence of investors, leading to large increases in asset prices and favoring the formation of financial bubbles. For another example, the survey by Case and Shiller (2003) to new homeowners revealed very optimistic expectations regarding future housing prices. In fact, 90 % of the surveyed believed that housing prices in their cities would keep increasing for the next 10 years, at an average estimated rate of between 9 and 15 %, depending on the city. An excess of optimism can also have detrimental consequences for workers. Using data from the American labor market, Spinnewijn (2015) shows that more than 80 % of the surveyed job seekers underestimated the time it would take to find a new job. Consequently, the job seekers devote little effort to search for a new job, save little for unemployment and consume these savings too quickly while they are unemployed. Regarding health, the study by Oster et al. (2013) with individuals at risk for Huntington disease, a degenerative neurological disorder associated to

[1] To prevent confusion, we stress that optimism is not the only bias which psychologists and other researchers have documented. Specifically, people may process information in a statistically "incorrect" manner for non-motivational reasons, e.g., not attending to some signal or forgetting it due to the mere effect of time. The literature about these "cold" biases is inextricably associated to the work by Daniel Kahneman and Amos Tversky, who developed the idea, introduced by Herbert Simon, that humans reason and solve problems by means of heuristics, i.e., intuitive rules –Simon, 1956; Tversky and Kahneman (1973, 1974); Kahneman and Tversky 2000. These rules, examples of which include availability and representativeness, are relatively efficient because they are easy to apply, but at the same time lead to systematic mistakes in our inferences. Benjamin (2019) extensively surveys the evidence on "cold" inference.

a genetic alteration, reveal that many of them underestimate their true risk, reject being tested and behave as if they do not have the disease, even when diagnosed people behave quite differently regarding decisions. For example, individuals that are certain that they have the genetic alteration are more likely to get married, to retire and to change their recreational activities and financial decisions than those who know not to carry the alteration. Interestingly, the behavior of the untested individuals mimics the behavior of the latter.[2]

A literature in Economics has proposed different models of motivated inference –e.g., Akerlof and Dickens (1982), Rabin (1994), Bénabou and Tirole (2002), Brunnermeier and Parker (2005), Möbius et al. (2014). In a nutshell, these models of non-Bayesian optimism (NBO henceforth) propose that motivated inference presents analogies with rational choice: people tend to hold the beliefs that maximize their utility, i.e., the beliefs they would hold if they could freely choose them, taking into account their costs and benefits. The NBO models differ mainly in the constraints that they posit people face when "choosing" their beliefs. First, papers like Akerlof and Dickens (1982) and Brunnermeier and Parker (2005) assume no restrictions: the wish begets the belief. Specifically, the optimal expectations conveniently balance the benefits from keeping pleasant beliefs and the advantages of belief accuracy in terms of better decision making. A second group of models posit the existence of cognitive costs of holding inaccurate beliefs as well as costs associated to information avoidance, e.g., Rabin (1994) and Bénabou and Tirole (2002). For example, Rabin (1994, p. 180) argues that departing from the "natural, intellectually honest set of beliefs" may require a lot of intellectual effort. It could imply also some opportunity costs in terms of refraining from pleasant activities in order to avoid challenging information. In addition, Bénabou and Tirole (2002) and Bénabou (2015) introduce sophisticated agents who somehow anticipate the existence of a positivity bias in their inferences and can hence limit the extent of such bias.

Since economists and psychologists have gathered a substantial amount of evidence on motivated inference, a natural question arises: how well do the NBO theories explain that evidence? An exploration of the limits of the models is essential to eventually improve them, if necessary, and to better understand the phenomenon. With this in mind, this paper presents several predictions of the NBO models and compares them with relevant evidence, incidentally offering a (inevitably partial) review of the empirical literature. In contrast with previous surveys on motivated inference, therefore, the theoretical predictions organize the review, which leads to the

[2] While the prior discussion stresses the downsides of optimism, the bias seems to have a positive side as well. For instance, convenient ego-beliefs or beliefs about the future can bolster our self-esteem, confidence, and hope (think of religion), hence giving us peace of mind and serenity. In addition, Festinger (1957) proposes that choosing an option that is good in some respect but bad in another one triggers a painful cognitive dissonance. This can be prevented if beliefs are tuned accordingly, e.g., believing that the payoff-maximizing, self-interested choice is also the fair one. Consider also a third, more indirect example: Sometimes having strong beliefs on some issue motivates us to be bolder, taking more risks or undertaking difficult tasks; determination can be helpful as well if we want to convince others to join some collective project (e.g., von Hippel & Trivers, 2011).

analysis in detail of issues like, say, the interaction between accuracy and monetary incentives. As we will argue, NBO models (as any simplification of reality) face some issues, and the key one seems to be an inaccurate account of the restrictions or limits of motivated inference. In effect, the models predict that having an interest in believing that some event is true is sufficient to inflate such belief, provided that it is not "too" inaccurate (Rabin, 1994) or that the person is sufficiently naïve (Bénabou & Tirole, 2002). This contrasts with evidence showing that, as we will see, people are sometimes surprisingly accurate even if they have an interest in holding some biased belief, and in some cases hold non-Bayesian pessimistic beliefs, i.e., assign a "too" low probability to the most desirable events. In summary, economic models have mixed support, and their predictions are quite often hardly supported or even challenged by the experimental evidence.

The rest of the paper is organized as follows. The next section presents for illustrative purposes a toy model of NBO and discusses its empirical relevance. Section 3 concludes by mentioning open questions and potential future venues of research.

## 2. NBO: a toy model and a review of the evidence

The goal of this section is to evaluate the empirical performance of the NBO models, taking into account some of the literature on motivated inference. To organize the discussion, we first introduce a toy model of NBO, building on the theoretical framework from Akerlof and Dickens (1982) and Brunnermeier and Parker (2005). From such model we derive a series of implications. We then analyze whether the existing evidence supports each implication, based on our reading of the literature.

### *2.1. A simple framework*

Let $S = \{\text{blue, red}\}$ denote the state space, where any state $s$ fully specifies all relevant features of the decider's environment. We describe the uncertainty involved by introducing the *objective* probability $\pi$ that the state is red. The decider's objective is to maximize a weighted sum of the expected utility of her material payoff plus the utility of holding the beliefs characterized by $\hat{\pi} \in [0, 1]$, which is interpreted as the decider's subjective degree of confidence that the red state is true. Implicitly, the red state is a positive one in which, say, the decider (i) is successful in some task, (ii) has some "good" personal quality, e.g., physical attractiveness, intelligence, honesty, etc., or (iii) gets some material reward. We define the decider's utility as:

$$U(a, s, \hat{\pi}) = u(a, s) + v(\hat{\pi}) \qquad [1]$$

where $a$ is the decider's action. The first term represents the utility associated to the

material payoff, which is possibly determined by the state and the specific action chosen by the decider. The second term captures the utility derived directly from holding beliefs $\hat{\pi}$. Without loss of generality, we assume that $dv/(d\hat{\pi}) > 0$, so that believing that the actual state is red provides some utility. In Akerlof and Dickens (1982), for instance, a belief that some job is not hazardous can be relieving, leading to less stress and fear. Alternatively, an individual may feel comfortable when thinking that she will not suffer a health problem in the future or that the value of her assets will always increase.

In line with Akerlof and Dickens (1982), there are two periods. In the first one, the decider is aware of the objective probability $\pi$ and "chooses" her beliefs $\hat{\pi}$. In the second period, she chooses an action based on her beliefs $\hat{\pi}$, and gets some pecuniary and non-pecuniary benefits, possibly including some utility $v(\hat{\pi})$ from beliefs $\hat{\pi}$. In general, we say that the decider is biased if $\hat{\pi} \neq \pi$, inflating (deflating) her beliefs if $\hat{\pi} > \pi$ ($\hat{\pi} < \pi$); we also employ the term "optimism" ("pessimism") as a synonym of inflation (deflation). We assume that the decider in the second period ignores the possibility that she may have strategically chosen inflated beliefs in the first period. Thus, in the second period, her optimal action is:

$$a^*(\hat{\pi}) = \arg \max_a \hat{\pi} u(a, \text{red}) + (1 - \hat{\pi}) u(a, \text{blue}) + v(\hat{\pi}) \quad [2]$$

Importantly, the decider anticipates in period 1 that her choice of beliefs $\hat{\pi}$ will possibly affect her action in period 2, perhaps leading to suboptimal behavior from the point of view of an accurate observer. In Akerlof and Dickens (1982), for instance, a worker that underestimates the risk associated to some job may feel less anxious, but at the same time he is more likely to suffer an accident. In choosing $\hat{\pi}$, therefore, the decider will look for a compromise, trading off the utility from $\hat{\pi}$ with the cost imposed by future wrong actions. The decider's optimal belief is thus:

$$\hat{\pi}^*(\pi) = \arg \max_{\hat{\pi}} \pi u(a^*(\hat{\pi}), \text{red}) + (1 - \pi) u(a^*(\hat{\pi}), \text{blue}) + v(\hat{\pi}) \quad [3]$$

To repeat, note well that $\hat{\pi}$ determines the decider's expected utility at period 1 in two ways. First, directly, since keeping beliefs $\hat{\pi}$ provides utility $v(\hat{\pi})$. Second, indirectly, since beliefs $\hat{\pi}$ may affect the decider's future action. From this simple framework, let consider three possible scenarios. To start, if $v(\hat{\pi}) = 0$ for any $\hat{\pi}$, i.e. if beliefs do not provide utility by themselves, $\hat{\pi}^*$ is such that:

$$a^*(\hat{\pi}^*) = \arg \max_a \pi u(a, \text{red}) + (1 - \pi) u(a, \text{blue}) \quad [4]$$

Given our previous definition of $a^*$, it turns out that $\hat{\pi}^* = \pi$. That is, if beliefs are not a direct source of utility, the decider has no incentive to keep biased beliefs. More generally, this result applies if $v(\hat{\pi})$ is constant for any $\hat{\pi}$. In the second and third remaining scenarios, we assume otherwise. The second scenario is the simplest one to analyze and occurs when there are no material costs from choosing wrong,

inaccurate beliefs. More precisely, let $\hat{\pi}^b = \arg \max_{\hat{\pi}} v(\hat{\pi})$ denote the beliefs giving maximal "direct utility" within interval [0, 1], and $Eu(a^*(\hat{\pi}))$ denote the expected utility derived from the *material* payoff, given the optimal action $a^*$ for beliefs $\hat{\pi}$. If $Eu(a^*(\hat{\pi}^b)) = Eu(a^*(\pi))$, utility maximization can be conceived as a relatively simple two-step, backwards induction procedure: in period 1, the decider opts for the "most pleasing" beliefs $\hat{\pi}^b$, anticipating that she will later choose act $a^*(\hat{\pi}^b)$ in period 2. Since beliefs do not compromise future material payoffs, that is, she just opts for those beliefs that maximize $v(\hat{\pi})$.[3] Particular sub-cases within this second scenario include (i) $a^*$ maximizes $u(a, s)$ for any $s$, or (ii) $u(a_i, s) = u(a_j, s)$ for any pair $i, j$.

The third scenario appears when some beliefs provide more "direct utility" $v(\hat{\pi})$ than others but the decider can suffer a material loss for holding inaccurate beliefs. Here the analysis of NBO becomes more complex. Given some additional assumptions, however, the toy model still predicts some degree of optimistic bias regarding $\pi$. Recall first that $dv/d\hat{\pi} > 0$, as we assumed before. Assuming in addition that $Eu(a^*(\hat{\pi}))$ is continuous and twice differentiable and that it has a unique maximum, the problem's first order condition for the decider in the first period is:

$$\frac{dEu(a^*(\hat{\pi}))}{d\hat{\pi}} + \frac{dv(\hat{\pi})}{d\hat{\pi}} = 0 \qquad [5]$$

Recall from the first scenario that the first term is maximized for $\hat{\pi} = \pi$. For those beliefs, it turns hence out that $dEu(a^*(\hat{\pi}))/d\hat{\pi} = 0$ and that $d^2Eu(a^*(\hat{\pi}))/d\hat{\pi}^2 < 0$. Since the second term in [5] is strictly positive, optimal beliefs $\hat{\pi}^*$ are such that $dEu(a^*(\hat{\pi}))/d\hat{\pi} < 0$, which implies that $\hat{\pi}^* > \pi$ as long as $\pi < 1$. In other words, the decider finds optimal to form optimistic beliefs, inflating $\hat{\pi}$ up to the point in which the marginal "direct utility" of those beliefs is cancelled out by the marginal decrease in the expected utility from the material payoff.

### *2.2. Predictions and evidence*

The first prediction of NBO is direct: a person's beliefs should co-move with her preferences. If the decider maximizes utility by believing that some state $s$ is the actual one, that is, she will tend to believe it. In contrast, she will be more critical towards that idea if it does not give utility.[4]

*Prediction 1 (P1):* Deciders inflate their beliefs. Specifically, the difference between the subjective and the objective probabilities is correlated with the decider's preferences over the beliefs about the state space.

---

[3] The discussion presumes for simplicity that both $\hat{\pi}^b$ and $a^*(\hat{\pi}^b)$ are unique. A multiplicity of equilibria exists otherwise.

[4] This can be contrasted with a perhaps more "rational" account in which people believe $S$ because they have (proper) reasons for that, e.g., evidence, and not because it is functional.

*Discussion and evidence:* For a well-known illustration of *P1,* consider a scenario where the decider's utility depends on her expected material payoff and some direct utility from beliefs. More precisely, $\pi$ represents here the belief that the decider is a fair and honest person; intuitively she gets "ego-utility" from believing that she is such a person. In this type of scenarios, it is often safe to assume that holding inaccurate beliefs is not risky, that is, entails no material loss. As we have indicated before, the NBO maximization problem is hence straightforward: When people have to make a normative choice, intuitively, they first conceive the act that best serves their self-interest and *then* pick the beliefs about what is normative to make that choice seem correct or justifiable, inferring as a result that they are fair-minded people –i.e., the "good" state of the world. More specifically, consider a scenario with two acts $a_{blue}$ and $a_{red}$ and two states, blue and red, i.e., such that $a_{blue}$ ($a_{red}$) is the right, normative act in the blue (red) state. Let $\pi$ denote the objective probability of the red state. If choice $a_{red}$ maximizes material utility, say, the prediction is that $\hat{\pi}^* = 1$, so that the decider can earn as much as possible and feel no guilt for acting in an unfair manner. If $1 > \pi$, e.g., if an impartial observer is not totally convinced that is the normative choice, a self-serving bias in fairness judgments is predicted.[5]

Apparently in line with *P1*, many studies on self-serving biases report behavior that can be rationalized by convenient beliefs about the fairness of the acts chosen.[6] Since fairness judgments can often be conditioned on a myriad of reasons, the studies reflect this richness. In some cases, for instance, people seem to have self-serving beliefs about the applicable moral rules. In Shu and Gino (2012), participants read an honor code and were then asked to conduct a task and report their performance –they were allowed to cheat by over-reporting, thus getting a higher payoff. Those participants who cheated tended to forget the moral rules from the honor code even when monetary incentives were provided. In contrast, they were equally likely to remember information unrelated with moral. In the control condition in Babcock et al. (1995), in turn, subjects are assigned the role of plaintiff or defendant and matched in pairs. All of them work with the same real case, in which the plaintiff demands a compensation of $100,000 after a traffic accident involving both parties. They are given real information regarding testimonies, police reports, and maps and informed that a real judge has been given the case materials to reach a decision about the compensation for the plaintiff. Each negotiator is then asked an incentivized, private estimate of the judge's decision, as well as an evaluation of what a fair settlement would be according to her own view. Afterwards, the negotiators bargain during 30 minutes over the amount that the defendant would pay to the plaintiff. If they do not reach an agreement over that time span, the plaintiff is awarded the compensation determined by the judge (for an amount unknown a priori by the

[5] These occur when, in the words of Babcock and Loewenstein (1997, p. 110), people "...conflate what is fair with what benefits oneself". Or as defined by Dahl and Ransom (1999, p. 703): "A self-serving bias occurs when individuals subconsciously alter their fundamental views about what is fair or right in a way that benefits their interest."

[6] In these studies, subjects' beliefs are often not elicited, though.

negotiators). In one treatment, subjects are informed of their roles only after reading the materials and estimating the judge's and fair settlements (but prior to negotiation). In line with *P1*, subjects in this veil-of-ignorance treatment are significantly more likely to settle and do so in less time. Further, they are significantly more likely to agree on the judge's and fair settlements and less likely to differ in the importance they attach to self-serving arguments.

In other occasions, a self-serving bias is manifested in the agents' selection of comparison groups, a point often relevant in fairness judgments. In a field study examining actual teacher contract negotiations in 500 school districts in Pennsylvania, indeed, Babcock et al. (1996) observe that unions tend to choose comparables (in other school districts) with higher salaries than those that school boards select. This has consequences, as the higher the difference between the proposed comparables, the higher the strike rate and hence the negotiation impasses.

In some studies, people apparently choose convenient beliefs about how deserving another person is. Di Tella et al. (2015) conduct a modified version of the dictator game, in which the receiver secretly decides the monetary value of the tokens allocated by the decider and receives an additional payoff if the lower value is chosen. Simultaneously, the decider chooses the allocation of 20 tokens, although the maximum number of tokens that deciders can keep for themselves varies among players (this information is known by the decider only). The main result was that those deciders who were allowed to keep a larger proportion of the tokens not only behaved more selfishly, but they were also more likely to believe that the receiver had behaved selfishly too, even when realistic beliefs were incentivized in some treatments. That is, deciders apparently form 'excusing' beliefs about receivers' behavior.[7] In Barkan et al. (2012), for another example in this vein, participants were presented a hypothetical situation in which they work in the HR department of a firm and must decide whether to hire a candidate whose ethical behavior is depicted as questionable but advantageous for the firm. In one of the treatments, participants were asked in a previous task to describe an unethical thing they had done. Those participants were significantly less likely to hire the candidate and judged him as potentially less loyal to the firm and less honest than the participants in the control treatment (which had not been asked previously to recall their own past misconduct). It seems therefore that people recalling their past misconduct tend to reduce ethical dissonance by judging others' moral behavior even more harshly.

The previous studies strongly suggest that self-serving biases in fairness judgments exist; see also Gino et al. (2016) for a review of additional evidence. It must be noted however that not all of the existing evidence in this realm is in line with *P1*. For instance, Dahl and Ransom (1999) find very limited evidence of a bias among Mormons in their own determination of what constitutes income for the purpose of tithing –e.g., having received a sizable gift or inheritance does not affect individuals' views of whether gifts and inheritances should be tithed.

[7] Yet Ging-Jehli et al. (2019) find no evidence of such kind of "strategic cynicism".

Honesty is not the only trait of intrinsic importance to the individual. Self-perceptions of one's intelligence and beauty, for instance, seem to affect ego-utility as well. In this sense, Moore and Healy (2008) distinguish three related types of overconfidence that have been commonly confounded. Two of them are particularly relevant in our context.[8] In first place, NBO models predict *overestimation* of one's abilities and inflated self-perceptions in these regards, which fits well with the results of different calibration studies (Lichtenstein & Fischhoff, 1977; Budescu et al., 1997, Brenner et al., 1996). The structure of these studies is usually quite similar: the participants are presented a series of multiple-choice questions. In addition to choosing the right answer, they must estimate the probability that their answer is correct. If a participant's mean estimation is p, overestimation occurs when the percentage of actually correct answers is lower than 100·p. Interestingly, the evidence suggests that this phenomenon is not universal and that it is conditioned to a large extent by the context (Gigerenzer et al., 1991). For instance, the degree of overestimation seems to increase with the difficulty of the task, which has been labeled the "hard-easy effect" –e.g., Brenner, 2003; Larrick et al., 2007.

As a second type of overconfidence in line with NBO, we have the phenomenon generally known as *over-placement* or *better-than-average effect:* when evaluating their own capacities and abilities in relative terms to others, a growing body of work from social psychology and economics suggests that people are often biased –consult Alicke and Govorun (2005) for a review. In Svenson (1981), for example, 93 % of U.S. students believe that they are both safer and better drivers than the median one. In the same vein, Cross (1977) find than more than 65 % of college teachers rank themselves in the top 25 % when asked about their teaching abilities relative to other professors.[9] Note yet that the evidence is not conclusive; for instance, the experiment by Clark and Friesen (2009), in which the participants face real effort tasks and must infer their absolute and relative performance, finds little evidence of either over-placement or over-estimation. In this vein, one feature that seems to be at odds with NBO is that over-placement seems particularly true for skills or tasks that are relatively easy to master. In this line, a survey conducted by Kruger (1999) reported significant above-average effects in skills like driving, riding a bicycle or saving money, but that people systematically underestimated their ability relative to others in more complex tasks like playing chess or programming a computer. Finally, two interesting aspects of over-placement are that (i) people tend to see themselves as less biased than others (Pronin et al., 2004) and that (ii) it makes people more effective in persuading or deceiving others. In line with (ii), subjects in Schwardmann and van der Weele (2019) first perform an intelligence test. Then

[8] Moore and Healy (2008) also discuss over-precision, that is, statistically unwarranted overconfidence on the truthfulness of one's inferences, typically exemplified by "too" narrow confidence intervals.

[9] As Benoit and Dubra (2011) argue, the evidence from some earlier social psychology studies on over-placement is not unequivocally at odds with Bayesian updating. Nevertheless, more recent studies in Economics like Eil and Rao (2011) and Möbius et al. (2014) eliminate these confounds; we review them below.

a random half of them are informed that they can earn a prize by convincing others of their superior performance in the test. It turns out that the privately elicited and incentivized beliefs of these subjects are significantly more confident than the beliefs of subjects in the control condition. Further, the more confident participants state a higher confidence during the face-to-face persuasion stage, and are significantly more likely to be rated as successful by their evaluators.

We turn now to another body of literature in line with NBO, this one suggesting that people tend to be optimistic about their future. In the survey conducted by Weinstein (1980), for example, people underestimate the chances of getting divorced, losing their jobs or suffering different health problems. In contrast, they overestimate the probability of living past eighty, owning a house or finding a good job. Far from anecdotical, the estimates suggest that about 80 % of the population show this kind of bias; Sharot, 2011. In some studies, for another example, people have inflated prospects about their material prospects. In the experiment in Mayraz (2011), subjects are shown a chart of historical wheat prices and have to predict afterwards the price at some future time point. This is iterated 12 periods, one of which is randomly selected afterwards for compensation; the size of the bonus depends on the accuracy of the corresponding prediction. Additionally, any subject gets a payoff that increases (decreases) with the future price of wheat if her randomly-selected role is "Farmer" ("Baker"). If subjects get sufficient utility from being optimistic and truthfully report their predictions, NBO predicts that the average farmer should make significantly higher predictions than the average baker, as it is indeed observed. Bayesian learning, in contrast, predicts no difference, given random role allocation. In turn, Forsythe, Rietz and Ross (1999) conduct an experiment in which the participants trade with two assets, namely Blue and Green certificates. At the end of each period, Blue certificates liquidate for some monetary value with probability 0.3 (Blue state), while Green certificates expire worthless. The opposite occurs with probability 0.7 (Green state). Each sequence of three trading periods has the following structure: in the first period, no additional prize is provided. In the second (third) period, all participants get $2 if the Blue (Green) state occurs. This prize is independent of the decision made by the traders. On average, the authors report that the price of Blue (Green) certificates increased when an additional prize was associated to the Blue (Green) state, suggesting an inflated probabilistic belief that the most favorable state occurs. In contrast, average prices tended to decrease in the subsequent periods with no state-contingent prize.

We note yet that negative evidence on optimism about financial prospects is offered by Heger and Papageorge (2018), where subjects sequentially completed several tasks, each one containing a number of questions –at the end of the session, one task was randomly selected for payment. In the Baseline task, subjects face several urns, each one containing a different proportion of white and black balls, which is common knowledge. Subjects report their probabilistic belief that a draw from each urn will be white, being paid for accuracy according to a quadratic scoring rule. To explore over-estimation of the probability of high-payoff outcomes, participants in the Payment task receive a side payment if a white ball is drawn from the urn.

Clearly, optimism implies higher average reported beliefs in this task, where white becomes payoff-favorable. Nevertheless, the data does not support this prediction. On the other hand, the authors also explore a Performance task in which white balls are not payoff-favorable, but ego-related. Specifically, each subject answers an IQ question and, if her answer is correct, one white ball is added to the urn. In this task, average beliefs are upwardly biased, in line with the evidence cited above about over-estimation. Further, the authors find that subjects who over/under-estimate the probability of white in the Payment task tend also to over/under-estimate their own performance in the Performance task. ■

Let us now turn on more specific predictions by the NBO models. In this vein, an implicit or explicit mechanism behind most models of NBO is asymmetric updating: Signal observations are over-weighted or under-weighted depending on the decider's goals, i.e., the target, optimal beliefs. While Akerlof and Dickens (1982) and Brunnermeier and Parker (2005) do not formalize how agents learn in their models (as neither does our toy model), Möbius et al. (2014) does analyze this point formally, building on the framework from Brunnermeier and Parker (2005). Möbius et al. assume that agents optimally choose the weight they give to bad and good news, predicting that (i) agents will be conservative, i.e., they will respond less to the evidence than a Bayesian, but at the same time (ii) they will be more conservative in response to bad news than good news. Benabou and Tirole (2002) propose an NBO model where agents display motivated recall, so that they recall better the evidence that supports their preferred beliefs. In their model, specifically, an individual that prefers to believe that she is good at certain task may repress or forget those pieces of information that question her ability. This manipulation is limited to some extent by cognitive costs. Further, their model allows for some sophistication in the sense that the individual can be aware up to a certain degree that her recalled sample is likely to be biased, which consequently may limit –but not eliminate– the effects of motivated recall.

*Prediction 2 (P2; asymmetric updating/inference)*: Individuals treat as more informative the signals in line with their preferred beliefs than those that challenge them. Thus, individuals process "good" and "bad" news asymmetrically in order to reinforce their favorite beliefs.

*Discussion and evidence:* Neuroscientist and psychologists have recently gathered some supportive evidence for asymmetric updating. In Sharot et al. (2011), participants (N = 19) are presented some adverse event, such as being diagnosed with Alzheimer's disease or being robbed, and have 6 seconds to estimate their chances of facing that event in the future, not being paid for accuracy. Afterwards, subjects are shown in the screen during 2 seconds the actual frequency with which that event happens among individuals living in the same socio-cultural environment as them. This procedure is repeated for 80 events, randomly ordered. To assess whether participants use the information on the actual empirical frequency to update their predictions, the authors subsequently ask them to provide estimates of their likelihood of encountering the events considered before. They find evidence for

asymmetric updating in favor of good news.[10] Interestingly, Sharot et al. (2011) obtain fMRI (functional Magnetic Resonance Imaging) data while subjects proceed with the experiment. This data shows that activity in a region of the frontal cortex (right IFG) identified as sensitive to negative estimation errors predicts the extent to which participants selectively update their beliefs, (relatively) dismissing the information conveyed by the empirical frequency when it was bad news.

Economists have explored as well the relevance of asymmetric updating. An interesting aspect of the experiments cited below is that, in contrast to our knowledge to the psychological studies, both signals and priors are strictly controlled by the experimenter. This makes the Bayesian benchmark readily computable, something unfeasible in the experiment just cited above by Sharot et al., 2011, for instance. Until recently, this was not very common in the literature, particularly when dealing with beliefs about aspects that are of every day concern –health, intelligence, beauty, academic performance, driving ability, etc. While these prior studies were tremendously suggestive, they failed to clarify the existence of biases, as the knowledge that subjects bring to the studies was hardly verifiable.[11]

Recent studies also allow pinpointing some mechanisms through which these biases develop. In this regard, we must note that experimental attempts in Economics to test *P2,* i.e., asymmetric updating, have concluded with mixed results. While some works find a significant overweight of good news relative to bad news (Eil & Rao, 2011; Möbius et al., 2014; Charness & Dave, 2017), other scholars find no evidence of asymmetric updating (Grossman & Owens, 2012; Gotthard-Real, 2017; Buser et al., 2018; Barron, 2019) or even report the opposite phenomenon: that individuals overweight negative information relative to positive one (Ertac, 2011; Kuhnen, 2015; Coutts, 2019).

More specifically, Eil and Rao (2011) explore in two separate treatments how subjects update their beliefs about their ranking within a group of 10 people according to intelligence, as measured by score on an IQ test, or physical attractiveness, as rated by 5 subjects of the opposite sex after an "speed-dating" exercise. Subjects participated as well in a randomly-ordered control, where they updated their beliefs about their position in a "neutral" ranking, determined by a randomly assigned integer ranging from one to ten. In all treatments, subjects first revealed their prior probability distribution over the 10 possible ranks.[12] Afterwards, each subject received a series of three objective signals, i.e., whether he/she was ranked higher or lower than a randomly selected comparison subject, and had his/her (posterior) probabilistic beliefs elicited after each one. Accuracy was rewarded according to a

[10] See Wiswall and Zafar (2015) for similar results regarding college students' beliefs about their future earnings.

[11] For example, Kunda (1987) reports that, when given some evidence that caffeine drinking leads to negative health effects for women, female caffeine drinkers are less likely to rate it as valid than male ones. Yet this is not a clear signal that female drinkers are biased, as they could have gathered some disconfirming evidence in the course of their lives.

[12] These priors were non-incentivized. As the authors recognize, an issue here is that subjects might conceal their true subjective priors. For instance, they could reveal low probabilities for the top ranks, perhaps out of some sense of modesty. A Bayesian who acted in this manner would later under-respond to any negative signal.

quadratic scoring rule; consequently, a subject's expected payoff is maximized if she honestly reports her beliefs (assuming she is risk-neutral and understands the implications of the procedure). The authors find evidence for asymmetric updating in the IQ and beauty treatments (although the effect is significant only in the latter one), but not in the control. Further, agents' posteriors in the ego-relevant treatments are less predictable and less sensitive to signal strength after receiving negative feedback.

Ertac (2011) considers a relatively similar design with a coarser ranking, as subjects must report beliefs over which third of the distribution they occupy. Each subject observes one signal indicating whether she is in the highest rank (the bottom one in some sessions). Beliefs are elicited both ex ante an ex post, and compensated by means of a quadratic scoring rule. She finds asymmetric updating but in favor of the *unfavorable* feedback, thus leading to overly pessimistic beliefs. The main exception are those individuals who have relatively optimistic priors. Further, subjects are relatively less pessimistic about their relative performance in a math quiz than in a verbal task that they found generally harder than the former. Hence, these findings are in line with the evidence cited above that people are more likely to be overconfident in domains that they perceive to be easy (e.g. Kruger, 1999; Moore & Cain, 2007).

In the web-based study by Möbius et al. (2014), in turn, subjects are ranked according to their score in an IQ test. Participants report beliefs about which half of the distribution they occupy, measured both before and after the test and then again after receiving each of four noisy signals. At each update, subjects are reminded of the full history of signals. The belief elicitation method used was the so-called crossover method, which is incentive compatible even for risk-averse agents. The authors report evidence that updating is asymmetric but also conservative, in other words, people respond less to feedback (of any sign) than Bayes' rule prescribes. These findings are compatible with the NBO theory that they propose. Further, the biases are substantially less pronounced in a follow-up where the conditions are as closely as possible to those in the baseline, except that ego is not at stake, as subjects updated their beliefs about the performance of a "robot" rather than their own performance.[13]

A distinctive feature of Grossman and Owens (2012) is that they explore learning about *absolute* performance. Participants take an incentivized ten-question quiz on logic and reasoning and estimate the entire probability distribution of the 11 potential scores, i.e., 0 to 10, before and after taking the quiz, and after receiving one noisy signal. Participants were not given feedback on prior estimates. One of the three belief elicitations, randomly selected, was compensated with a quadratic scoring rule.

[13] Eil and Rao (2011) and Möbius (2014) check for asymmetric updating but also for the possibility of a confirmation bias. i.e., the tendency to place more weight on signals that confirm one's prior and underweight disconfirming signals; see Rabin and Schrag (1999) for a review. Neither of these studies finds evidence for such phenomenon. In Eil and Rao (2011), for example, subjects with negative priors also exhibited asymmetric updating, i.e., put relatively more weight on the positive, disconfirming signals.

Participants estimate their own score in treatment *Self* and the score of an anonymous participant in a separate condition called *Other*. This makes a difference with the studies cited before, where the controls feature events with objective probabilities, which may be interpreted very differently by subjects. Comparing a subject's expected score with her actual one, the authors find evidence of overconfident prior beliefs, particularly in the Self condition. However, non-Bayesian updating in the form of conservatism or asymmetric updating is not observed in the data, except for a minority of subjects. Taking into account the results from Eil and Rao (2011) and Möbius et al. (2014) cited before, the authors conjecture that asymmetric responsiveness to feedback might be more likely in beliefs about relative performance.

In the studies previously considered, the desired states are ego-relevant, e.g., a subject's self-esteem possibly depends on how well he/she performs in the IQ test relative to others, or in absolute terms. This is problematic in that people with different personal characteristics (e.g., intelligence) may update differently, and moreover receive different evidence –for example, a subject with a high IQ score is likely to receive good news, even if the signal is noisy. Under these circumstances, aggregate behavior may exhibit apparent asymmetries that are not present at the individual level; see Barron (2019) for a discussion of this point. In addition, one could wonder whether asymmetric updating extends to other domains. In this regard, several papers consider experiments in which subjects have a preference for some states simply because they give them a higher monetary reward. In Gotthard-Real (2017), there are two states (bad/good), priors are exogenously given and uniform, and a subject gets $7 if the good state occurs (and zero otherwise). As in Möbius et al. (2014), each subject received four noisy signals, always with feedback, and reported her belief about the probability of the good state after each signal. Further, a randomly selected belief was compensated using a quadratic scoring rule. To diminish any suspicion from the subjects, the random selection of an event was always done mechanically (contrary to other studies). The author finds no evidence that subjects inflate the probability of the good state, relative to a Bayesian.

In Coutts (2019), subjects update their probabilistic beliefs about four different binary events, randomly presented. One refers to the mean temperature on a randomly selected day in New York City, another one is ego-relevant, i.e., relative performance on a math and verbal skills quiz, and the other two are neutral, as they involve rolling dice. In each scenario, a subset of the participants had a 50 % chance of receiving $80 if some target event occurred; hence they had a financial interest in this respect. In each scenario, further, subjects first stated their subjective priors, then received three noisy signals, and reported her belief about the probability of the corresponding good state after each signal –one belief, randomly selected across all scenarios, was compensated using the crossover method, as in Möbius at al. (2014). Coutts (2019) reports that updating is conservative, with many non-updates, and asymmetric. However, it is *negative* signals that are over-weighted, not positive ones. These phenomena are observed both in the valenced and neutral contexts. Any bias, therefore, seems not to be of a motivated nature. Since the analysis further reveals

that, after receiving two signals, subjects overweight a second positive (negative) signal if the first one was positive (negative) signal, the author suggests the potential role of a confirmation bias. As we saw before, however, neither Eil and Rao (2011) nor Möbius et al. (2014) find evidence for such phenomenon.

Barron (2019) considers asymmetric treatments, in which subjects get a (small) bonus payment if one state of the world (out of two) is realized. Priors of the two states are exogenously fixed, subjects observe a sequence of five independent but noisy signals (feedback about prior signals is always provided), and state their beliefs about the probability of the "desired" state upon receipt of each signal (the quadratic scoring rule is used for compensation). Subjects repeat the exercise five times with different priors in each case, chosen in random order from the set {1/6, 2/6, 3/6, 4/6, 5/6}. Barron (2019) finds no evidence of asymmetric overweighting of good news or alternatively confirmation bias; indeed, average updating behavior is well approximated by Bayes' rule. This is also true at an individual level –this type of checking is relevant because, if some individuals are optimistic updaters and others pessimistic updaters, one could observe no asymmetry at the aggregate level if both groups were of a similar size and their bias of a similar magnitude.

In Charness and Dave (2017), there are two possible states, equally likely a priori. In one state, a container has seven black balls and three white balls, while these numbers are reversed in the other case. In each of 10 rounds, the actual state is first determined and then participants observe a sequence of six random draws from the container; after each draw, they state the posterior that each state has been realized, with a payoff for accuracy. At the end of each round, the true state is revealed and (in the Strategic condition) each subject is randomly matched with another one to play a $2 \times 2$ game; which payoff matrix is in force depends on which state has manifested. Subjects in the "Odd" role receive the same equilibrium payoff in both matrices, while those in the "Even" role receive a higher equilibrium payoff in one of the two matrices. The Even role, therefore, has an interest that one of the states occurs. On the other hand, in the "Non-strategic" treatment, each state is associated to a fixed payoff, with an identical ordinal structure for both roles as in the Strategic condition. The regression results indicate that Odd players learn differently than Even ones, as the latter give less weight to negative signals, i.e., those confirming the unfavorable state. To our knowledge, this seems to be the only economic study offering detailed evidence of asymmetric updating when subjects have a financial interest for some state.

In Buser et al. (2018), participants perform three tasks. The first one is announced to be designed so as to measure the individual's IQ (Raven's matrices). In the second, participants must form anagrams of different words. Subjects are informed that the task assesses their verbal ability. In the last task, they observe several matrices filled with numbers; in each one, they must select the two numbers that add up to 10. Participants are told that this third task measures their mathematical ability. In each task, groups of 8 participants are randomly selected. After completing a task, each participant is asked for her (incentivized) belief that she is in the top-performing

half in her group. Then, the participant observes a total of six informative signals about her relative performance, and her beliefs are elicited after each observation. At the aggregate level, Buser et al. (2018) find evidence for conservatism but not for asymmetric updating. Further, individuals tend to be more conservative in tasks that they perceive as more ego-relevant. While an individual's conservatism is correlated across tasks, finally, asymmetric updating was less stable. ■

We move to a slightly different issue. The framework model presented in *Section 2.1* captures in a simple way the reasons why an individual may prefer keeping biased, optimistic beliefs rather than realistic expectations. The idea of asymmetric updating, further, characterizes learning by optimistic agents. However, the question of which specific cognitive mechanism generates the optimistic beliefs remains unanswered. In this sense, Bénabou and Tirole (2012) discuss three non-exclusive alternatives. The first one assumes that, through evolution, a cognitive bias has persisted so that humans systematically overweight favorable information and underweight adverse one. The authors note, however, that this hypothesis is limited as long as people are not always overoptimistic and, in fact, some of them can be quite pessimistic. Second, people may prefer to ignore new information if that could challenge their favored beliefs. Thus, they could attend only those sources of information that they expect to reinforce their preferred beliefs. Finally, selective memory (denial, repression, etc.) could help manage those signals that are unavoidable and that may challenge the individual's preferred beliefs. This latter possibility leads naturally to the next prediction.

*Prediction 3 (P3; Bénabou and Tirole, 2012):* Asymmetric updating operates through biased recall and selective attention.

*Discussion and evidence:* there is some suggestive evidence in line with *P3,* but we believe that it is hardly conclusive. For instance, Thompson and Loewenstein (1992) explore labor negotiations, and find that subjects representing opposite sides later remember, from the same case file, more facts favoring their position than going the other way. The more divergent their recalls, moreover, the longer and costlier is the delay to agreement in the bargaining phase –see also Loewenstein et al., 1993. Recall also the evidence from Shu and Gino (2012) cited in the discussion of *P1,* where those participants who cheated tended to forget the moral rules from the honor code. In Chew et al. (2019), in turn, subjects (N = 1143) answer in an incentivized manner four questions from a Raven's IQ test. Some months later they are shown the same four questions, plus two which are new but similar, together with all the correct answers, and are asked to recall for each of the six questions whether they (a) answered it correctly, (b) incorrectly, (c) never saw it, or (d) just do not remember. For each question, subjects receive a prize for being accurate, suffer a loss if they recall incorrectly, and get nothing if they cannot remember. Subjects often make errors, which show some systematic patterns. First, the error by which one inaccurately "remembers" having correctly answered one question is significantly more likely than the reverse error, that is, a wrong memory of an incorrect answer. Second, there

is significant evidence of positive delusion. i.e., fabricating a positive event that did not actually happen. In effect, subjects had never seen questions 5 and 6, but more than 56% of them "remembered" answering any of them correctly, versus less than 6% incorrectly. Third, positive amnesia, i.e., forgetting a past negative event, was also very significant. In effect, the probability of not remembering one's answer, or whether one saw a question, is on average twice as high for those individuals who did a question incorrectly, relative to those who answered it correctly.[14]

On the negative side, the above cited study by Sharot et al. (2011) reports that asymmetric updating in their data cannot be explained by differential memory for desirable compared to undesirable information. In effect, after the scanning session, participants had to recall the (previously presented) actual probability of each of the 80 events occurring to an average person in the same socio-cultural environment. The errors thus committed did not depend on whether the actual probability was better or worse than initially expected by the participants, i.e., whether it was bad or good news. In most of the studies discussed in relation to *P2,* finally, subjects receive feedback so that biased recall should play no role. Still, some of them, e.g., Eil and Rao (2011) and Möbius (2014), find a positivity bias. This is perhaps a signal that NBO does not require that subjects "forget" or "misinterpret" signals. ■

For the following *Prediction 4,* consider two scenarios S1 and S2. In S1, there is a risk of a substantial material loss for having mistaken beliefs. In S2, the expected loss is small or there is no such risk, e.g., in a lab scenario, belief elicitation is incentivized in S1 but not in S2. Economic models of NBO predict that people will infer in a more Bayesian manner in S1, so as to reduce mistakes.[15]

*Prediction 4 (P4):* The correlation between beliefs and preferences will be weaker if the expected material loss for holding inaccurate beliefs is high.

*Discussion and evidence:* several of the papers cited above have systematically analyzed whether the size of the expected loss reduces the degree of NBO. Coutts (2019) runs sessions with different accuracy payments, i.e., low ($3), medium ($10), or high ($20). In addition, participants can either get a nil or high prize ($80) if some target event E occurs. According to *P4,* subjects have no incentive to distort their beliefs about the probability of E in case the prize is $0, provided that they get no other utility from the occurrence of E, e.g., if E is not ego-relevant. In contrast, distortion should be maximal if the prize is high and the accuracy payment low, i.e., $3. In spite of this, the author concludes that neither prizes nor accuracy payments alter updating. Mayraz (2011) varies the size of the accuracy bonus from £1 to £5 and reports that

[14] This is highly suggestive, but is it asymmetric updating? Maybe people use the following heuristic: "If the answer sounds intuitive and hence familiar with my preconceptions, I answered it correctly". Familiarity, not convenience, might explain recall here. On the other hand, is the observed pattern of recall optimal? Over-estimating the rate of correct answers might bolster ego-utility, but why should subjects commit the (costly) mistake of saying that they saw one question that in fact they did not?

[15] Something similar should be expected if individuals get some intrinsic utility from being accurate in S1, but not in S2. The literature on motivated inference has to our knowledge rarely considered this possibility, though.

the magnitude of the bias does not depend on the scale of the bonus (neither on the size of the prize associated to the desirable event). Similarly, Ertac (2011) studies the effect of rewarding accurate beliefs and finds no significant difference across compensated and non-compensated sessions in terms of the distribution of priors and the absolute value of the bias. If people suffer from NBO, in summary, the cost of such bias does not seem to set limits on its size, at least given the parameterizations that have been considered so far.

More indirect evidence is also at odds with *P4*. In the surveys conducted by Case and Shiller (2003), respondents showed an exaggerated optimism regarding the future evolution of house prices. Also, people tend to be too optimistic regarding their future health (Weinstein, 1987; Caponecchia, 2010). A survey conducted by Weinstein et al. (2005), for instance, find that smokers underestimate their risk of developing lung cancer and that they tend to see themselves as having a lower risk than the average smoker. Given the huge losses potentially involved in all these scenarios, one would expect people to verge more on the Bayesian side than the wishful one. ■

The following prediction is a trivial corollary from *P1*. Consider the simple framework presented in *Section 2.1* and assume that there are two deciders whose utility has the functional form defined in [1]. For one decider, we assume that $dv/(d\hat{\pi}) > 0$, so believing that the state is red gives her some utility. For the other one, however, $dv/(d\hat{\pi}) < 0$, so she gets utility from thinking that the true state is blue. Assuming that both deciders have the same priors, the toy model predicts that the deciders' optimal beliefs will differ. Specifically, the first (second) decider will inflate (deflate) the probability of the state Red.

*Prediction 5 (P5):* If two deciders have sufficiently opposed preferences, their beliefs will be opposed as well, even if they have access to the same set of information.

*Discussion and evidence:* There are to our knowledge few controlled studies on this issue. The existing evidence seems mixed. On the positive side, recall Mayraz (2011), discussed after *P1*, where Bakers and Farmers had opposed preferences and expressed different beliefs given the same information –a time series of "wheat" prices. Regarding negative evidence, Bar-Hillel and Budescu (1995) ask subjects each week over a four-week period to estimate the probability that the Dow Jones Industrial Average (DJIA) would change by more than 20 points in a week. Half of the subjects were eligible for a cash prize if the DJIA changed by more than 20 points, while the other half were eligible if it changed by less than 20 points. Obviously, each group had opposing interests in what regards the state of the world, and according to NBO this should translate into some differences in their probability judgments. Yet no significant differences were found in any of the four weeks of their study. Finally, some indirect evidence comes from comparing valenced and neutral treatments, i.e., treatments with and without a desired state of the world, respectively. Most of the papers cited after *P2* consider both types of treatments and it must be noted that few studies report a difference across treatments –e.g., Eil

and Rao (2011) and Möbius et al. (2014). Barron (2019), for a negative example, compares belief updating by two groups of individuals who share the same prior belief and receive an equally informative signal, but for one group the signal goes in line with their interests while for others the signal has no positive or negative value. His results show that there are no significant differences across groups. ■

As we have noticed before, few NBO models formalize in detail the learning process, with exceptions like Möbius et al. (2014). Suppose however that agents can choose whether or not to process a signal that has been observed (i.e., to not pay attention to it, explain it away, or not think about it); if the agent gets anticipatory utility from putting high probability on the good state, then agents may selectively ignore bad news. If people systematically choose to under-update to negative news, in other words, a natural response could be a biased search for information, i.e., avoiding those sources that typically bring negative news. This seems rational given that the bad news are hardly taken into account. The next prediction is then a straightforward one.

*Prediction 6 (P6; information avoidance):* Deciders avoid evidence or data that might challenge their favorite beliefs, even if they are aware that such information exists and it is costless to obtain. In contrast, they actively search for evidence that is expected to be consonant with their favorite beliefs.

*Discussion and evidence:* In each treatment of Eil and Rao (2011), subjects are asked at the end of the experiment their willingness-to-pay (WTP) to learn whether they are in the top half of the IQ or beauty ranking. Their WTP was positive for those who believed to be in the top ranks, whereas those who had arrived at "bad" posteriors required a subsidy. In the control, WTP was not affected by beliefs or signal history. In a previous experiment by Brock and Balloun (1967), participants are presented several sound recordings, some of them with speeches supporting or challenging the link between smoking and lung cancer. The recordings were masked by static noise that the participants can remove by repeatedly pressing a button. Interestingly, the smokers pressed the button significantly more than nonsmokers during the speeches that questioned the link between smoking and lung cancer, but less in the speeches that supported that link. The range of situations in which similar patterns of behavior have been reported is quite heterogeneous. In financial contexts, for example, Karlsson et al. (2009) observe that investors check the evolution of their portfolios more frequently in rising markets than when they are falling. This kind of behavior has been often called "ostrich effect". Also, some people at risk of suffering some health problems seem reluctant to get diagnosed. For example, in the survey conducted by Kellerman et al. (2002) with populations at high risk for HIV, the main reasons reported by individuals refusing to get tested were denial of risk and fear of being HIV-positive. Similarly, in their study with individuals at risk for Huntington disease (HD), Oster et al. (2013) find that some of the individuals avoid to get tested. Moreover, these individuals tend to be overoptimistic about their probability of having HD and behave as if they do not have HD. In the study by

Ganguly and Tasoff (2016), some participants were reluctant to get tested for the HSV-1 and the HSV-2 (herpes simplex virus) and were even willing to renounce to some payoff in order to avoid the test. Further, the avoidance rate was three times higher for the HSV-2, which was perceived as worse than the HSV-1.

Other works provide evidence of information avoidance in problems in which one's actions may have an impact on others and in which moral principles play a role. In those situations, moral rules may constrict individual action. However, if individuals are not sure about the consequences of their actions on others, some of them may choose the most convenient action without feeling their moral principles compromised. This is often known as the "moral wiggle room". In Dana et al. (2007), for instance, subjects play a binary dictator game with (decider, receiver) allocations A: (6, 1) and B: (5, 5). In an additional treatment labeled *hidden information,* the decider does not know a priori the receiver's payoffs in options A and B, knowing only that they can respectively equal $1 and $5 or $5 and $1, as determined by a coin flip before the session –deciders could choose to reveal the true payoff scheme at no cost before choosing. The idea of information avoidance implies that deciders will believe that option A leads to the socially optimal allocation (6, 5), having no reason to reveal the true payoffs, since they are sure that A is the 'good' option. Indeed, only 26 % of the deciders chose option A in the baseline, whereas 44 % of them preferred to remain ignorant in *hidden information,* where around 70 % chose A. To test the robustness of these results, Larson and Capra (2007) conduct a slightly different experiment, obtaining similar results (in this case, 53 % of participants in the hidden payoff treatment chose to remain ignorant). ■

*Prediction 7 (P7):* Predictions *P1* to *P6* above do not depend on the complexity of the data available or the attention paid to it.

*Discussion and evidence:* the available evidence on self-serving biases does not seem to support *P7,* as the extent of the bias seems to depend on the quality of the data sample. For instance, Haisley and Weber (2010) find that the behavior observed in a scenario similar to that in Dana et al. (2007) above is even more pronounced in more ambiguous contexts. Specifically, the experiment includes a "simple risk" and an "ambiguous" treatment. In the former, the receiver's payoff scheme is determined by a random draw from an urn which contains 10 red and 10 blue chips. In the ambiguous" treatment, the specific content of that urn was unknown, although its composition was determined randomly using a uniform distribution from 0 to 20 (note, however, that the a priori probability of sampling a red chip from the ambiguous urn is 0.5 as in the "simple risk" treatment). Haisley and Weber (2010) report data on beliefs, e.g., deciders systematically overestimate the receivers' expected payoff: the mean estimate bias was of $0.29 among the participants in the simple risk treatment and of $0.89 among those in the ambiguous treatment. For another example, Dahl and Ransom (1999) find that Mormons who attend church regularly, serve in volunteer church positions, and have previously served as missionaries tend to report an increased tithable income base, which suggests that knowledge prevents

biases.[16] Further, church members exhibit stronger biases in the tithing scenarios that they view more ambiguously, i.e., when they express more uncertainty, possibly as a result of having weak evidence on the issue. Finally, Babcock et al. (1997) explore several ways to reduce self-serving biases in pre-trial bargaining, using the *same* experiment as Babcock et al. (1995). In this study, as the reader may recall from the discussion of *P1*, subjects play either the role of plaintiff or defendant. Further, they read the same case materials, estimate the judge's decision and what a fair settlement would be like, and bargain over the case. Yet the authors find that having subjects who already know their roles list the weaknesses in their own case significantly decreases differences in their estimates of the judge's award and in the occurrence of impasse. This suggests that attention paid to the non-favorable evidence can reduce the extent of the self-serving biases. ■

Another important issue regards the sources of belief heterogeneity. In the models of NBO, including our toy model, two individuals may keep different beliefs if they have access to a different set of information (different priors or observed signals). This is in line with Bayesian models. However, in the models of NBO, belief heterogeneity can be the result as well of heterogeneous preferences. Suppose that each member of a group of people must rank the beauty of each member of the group, including themselves. If beauty is important for their self-esteem and the cost of making unrealistic estimates is low enough, the NBO models predict that their beliefs will differ. Specifically, everyone will overestimate their own beauty, even if all of them have access to the same information. Finally, in Bénabou and Tirole (2002) and Bénabou (2015), different individuals may keep different beliefs depending on how aware they are that their recalled sample is systematically biased. Some individuals may be fully aware that the evidence they recall is a biased sample that favors their preferred state. In this case, they will adequately incorporate this in their beliefs, alleviating the bias to some extent, although not completely. On the other extreme, fully naïve individuals will form their beliefs as if their recalled sample was unbiased.

*Prediction 8 (P8; heterogeneity and individual co-variates):* In models of optimism, the sources of belief heterogeneity are the evidence observed, the individual preferences and the individual's degree of awareness about their bias.

*Discussion:* Some evidence suggests a more complex picture. For instance, some studies point out that individuals with a higher ability to solve complex problems are apparently more likely to act as motivated believers. In Kahan, Peters, Dawson, and Slovic (2017) subjects are presented with a problem that requires drawing valid causal inferences from empirical data. One problem is ideologically neutral, dealing with a new skin-rash treatment. Here subjects who scored highest on numeracy tests did substantially better than less numerate ones. A second problem is more ideologically

[16] Indeed, more active church members are significantly less uncertain about what income sources are tithable.

charged, having to do with a study on a gun-control ban. The impact of numeracy on performance in this condition was minimal. In fact, higher numeracy improves a subject's performance in detecting covariance only when the correct response is congenial to her subjects' political outlooks. Apparently, numerate subjects use their capacity selectively to conform their interpretation of the data to the result most consistent with their political outlooks. In Möbius et al. (2014), in contrast, high and low performers on the IQ test do not update differently: Both exhibit conservatism and asymmetric updating. To complicate matters further, Coutts (2019) reports that performance in his math and verbal skills quiz is (weakly) related to both conservatism and asymmetry; specifically, subjects with higher ability appear to be less pessimistic and conservative. Grossman and Owens (2012), in turn, observe asymmetry only among those participants with the lowest quiz scores and unrealistically high priors. Similarly, Ertac (2011) reports that self-serving tendencies are more prevalent among those subjects who exhibit most confident priors. Perhaps this signals some kind of confirmation bias, although the evidence is not conclusive, as we have noted earlier.

Gender stands out as another individual variable whose potential correlation with positive thinking deserves to be explored. In this regard, Ertac (2011) finds that men interpret positive feedback in the verbal task less conservatively than women, leading to less pessimistic posteriors. This was not observed in the math task, though, which both genders regarded as much less difficult. Möbius et al. (2014) also report that men update beliefs about performance in the IQ quiz significantly less conservatively than women, although no significant differences in asymmetry are found across genders. Similarly, Coutts (2019) finds no difference in asymmetry between genders. In contrast to Möbius et al. (2014), he reports no significant effect of gender on conservativeness, at least when interacted with ability in the quiz. Finally, Mayraz (2011) finds no difference across genders in the size of the bias.

## 3. Conclusion

Suppose that you get a nicely written e-mail. It includes a list with five names and promises that if you send one Euro to each of these five guys, put your own name on the list, and forward the new list to as many people as you can, you will receive in 20-60 days around 10,000 Euros in the mail. If you do not send the e-mail, you will not only miss out on the money, but bad fortune will befall to you. Encouragingly, the e-mail states that this is not a scam or a pyramid scheme. Indeed, it includes the testimony of some fellow who claims to have earned a HUGE amount of money and assures you that this program works perfectly every time. How does this sound to you?

Economic theories of optimism predict that people will fall victim of illusions when the cost of holding such illusions is not high. As our review of the literature suggests, there seems to be a puzzle in this respect. On one hand, the abundant evidence reviewed after *Prediction 1* suggests that the phenomenon of motivated

inference is a real one. Indeed, alternative, cognitive explanations of the fact that our preferences frequently interact with our beliefs appear to be overall less convincing. On the other hand, the more controlled tests surveyed after *Prediction 2* tend to find little evidence for asymmetric updating, a point made as well by Benjamin (2019). In general terms, the economic models of NBO do not seem to fully capture the phenomenon.

This means that the precise causes of the optimistic biases are still unclear. We find several open questions. First, under what conditions do people update in an asymmetric manner? Contrary to what the NBO models state, the mere existence of an interest and a small risk of a loss are not enough –e.g., Gotthard-Real, 2017; Barron, 2019. In this regard, second, the conditions required on the signal for the occurrence of a bias are still in doubt. One conjecture is that timing and salience are relevant, because the biases might be the result of biased recall and selective attention –Kunda (1990), Benábou and Tirole (2012). This is a point that has been relatively underexplored: in all the economic studies cited in the discussion of *P2,* for instance, subjects get feedback at all times about previous signals. Some of these studies, e.g., Eil and Rao (2011) and Möbius (2014), interestingly, still find a bias. This could be a signal that NBO does not require that subjects "forget" or "misinterpret" signals. From our point of view, however, these studies present confounds because the subjects' priors are not exogenously fixed (see footnote 12). It must be noted, third, that the studies differ in the characteristics of the event/state about which subjects' beliefs are controlled. Eil and Rao (2011), Ertac (2011), and Möbius et al. (2014) elicit beliefs about self-relevant states, e.g., one's intelligence or attractiveness, whereas other studies focus more on financial domains. Perhaps people are relatively more optimistic about some issues. In some studies, fourth, the priors are objective whereas they are of a more subjective nature in others, particularly in those about ego-relevant events. As Coutts (2019) notes, this could affect updating behavior. Further, the signal generating process is always explicit and objective in these studies. In the field, however, inference scenarios are often characterized by high ambiguity.

A fifth question is heterogeneity. What (if any) personal characteristics reinforce asymmetric updating? As we discussed after *P8*, the NBO models are relatively agnostic in this regard, but there are several variables that might be worth considering. Gender stands as a very prominent candidate. While it seems well-established that men are more self-confident than women, even when actual performance is controlled for (e.g. Niederle & Vesterlund, 2007; Barber & Odean, 2001), it is still unclear whether women and men differ in the way they process information. Sixth, a potential confound in the existing studies regards the belief elicitation mode. In the vast majority of the studies cited in *P2*, specifically, subjects must estimate posterior probabilities using incentive compatible methods like the quadratic scoring rule or the lottery method. A concern therefore is that subjects are required to exert substantial mental effort, as conditional probabilities are difficult to compute, and the elicitation methods are complex –indeed, they often require intuitive additional explanations,

e.g., Coutts (2019), Barron (2019).[17] Could people infer in a very reflexive manner in these scenarios, perhaps diminishing optimistic biases? Although the evidence so far seems negative (see *P4*), seventh, the role of material incentives on motivated inference possibly warrants further research. Does the risk of a (significant) loss lead to more Bayesian updating, as the NBO models predict?

Another question, eighth, refers to the persistence of the biases, particularly when people interact and exchange ideas. NBO models predict that optimistic biases should be relatively resistant, at least if people do not change their perceptions about their present and future preferences.[18] This means a reluctance to admit mistakes and change opinion when others (including experts) think otherwise; although this resistance would be moderated if people are sophisticated and anticipate their biases, Bénabou and Tirole (2002). Arguably, a policy designed to make citizens' beliefs more accurate would be useless, since beliefs will hardly change, and harming, because it would make people unhappier (if it works). Further research should be devoted to explore the relationships between optimism and socialization: Are (optimistic) people immune to others' opinions?[19] When do people expect to be able to convince others, if they are thought to be wishful thinkers? Last, but not least, there is the question of functionality and welfare. That is, economists have long considered Bayes' rule as a normative benchmark for belief updating, based on an instrumental argument. If beliefs yield intrinsic utility, however, it might be optimal in some circumstances to hold distorted, non-Bayesian beliefs. What are these circumstances, that is, when does the positive side of distortion dominate the negative one?

## References

Akerlof, G. A., & Dickens, W. T. (1982). The economic consequences of cognitive dissonance. *The American Economic Review, 72*(3), 307-319.

Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. *The self in social judgment, 1* (pp. 85-106).

Babcock, L., & Loewenstein, G. (1997). Explaining Bargaining Impasse: The Role of Self-Serving Biases. *Journal of Economic Perspectives, 11*(1), 109-126.

Babcock, L., Loewenstein, G., & Issacharoff, S. (1997). Creating convergence: Debiasing biased litigants. *Law and Social Inquiry, 22*(4), 913-925.

[17] Ertac (2011) runs some non-incentivized sessions and concludes that her results are not an artifact of the belief elicitation method. Still, subjects must report numerical likelihoods.

[18] Conversely, NBO predicts that beliefs should *rapidly* change if preferences over them change –e.g., if the monetary prizes associated to the states change their relative positions. We do not know of any evidence in this respect.

[19] In Dahl and Ransom (1999), to illustrate some issues here, slightly over half of the survey respondents discuss what items to tithe with their spouse, while around 40% have sought outside advice about tithing. Further, individuals who seek advice from spouses are significantly more generous in their tithing donations.

Babcock, L., Loewenstein, G., Issacharoff, S., & Camerer, C. (1995). Biased judgments of fairness in bargaining. *American Economic Review, 85*(5), 1337-1343.

Babcock, L., Wang, X., & Loewenstein, G. (1996). Choosing the wrong pond: social comparisons in negotiations that reflect a self-serving bias. *Quarterly Journal of Economics, 111*(1), 1-19.

Barber, B. M., & Odean, T. (2001). Boys will be boys: gender, overconfidence, and common stock investment. *Quarterly Journal of Economics, 116*(1), 261–292.

Bar-Hillel, M., & Budescu, D. (1995). The elusive wishful thinking effect. *Thinking and Reasoning, 1*(1), 71-103.

Barkan, R., Ayal, S., Gino, F., & Ariely, D. (2012). The pot calling the kettle black: distancing response to ethical dissonance. *Journal of Experimental Psychology: General, 141*(4), 757-773.

Barron, K. (2019). *Belief updating: Does the 'good-news, bad-news' asymmetry extend to purely financial domains?* Unpublished manuscript.

Baumeister, Roy F. (1998). The Self. In D. Gilbert, S. Fiske & G. Lindzsey (Eds.), *The Handbook of Social Psychology*. McGraw-Hill.

Bénabou, R. (2015). The economics of motivated beliefs. *Revue d'économie politique, 125*(5), 665-685.

Bénabou, R. & Tirole, J. (2002). Self-Confidence and Personal Motivation. *Quarterly Journal of Economics 117*(3), 871-915.

Bénabou, R. & Tirole, J. (2016). Mindful economics: the production, consumption, and value of beliefs, *Journal of Economic Perspectives, 30*(3), 141-164.

Benjamin, D. J. (2019). Errors in probabilistic reasoning and judgment biases. In *Handbook of Behavioral Economics: Applications and Foundations 1* (Vol. 2, pp. 69-186). North-Holland.

Benoit, J. & Dubra, J. (2011). Apparent Overconfidence. *Econometrica, 79*(5), 1591-1625.

Brenner, L. A. (2003). A random support model of the calibration of subjective probabilities. *Organizational Behavior and Human Decision Processes, 90*(1), 87-110.

Brenner, L. A., Koehler, D. J., Liberman, V., & Tversky, A. (1996). Overconfidence in probability and frequency judgments: A critical examination. *Organizational Behavior and Human Decision Processes, 65*(3), 212-219.

Brock, T. C., & Balloun, J. L. (1967). Behavioral receptivity to dissonant information. *Journal of Personality and Social Psychology, 6*(4), 413-428.

Brunnermeier, M. K. & J. A. Parker (2005). Optimal Expectations. *American Economic Review, 95*(4), 1092–1118.

Budescu, D. V., Wallsten, T. S., & Au, W. T. (1997). On the importance of random error in the study of probability judgment. Part II: Applying the stochastic judgment model to detect systematic trends. *Journal of Behavioral Decision Making, 10*(3), 173-188.

Buser, T., Gerhards, L., & Van Der Weele, J. (2018). Responsiveness to feedback as a personal trait. *Journal of Risk and Uncertainty, 56*(2), 165-192.

Caponecchia, C. (2010). It won't happen to me: An investigation of optimism bias in occupational health and safety. *Journal of Applied Social Psychology, 40*(3), 601-617.

Case, K. E., & Shiller, R. J. (2003). Is there a bubble in the housing market? *Brookings Papers on Economic Activity* (2) (pp. 299-362).

Charness, G., & Dave, C. (2017). Confirmation bias with motivated beliefs. *Games and Economic Behavior, 104* (pp. 1-23).

Chew, S. H., Huang, W., & Zhao, X. (2019). Motivated False Memory. https://ssrn.com/abstract=2127795 or http://dx.doi.org/10.2139/ssrn.2127795

Clark, J., & Friesen, L. (2009). Overconfidence in forecasts of own performance: an experimental study. *Economic Journal, 119*(534), 229-251.

Coutts, A. (2019). Good news and bad news are still news: Experimental evidence on belief updating. *Experimental Economics, 22*(2), 369-395.

Cross, K. P. (1977). Not can, but will college teaching be improved? *New Directions for Higher Education* (17) (pp. 1-15).

Dahl, G. D., & M. R. Ransom (1999). Does Where You Stand Depend on Where You Sit? Tithing Donations and Self-Serving Bias. *American Economic Review, 89*(4), 703-727.

Dana, J., Weber, R. A., & Kuang, J. X. (2007). Exploiting moral wiggle room: experiments demonstrating an illusory preference for fairness. *Economic Theory, 33*(1), 67-80.

Di Tella, R., Perez-Truglia, R., Babino, A., & Sigman, M. (2015). Conveniently upset: Avoiding altruism by distorting beliefs about others' altruism. *American Economic Review, 105*(11), 3416-42.

Eil, D., & Rao, J. M. (2011). The Good News–Bad News Effect: Asymmetric Processing of Objective Information about Yourself. *American Economic Journal: Microeconomics, 3*(2), 114-38.

Epley, N., & Gilovich, T. (2016). The mechanics of motivated reasoning. *Journal of Economic Perspectives, 30*(3), 133-40.

Ertac, S. (2011). Does self-relevance affect information processing? Experimental evidence on the response to performance and non-performance feedback. *Journal of Economic Behavior and Organization, 80*(3), 532-545.

Festinger, L. (1957). A theory of cognitive dissonance. Stanford University Press.

Forsythe, R., Rietz T., & Ross, T. (1999). Wishes, Expectations, and Actions: Price Formation in Election Stock Markets. *Journal of Economic Behavior and Organization, 39* (pp. 83-110).

Ganguly, A., & Tasoff, J. (2016). Fantasy and dread: The demand for information and the consumption utility of the future. *Management Science, 63*(12), 4037-4060.

Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond "heuristics and biases". *European review of social psychology, 2*(1), 83-115.

Ging-Jehli, N. R., Schneider F. H., & Weber, R. A. (2019). *On self-serving strategic beliefs* (Working Papers No 315). University of Zurich.

Gino, F., Norton, M. I., & Weber, R. A. (2016). Motivated Bayesians: feeling moral while acting egoistically. *Journal of Economic Perspectives, 30*(3), 189-212.

Gotthard-Real, A. (2017). Desirability and information processing: An experimental study. *Economics Letters, 152* (pp. 96-99).

Grossman, Z., & Owens, D. (2012). An unlucky feeling: Overconfidence and noisy feedback. *Journal of Economic Behavior and Organization, 8*4(2), 510-524.

Haisley, E. C., & Weber, R. A. (2010). Self-serving interpretations of ambiguity in other-regarding behavior. *Games and Economic Behavior, 68*(2), 614-625.

Heger, S. A., & Papageorge, N. W. (2018). We should totally open a restaurant: How optimism and overconfidence affect beliefs. *Journal of Economic Psychology, 67* (pp. 177-190).

Kahan, D. M., Peters E., Dawson E. C., & Slovic P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy, 1*(1), 54-86.

Kahneman, D., & Tversky A. (Eds.) (2000). *Choices, Values, and Frames*. Cambridge University Press and the Russell Sage Foundation.

Karlsson, N., Loewenstein, G., & Seppi, D. (2009). The ostrich effect: Selective attention to information. *Journal of Risk and Uncertainty, 38*(2), 95-115.

Kellerman, S. E., Lehman, J. S., Lansky, A., Stevens, M. R., Hecht, F. M., Bindman, A. B., & Wortley, P. M. (2002). HIV testing within at-risk populations in the United States and the reasons for seeking or avoiding HIV testing. *Journal of Acquired Immune Deficiency Syndromes, 31*(2), 202-210.

Kindleberger, C. P., & Aliber, R. Z. (1978). *Manias, Panics, and Crashes: A History of Financial Crises*. 4th ed. Wiley, 2000.

Kruger, J. (1999). Lake Wobegon be gone! The 'below-average effect' and the egocentric nature of comparative ability judgments. *Journal of Personality and Social Psychology, 77*(2), 221-232.

Kuhnen, C. M. *(2015)*. Asymmetric learning from financial information. *The Journal of Finance, 70*(5), 2029-2062.

Kunda, Z. (1987). Motivated Inference: Self-Serving Generation and Evaluation of Causal Theories. J*ournal of Personality and Social Psychology 53*(4), 636-47.

Kunda, Z. (1990). The case for motivated reasoning. *Psychological Bulletin, 108*(3), 480-498.

Larrick, R. P., Burson, K. A., & Soll, J. B. (2007). Social comparison and confidence: When thinking you're better than average predicts overconfidence (and when it does not). *Organizational Behavior and Human Decision Processes, 102*(1), 76-94.

Larson, T., & Capra, C. M. (2009). Exploiting moral wiggle room: Illusory preference for fairness? A comment. *Judgment and Ddecision Making, 4*(6), 467-474.

Lichtenstein, S., & Fischhoff, B. (1977). Do those who know more also know more about how much they know? *Organizational Behavior and Human Performance, 20*(2), 159-183.

Mayraz, G. (2011). *Wishful Thinking* (Working Paper).

Möbius, M. M., Niederle, M., Niehaus, P., & Rosenblat T. (2014). *Managing Self-Confidence* (Working Paper).

Moore, D. A., & Cain, D. M. (2007). Overconfidence and underconfidence: when and why people underestimate (and overestimate) the competition. *Organizational Behavior and Human Decision Processes, 103*(2), 197-213.

Moore, D. A., & Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review, 115*(2), 502-517.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *Quarterly Journal of Economics, 122*(3), 1067-1101.

Oster, E., Shoulson, I., & Dorsey E. R. (2013). Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease. *American Economic Review, 103*(2), 804-830.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: divergent perceptions of bias in self versus others. *Psychological Review, 111*(3), 781.

Rabin, M. (1994). Cognitive Dissonance and Social Change. *Journal of Economic Behavior and Organization, 23*(2), 177-94.

Rabin, M., & Schrag, J. (1999) First Impressions Matter: A Model of Confirmatory Bias. *The Quarterly Journal of Economics, 114*(1), 37-82.

Samuelson, L. (2004). Modeling Knowledge in Economic Analysis. *Journal of Economic Literature, 42*(2), 367-403.

Schwardmann, P., & van der Weele, J. (2019). Deception and self-deception. *Nature Human Behavior, 3*(10), 1055-1061.

Sharot, T., Korn, C. W., & Raymond Dolan (2011). How Unrealistic Optimism is Maintained in the Face of Reality. *Nature Neuroscience 14*(11), 1475-79.

Sharot, T., & Garrett, N. (2016). Forming Beliefs: Why Valence Matters. *Trends in Cognitive Science, 20*(1), 25-33.

Shiller (2000). *Irrational Exuberance*. Princeton University Press..

Simon, H. A. (1956). Rational Choice and the Structure of the Environment. *Psychological Review, 63*(2), 129-138.

Spinnewijn, J. (2015). Unemployed but optimistic: Optimal insurance design with biased beliefs. *Journal of the European Economic Association, 13*(1), 130-167.

Svenson, O. (1981). Are we all less risky and more skillful than our fellow drivers? *Acta Psychologica, 47*(2), 143-148.

Thompson, L., & Loewenstein, G. (1992). Egocentric Interpretations of Fairness and Interpersonal Conflict. *Organizational Behavior and Human Decision Processes, 51*(2), 176-197.

Tversky, A., & Kahneman, D. (1973). Availability: A Heuristic for Judging Frequency and Probability. *Cognitive Psychology, 5*(2), 207-232.

Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science, 185*(4157), 1124-1131.

von Hippel, W., & Trivers, R. (2011). The evolution and psychology of self-deception. *Behavioral and Brain Sciences, 34*(1), 1-16.

Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and social psychology, 39*(5), 806-820.

Weinstein, N. D. (1987). Unrealistic optimism about susceptibility to health problems: Conclusions from a community-wide sample. *Journal of Behavioral Medicine, 10*(5), 481-500.

Weinstein, N. D., Marcus, S. E., & Moser, R. P. (2005). Smokers' unrealistic optimism about their risk. *Tobacco Control, 14*(1), 55-59.

Wicklund, R. A., & Brehm, J. W. (1976). *Perspectives on cognitive dissonance*. Lawrence Erlbaum.

Wiswall, M., & Zafar, B. (2015). How do college students respond to public information about earnings? *Journal of Human Capital, 9*(2), 117-169.

CUADERNOS ECONÓMICOS ICE INFORMACIÓN COMERCIAL ESPAÑOLA

# Sobornos y contratación pública: un experimento económico*

# *Bribes and Public Procurement: An Economic Experiment*

Álvaro Javier Vargas Villamizar
Universidad Santo Tomás

Luis Alejandro Palacio García
Universidad Industrial de Santander

***Resumen***

*La presente investigación pretende determinar si el pliego de condiciones genera incentivos para la entrega de sobornos en la contratación pública. En este sentido, el objetivo es analizar cómo la discrecionalidad del Burócrata y las asimetrías de información afectan el nivel de entrega de sobornos y la discriminación de concursantes en una subasta para la adjudicación de contratos públicos. Para ello se realizó un experimento económico de laboratorio, cuyo diseño experimental simuló una subasta por contratos públicos bajo cuatro tratamientos. Se encontró que las asimetrías de información son una herramienta útil para disminuir la entrega de sobornos y que existe discriminación negativa por parte del Burócrata al asignar los contratos. Por lo cual se logra concluir que el poder que posee el Burócrata para decidir sobre el ingreso de los participantes a la subasta es el principal detonante para la entrega de sobornos.*

***Palabras clave:*** *sobornos, pliego de condiciones, contratos públicos, economía experimental.*
***Clasificación JEL:*** *C70, C92, D73.*

***Abstract***

*The present research aims to determine whether tender specifications generate incentives for the delivery of bribes in public procurement. In this sense, the objective is to analyze how bureaucratic discretion and asymmetries of information affect the level of bribe delivery and the discrimination of bidders in a public contract auction. To achieve this, an economic experiment in the laboratory was conducted, in which the experimental design simulated an auction for public contracts under four treatments. It was found that information asymmetries are a useful tool to reduce the delivery of bribes and that there is negative discrimination by the bureaucrat when assigning contracts. Therefore, it was concluded that the power of the bureaucrat to choose over the admission of the participants to the auction is the main trigger for the delivery of bribes.*

***Keywords:*** *bribes, tender specifications, public contracts, experimental economics.*
***JEL Classification:*** *C70, C92, D73.*

* Los autores agradecen los comentarios de Daniel Parra, del Wissenschaftszentrum de Berlín; Alexandra Cortés Aguilar, de la Universidad Industrial de Santander, y María Camila Álvarez Lozano, del Departamento Administrativo Nacional de Estadística (DANE). Todas sus sugerencias fueron bien recibidas. Este artículo está basado en el trabajo de grado de maestría titulado «Análisis de la entrega de sobornos en la contratación pública: evidencia experimental», que se desarrolló en el marco del proyecto titulado «Juegos de negociación: una aplicación de la economía experimental en el aula», con código 1877, financiado por la Vicerrectoría de Investigación y Extensión de la Universidad Industrial de Santander.


DOI: https://doi.org/10.32796/cice.2020.99.7009
Recibido: Enero 2020 · Aceptado: Febrero 2020

## 1. Introducción

La entrega de sobornos en los contratos públicos puede afectar a las economías a través de la desviación de fondos, que a su vez podrían ayudar a disminuir la pobreza, el hambre, los problemas de salud y favorecer el crecimiento económico. ¿Se imagina cuáles serían las posibles consecuencias de elegir discrecionalmente sobre el destino de los recursos públicos? Aunque suene difícil de creer, este fenómeno se presenta en múltiples entornos de la contratación pública, lo cual convierte la posibilidad que tiene el Burócrata de decidir quién puede contratar con el Estado en uno de los determinantes de la entrega de sobornos. No obstante, uno de los principales inconvenientes a la hora de estudiar la entrega de sobornos es que esta se da en entornos clandestinos; inclusive, su correcto funcionamiento depende directamente del nivel de confidencialidad en que se mantenga. Por tal motivo, los datos a los que acceden los investigadores generalmente están relacionados con casos que salieron a la luz pública, por alguna denuncia, o indicadores de percepción subjetiva que son construidos a partir de la opinión ciudadana.

Es necesario tener en cuenta que, en la mayoría de los países del mundo, las inversiones de los recursos del Estado se realizan a través de empresas privadas que ganan una subasta (licitación o concesión). Autores como Vickrey (1961) y Binmore y Kemplerer (2002) han realizado aportes significativos a esta literatura, mostrando cómo estos mecanismos, de acuerdo con su diseño, pueden resultar herramientas útiles a la hora de realizar intercambios, o en su defecto, asignar recursos públicos o derechos de propiedad. Por tanto, subastar el acceso a los recursos públicos tiene como principal propósito buscar una asignación eficiente, lo que, en términos del Teorema de Coase, sería asignar los derechos de propiedad al agente que más valore el bien.

Por tanto, dado que las inversiones de recursos públicos se realizan en un mercado imperfecto, la subasta ha servido como un mecanismo eficiente para la asignación de recursos. Sin embargo, en muchas ocasiones el mecanismo habilitante (pliego de condiciones[1]) para poder entrar en la subasta carece de objetividad debido a que recae en la discrecionalidad de los Burócratas encargados de adelantar el proceso licitatorio. Por tanto, podrían existir distorsiones a la hora de definir qué requisitos debe cumplir el agente para entrar a concursar. Dicha discrecionalidad puede ser explotada como una herramienta para excluir de la subasta a posibles competidores y así favorecer al agente que esté en la capacidad de entregar un soborno. De esta forma, agentes privados tienen la posibilidad de sobornar al Burócrata para que este evite la entrada de más participantes y, por ende, puedan ganar la subasta, y a su vez, obtener el mayor beneficio posible al no existir competencia por parte de terceros.

En este sentido, la pregunta de investigación que se pretende responder en este trabajo es si existe algún tipo de discriminación hacia los concursantes cuando el

[1] El pliego de condiciones en Colombia es un documento a través del cual el Burócrata define cuáles son los requisitos mínimos para que el concursante pueda entrar en la subasta. Dentro de dichos requisitos generalmente se exige experiencia, capacidad operativa, músculo financiero y nivel de riesgo.

Burócrata puede observar el soborno, el nivel de tecnología y la apropiación de recursos. Con base en lo anterior, la presente investigación busca obtener evidencia experimental que muestre la incidencia de las asimetrías de información del Burócrata en la entrega de sobornos, la apropiación de recursos y los niveles de participación. Específicamente, se desea conocer si observar el soborno, el nivel de tecnología y la apropiación de los concursantes juegan un papel importante a la hora de entregar sobornos, porque a pesar de que la asimetría recae en el Burócrata, quienes entregan el soborno son los licitantes, lo cual permite conocer si la existencia de dichas asimetrías generan cambios en el comportamiento de los concursantes a la hora de ofertar un soborno, y si existe algún tipo de comportamiento recíproco por parte del subastador.

En la segunda sección se realiza una aproximación al concepto de corrupción y la entrega de sobornos. En la tercera, se realiza un abordaje sobre el fenómeno específicamente en entornos experimentales. En la cuarta, se presentan el diseño experimental y los procedimientos. En la quinta y última sección se muestran los resultados obtenidos por el experimento de laboratorio y las conclusiones.

## 2. Una aproximación al fenómeno de la corrupción y la entrega de sobornos

El trabajo seminal realizado por Rose-Ackerman (1975) ha sido el punto de partida para los estudios contemporáneos sobre corrupción. Si bien Becker (1968) fue pionero en la aplicación del sistema de elección racional en la economía del crimen, Rose-Ackerman (1975) presentó los fundamentos teóricos para el estudio específico en temas de corrupción. A pesar de estos aportes seminales, el estudio empírico de la corrupción es una actividad compleja debido a su naturaleza oculta, de hecho, una parte importante de los trabajos realizados se basan en criterios de percepción que generan una distorsión subjetiva. De ahí la importancia de utilizar las herramientas de la economía experimental para simular entornos de negociación, pues de esta manera, es posible obtener información útil que dé luz hacia una comprensión más exacta de cómo funciona la entrega de sobornos.

La corrupción es un fenómeno amplio que puede presentarse de diferentes maneras, mediante el soborno, la malversación y apropiación de fondos, el nepotismo, el tráfico de influencias, etc. Todas sus manifestaciones se pueden condensar en la definición dada por Transparencia Internacional (2016): «abuso del poder para el beneficio propio». Etimológicamente, corrupción significa «contaminar, menoscabar la pureza», del latín es el participio pasado de *corrumpere* que significa «destruir». Una segunda definición es la presentada por Khan (2004), quien entiende la corrupción como un intercambio entre un agente privado (o un grupo) y un agente público (Burócrata), donde el agente público ofrece algo a un privado que no podría obtener de otra manera. De modo que es posible resumir tres condiciones básicas para que se pueda presentar la corrupción: 1) poder discrecional de un funcionario público; 2) rentas económicas que puedan ser extraídas, y 3) debilidad institucional (Aidt, 2003).

Es evidente que estas definiciones abarcan una gran cantidad de situaciones y diversas prácticas corruptas que van más allá del soborno, por lo cual es conveniente tener en cuenta las posibles maneras en que agentes privados pueden relacionarse con funcionarios públicos. En este contexto, se puede usar como referencia una tipología presentada por Hellman *el al*. (2000), donde dividen los tipos de corrupción según los mecanismos a través de los cuales las empresas o particulares pueden ejercer control sobre el Estado. Las categorías son: 1) *Grand Corruption*: pagos privados, ilícitos y no transparentes a los funcionarios públicos que tratan de modificar la elaboración de las leyes, normas, decretos y reglamentos; 2) *Petty Corruption*: es una forma de corrupción perseguida por agentes de nivel medio o bajo, que pueden estar mal pagados, y que recurren a estas rentas adicionales, relativamente pequeñas, pero ilegales. La pequeña corrupción afecta desproporcionadamente a los miembros más pobres de la sociedad, que pueden experimentar regularmente peticiones de dinero en sus encuentros con la Administración pública[2] y para acceder a servicios públicos básicos (hospitales, escuelas, autoridades locales de licencias, policía y autoridades tributarias, entre otros).

El estudio de la corrupción cobra especial importancia y es de sumo interés social comprender el fenómeno para poder tomar acciones que eviten socavar la estabilidad institucional. Como pionero en el tema, el profesor Johann Graf Lambsdorff ha sido uno de los académicos más reconocidos en el mundo por estudiar de forma continuada el fenómeno de la corrupción. Él es el creador del Índice de Percepción de Corrupción, y como menciona en su libro *Institutional economics and reform: Theory, evidence and policy,* el propósito de crear este indicador es el de dar información elemental sobre el fenómeno para fomentar la investigación en la temática.

Lambsdorff (2007) manifiesta que la corrupción es el abuso del poder público para el beneficio privado, y que tal factor representa un gran reto de estudio para las ciencias sociales. El problema de la corrupción ha estado en el debate público desde hace siglos, pero su abordaje científico es relativamente nuevo. Además, como efecto de una falta de consenso acerca de la conceptualización de lo que es la corrupción, se han creado múltiples teorías, que con el tiempo se van quedando obsoletas a medida que se hacen nuevos hallazgos científicos. El problema de que esto suceda es que, al final, se han terminado realizando avances en la comprensión del fenómeno, pero todos apuntan a vías diferentes.

Para entender la corrupción desde la perspectiva económica, Lambsdorff (2007) acude al uso de la teoría de juegos, modelando un escenario en el que los concursantes en una licitación pública deben crear una estrategia para ganar el contrato sin

[2] Este tipo de práctica es muy común en los países en vías de desarrollo, en donde aprovechando la ineficiencia de las instituciones y la inoperancia del sistema, los Burócratas piden sumas de dinero a cambio de realizar trámites que de acuerdo con el ordenamiento jurídico deberían ser gratuitos o de bajo costo. En tal sentido, las personas en dichas situaciones se ven obligadas a acceder debido al monopolio estatal para la realización de dichos trámites. A efectos teóricos, cuando es el Burócrata el que solicita la entrega de dinero a cambio de un favor en la Administración pública Lambsdorff (2007) lo define como extorsión y no como un soborno.

los recursos suficientes para hacerlo por la vía legal. En este tipo de escenarios, es posible evidenciar que los participantes harán propuestas muy innovadoras, de características inusuales, pero altamente criminales. Para Lambsdorff (2007) esto es lo que significa realmente la corrupción: lograr violar las reglas del juego de tal forma que nadie lo hubiese podido anticipar[3]. Por consiguiente, la corrupción no solo debería verse desde la óptica ortodoxa en la que alguien haría algo criminal únicamente cuando la utilidad esperada de hacerlo es mayor que el costo esperado[4].

Algunos economistas manifiestan que la corrupción en forma de soborno es un intercambio mutuamente beneficioso, particularmente, los microeconomistas plantean que mientras existan incentivos racionales a la entrega de sobornos es inevitable su aparición. En la misma línea, los funcionalistas suponen que la simple existencia del soborno es una prueba clara y fehaciente de su utilidad para alguien. Si después de la entrega de un soborno existen dos partes que se encuentran mejor que antes ¿por qué el soborno es perjudicial? Lambsdorff (2007) manifiesta que el problema de estas aproximaciones es que en ellas no se tienen en cuenta las posibles consecuencias colaterales del soborno, que pueden llevar a que, en algunos casos, la corrupción resulte perjudicial incluso para las partes que intervienen en la misma.

Un ejemplo de las consecuencias relacionadas con la corrupción es la imposibilidad de tomar decisiones acertadas debido al constante flujo de propuestas corruptas que son ofrecidas a los Burócratas que ostentan el poder. Por ejemplo, a toda la sociedad beneficia la construcción de una carretera de buena calidad, pero si el proyecto sale mal (a pesar de haber realizado una adjudicación transparente), el gobierno debe cancelar el proyecto. Si el contrato no fue asignado al contratista correcto, las falencias de éste harían que el contratante (Estado) decidiera cancelar el contrato. Pero en el escenario en el que la adjudicación se realice bajo un esquema de entrega de sobornos, lo más probable es que el proyecto no se detenga (porque el Burócrata que recibió el soborno está interesado en que el proyecto continúe), y que la calidad de los materiales para la construcción del proyecto se supediten a la cuantía de sobornos que fueron entregados para obtener el contrato.

Otro de los problemas de la existencia de la entrega de sobornos es que no genera incentivos para el correcto actuar de los demás agentes que hay en el mercado, dado que la competencia desleal lleva a un paulatino debilitamiento de la estrategia honesta, hasta el punto en que los que desean acceder a los contratos o beneficios que brindan los recursos públicos se vean obligados por inercia a entrar en actividades corruptas. Además, Huang y Li (2015) manifiestan que las personas que entregan sobornos estarán siempre dispuestas a hacerlo si y solo si pueden obtener algún beneficio de ello. Por lo que se asume que el fenómeno de la entrega de sobornos

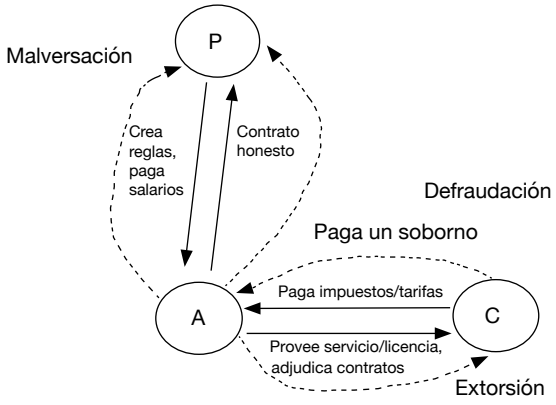
[3] En este contexto, buscar alternativas lo suficientemente creativas y poco esperadas podría ser un mecanismo para evitar la deducción de estrategias a través de la inducción hacia atrás. De este modo, al no poderse prever la contingencia, las probabilidades de éxito criminal aumentarán.

[4] Esta perspectiva es la presentada por Gary Becker (1968) en la que plantea que las actividades criminales son llevadas a cabo a través de un proceso de elección racional. Se esperaría entonces que un criminal no lleve a cabo sus fechorías si el costo esperado de la misma fuese superior a la utilidad esperada.

existirá en la medida en que la probabilidad de cumplimiento de la contraparte sea lo suficientemente alta.

Por otra parte, Andvig y Fjeldstad (2000) manifiestan que la corrupción es un proceso en el cual se da un intercambio de favores entre dos actores, un Agente y un Cliente. Este abordaje permite observar el fenómeno de la corrupción en un Modelo de Agencia, en donde el Principal, que es un superior, delega su poder en el Agente y establece las reglas acerca de cómo deben cumplirse las tareas asignadas. Se supone que el Agente debe servir al Cliente obedeciendo las reglas que fueron establecidas por el Principal. La malversación, el soborno y la extorsión son variantes de un comportamiento corrupto del Agente[5] que abandonó el cumplimiento de las reglas que fueron preestablecidas. En el soborno, el Cliente es un actor que realiza un pago al Agente a cambio de obtener un favor al que legalmente no tiene derecho, por ejemplo, una devolución de impuestos o un contrato público, por lo que toma un rol de sobornador. En la extorsión, el Agente utiliza el poder que le fue concedido para extraer dinero o algún otro tipo de beneficios del Cliente. En cuanto a la malversación, se entiende como un proceso en el cual, sin la necesidad de un Cliente, el Agente destina recursos con un fin específico a otros rubros que legalmente no lo merecían, y que habitualmente generan un beneficio privado. Es así como el soborno, la malversación y la extorsión implican que las reglas que fueron establecidas inicialmente sean transgredidas lesionando los intereses del Principal.

**FIGURA 1**
**MODELO DE CORRUPCIÓN PRINCIPAL-AGENTE**




**FUENTE:** Elaboración propia, adaptado de Lambsdorff (2007).

[5] En el Modelo de Agencia presentado, el Agente (Burócrata) es el que tiene la condición de administrar los recursos que le fueron otorgados por el Principal (Estado) de acuerdo con la ley. Además, existe un Cliente (ciudadano) que desea interactuar con el Principal pero solo puedo hacer a través del Agente debido a la naturaleza metafísica del Estado.

En la Figura 1 se observa el fenómeno explicado anteriormente, en donde se establece la relación existente entre el Principal (P), el Agente (A) y el Cliente (C). En principio, si cada uno de ellos decide obedecer las reglas, podrían seguir el conducto legal representado por la línea continua, que, para este caso en particular, sería obtener los servicios estatales sin realizar pago alguno. Por otro lado, el Agente y el Cliente pueden decidir romper las reglas y acceder de forma fraudulenta a los servicios o contratos estatales a través del pago de un soborno o una extorsión como se observa a través de las líneas punteadas de la Figura 1. En síntesis, la existencia de actos corruptos por parte del Agente y del Cliente obedecen a la lógica individual de maximización, en la que se explota el poder concedido para el beneficio propio, llevando a resultados poco beneficiosos a nivel social pero cuya existencia es justificada en la medida en que llevar a cabo dichos actos sea rentable para los involucrados.

## 3. Evidencia experimental sobre la entrega de sobornos en subastas

El primer experimento de laboratorio sobre corrupción fue el desarrollado por Abbink *et al*. (2002) y fue llamado el Juego del Soborno. Ellos partieron del hecho que la corrupción tiene tres factores claves: reciprocidad, efectos negativos a terceros (externalidades negativas) y penas altas al ser descubiertos. En ese sentido, generalmente la entrega de sobornos se da en círculos sociales en los que existe cierto nivel de confianza, y por esta razón existe un alto nivel de reciprocidad que conlleva el posterior cumplimiento de los acuerdos. Estos autores también utilizaron experimentos para aislar cada uno de los tres efectos, donde se pudo encontrar que la reciprocidad puede establecer relaciones de soborno, en las cuales, las externalidades negativas no tienen ningún efecto aparente en el comportamiento de los agentes. La razón de esto es sencilla: quienes pagan sobornos no están seguros de si van a obtener el servicio prometido o simplemente serán timados; por lo que pagar un soborno significa esperar que el receptor tenga un sentido de reciprocidad y actúe consecuentemente con este. Este hallazgo es muy importante, pues evidencia la importancia de la reciprocidad al establecer redes criminales y la poca relevancia del hecho de causar daños a la sociedad como factor clave para considerar la decisión.

El trabajo de Büchner *et al*. (2008) realizó una simulación del sistema de subasta de contratación pública, en el que los licitantes efectúan ofertas con sobornos sumergidos con el propósito de aumentar sus ganancias. Si bien en este trabajo no se tiene en cuenta el comportamiento del contratante (Burócrata), se realiza una aproximación hacia el comportamiento de los licitantes en el que se observa que el conocimiento sobre el daño social no disminuye la entrega de sobornos de manera significativa. Podría entenderse que en este tipo de contextos las decisiones de los participantes se inclinan más hacia la búsqueda del beneficio individual que al colectivo.

También Vetter (2013) llevó a cabo un estudio en el que analizó de manera experimental si existen mecanismos legales alternativos al soborno que conlleven los mismos resultados. Encontró que la anticipación de recompensas futuras por la entrega voluntaria de algún tipo de dádiva por parte de un concursante, basta para sesgar los procesos de selección futuros a su favor**.** A pesar de que recompensas futuras no sean exigibles, su anticipación produce un intercambio mutuamente beneficioso entre las partes a costa del detrimento del bien público. De esta forma, se puede llegar a un resultado similar al que se hubiese llegado con la entrega de un soborno, pero sin caer en la ilegalidad.

Además, Makowsky y Wang (2015) realizaron un experimento basado en el juego de bienes públicos y el ultimátum para estudiar la malversación de fondos y el comportamiento hacia la denuncia en diversas estructuras organizacionales. En un juego secuencial, los participantes pudieron decidir apropiarse de parte de los recursos disponibles o delatar a alguien con un comportamiento deshonesto, en cuyo caso, nadie del grupo obtenía ningún tipo de ganancia. Al final, el monto que quedó en el bien público, es decir, lo que no fue apropiado benefició a todos los miembros de cada grupo por igual. Makowsky y Wang (2015) muestran que las estructuras planas y de pirámide son más eficaces en la reducción de malversación de fondos. Además, evidencian que las tasas de malversación de fondos y la denuncia de irregularidades aumentan con el número de niveles en la estructura.

Por último, Albano *et al*. (2008) realizan un análisis acerca de los efectos de la colusión entre los licitantes al disminuir el número de participantes en el proceso de subasta. Su estudio se centra en encontrar los efectos negativos de la colusión y evidencia que una reducción del número de participantes en los procesos de contratación pública lleva a resultados anticompetitivos y poco eficientes, pues si se restringe la participación de licitadores en las subastas, se pierden las propiedades de esta, lo que generaría una tendencia a un mercado monopólico.

El aporte sustantivo que realiza el presente trabajo está relacionado con brindar información acerca de cómo funciona el fenómeno de la entrega de sobornos en contextos de contratos públicos a partir de la simulación en un entorno controlado. Por ende, en el diseño experimental se simula la posibilidad que tienen los Burócratas de excluir competidores. Lo novedoso es que, a pesar de que la literatura ha venido realizando avances en el estudio de los sobornos, no se han estudiado los efectos de la existencia del pliego de condiciones como mecanismo excluyente. Si bien, en la legislación, se justifica su existencia con el propósito de realizar discriminación positiva (dejar entrar en la subasta únicamente a los concursantes más aptos), dicha figura se podría estar explotando de tal manera que estaría desdibujando su verdadero propósito. Por eso, es importante entender si el poder habilitante del Burócrata en las subastas es un generador de incentivos a la entrega de sobornos por parte de agentes privados.

## 4. Modelo, diseño y procedimiento experimental

El experimento fue llevado a cabo entre el 5 y 7 de septiembre de 2018. Participaron estudiantes de diversas carreras de la Universidad Industrial de Santander, la Universidad Santo Tomás y la Universidad Autónoma de Bucaramanga en Colombia. En total participaron 96 personas en el experimento, 24 jugadores en cada tratamiento, de los cuales 18 fueron concursantes y seis Burócratas. Adicionalmente, los grupos estuvieron conformados por las mismas personas a largo de diez periodos repetidos, con edades que oscilaban entre los 18 y 26 años. Las instrucciones fueron dadas en voz alta para todos los participantes en la sala. Se utilizó el software O-Tree (Chen *et al*., 2016). Al final de cada sesión se pagó a los participantes lo correspondiente en pesos colombianos a la sumatoria de los puntos ganados en cada periodo. Los pagos oscilaron entre los 15.000 y 30.000 pesos (cinco y diez dólares respectivamente), con pagos promedio por persona de 21.000 pesos (siete dólares).

Se realizó el reclutamiento a través de las redes sociales del Laboratorio de Economía Experimental del Grupo EMAR (EMAR-Lab). La convocatoria se realizó respetando los tres principios éticos establecidos en el *Belmont Report* de 1979: respeto por las personas, beneficencia y justicia. Además, como parte de los protocolos internos del grupo de investigación EMAR, se contó con el aval del Comité de Ética en Investigación Científica (CEINCI) de la Universidad Industrial de Santander. Para garantizar los protocolos éticos, a cada uno de los sujetos experimentales se pidió firmar un consentimiento informado que explicaba cuáles eran los potenciales riesgos, los beneficios, la voluntariedad en la participación y la protección de datos personales para dar estricto cumplimento a la Ley de *Habeas Data* en Colombia.

### *4.1. Diseño experimental*

El experimento se organizó en tres etapas, con cuatro tratamientos y grupos de cuatro participantes. En cada grupo existían dos roles diferenciados, un Burócrata y tres concursantes. Cada grupo poseía una bolsa con 5.000 pesos[6] a ser subastados por el Burócrata. En cada grupo existían tres tipos de cuentas: la Cuenta Pública, la Cuenta del Burócrata y la Cuenta Personal de cada concursante. Cada uno de los concursantes poseía un nivel de tecnología $T_G$ que tomaba valores aleatorios entre 1 y 3 con datos continuos.

**Etapa 0:** De forma aleatoria se asignan los roles, un Burócrata y tres concursantes. Luego, de forma también aleatoria se asigna un nivel de tecnología variable a lo largo de los diez periodos. Los grupos y los roles asignados en la Etapa 0 se mantuvieron a lo largo de las diez rondas de la sesión, por lo que los participantes

[6] Se decidió que cada punto experimental (tradicionalmente llamados ECU en la literatura) equivalía a un peso colombiano.

únicamente pudieron interactuar con las personas que se encontraban en su mismo grupo y además conservaron sus roles de Burócrata y concursante a lo largo de toda la sesión del experimento.

**Etapa 1. Concursantes:** Los concursantes de forma simultánea realizaban una propuesta al Burócrata (estrategia del concursante $S_{Ci}$) acerca de cómo se deberían asignar los 5.000 pesos del grupo. En este sentido, cada concursante debió elegir de los 5.000 pesos cuánto dinero destinaba a la Cuenta del Burócrata (soborno), a su Cuenta Personal (apropiación) y a la Cuenta Pública (legalidad).

**Etapa 2. Burócrata:** Decidía cuáles concursantes entraban a la subasta (estrategia del Burócrata $S_B$)[7]. Es decir, el Burócrata elige entre «Sí Participa» o «No Participa» para cada uno de los Concursantes del grupo en función de la información disponible en cada tratamiento. El Burócrata podía permitir participar a los tres concursantes, a dos de ellos, a uno solo o a ninguno. En este sentido, este proceso simula la capacidad que tiene para determinar cuáles propuestas pueden participar en la subasta, teniendo claro que el ganador de la subasta será el concursante cuya propuesta de asignación de dinero a la Cuenta Pública haya sido la más alta. En caso de empate, el ganador se elegirá al azar entre los empatados. En esta etapa, de acuerdo con el tratamiento, el Burócrata podía observar el soborno, la cantidad apropiada por el concursante, así como el nivel de tecnología. En ningún tratamiento el Burócrata pudo observar la cantidad de dinero enviada a la Cuenta Pública.

**Ganancias de los participantes:** Para calcular las ganancias, el dinero que se propuso asignar en la Cuenta del Burócrata (*CB*) y la Cuenta Personal (*CP*) no se multiplica por ningún factor. El dinero asignado de la propuesta ganadora se transfiere al dueño de la cuenta. Por su parte, el dinero asignado en la Cuenta Pública (*CPub*) por parte del ganador de la subasta se multiplicó por su tecnología ($T_G$) y se dividió en cuatro partes iguales. En síntesis, las funciones de pago para cada uno de los participantes del grupo teniendo en cuenta su rol y si ganó o no la subasta son las siguientes:

Las ganancias del Burócrata se expresan como $U_B$, donde

$$U_B(S_{Ci}, S_B) = CB + \frac{CPub * T_G}{4}$$

es la cantidad de dinero que el Concursante ganador ($C_G$) de la subasta propuso asignar a su Cuenta de Burócrata (*CB*) más la cantidad de dinero que el concursante

[7] El diseño contempló esta posibilidad para poder endogenizar el poder que tienen los Burócratas en la vida real de establecer un pliego de condiciones que pueda excluir a los competidores que desee de la subasta. Por tanto, en las subastas por contratos públicos reales, el Burócrata podrá decidir a cuáles concursantes permite participar en la subasta y a cuáles no.

ganador propuso asignar a la Cuenta Pública (*CPub*) multiplicado por el nivel de tecnología del concursante ganador de la subasta dividido en cuatro. La función de pago del concursante ganador

$$U_{C_G}(S_{C_i}, S_B) = CP + \frac{CPub * T_G}{4}$$

es la cantidad de dinero que él propuso asignar a su Cuenta Personal más la cantidad de dinero que propuso asignar a la Cuenta Pública multiplicado por su nivel de tecnología, dividido en cuatro. Por último, la ganancia de los concursantes no ganadores

$$U_{C_{NG}}(S_{C_i}, S_B) = \frac{CPub * T_G}{4}$$

es la cantidad de dinero que el concursante ganador propuso asignar a la Cuenta Pública multiplicado por el nivel de tecnología del concursante ganador de la subasta dividido en cuatro.

**Equilibrio de Nash:** $S_{C_i}$: $\{CB = 4999, CC = 1, CP = 0\}$, $S_B$: $\{C_i = Sí\ Participa\}$

Por inducción hacía atrás, el Equilibrio de Nash del juego se ubica en donde todos los concursantes proponen destinar 4.999 pesos a la cuenta del Burócrata, un peso a la Cuenta de Concursante y el Burócrata decidirá seleccionar la estrategia «Sí Participa» al concursante que tenga mayor nivel de tecnología en caso de que pueda verla.

**Óptimo de Pareto:** $S_{C_i}$: $\{CB = 0, CC = 0, CP = 5000\}$, $S_S$: $\{C_i = Sí\ Participa\}$

El Óptimo de Pareto se ubica donde todos los concursantes destinen cero pesos a la Cuenta del Burócrata, cero pesos a la Cuenta de Concursante y 5.000 pesos a la Cuenta Pública, y que el Burócrata elija la estrategia «Sí Participa» al concursante con mayor nivel de tecnología en caso de que pueda ser observada.

### *4.2. Procedimientos experimentales*

Tratamientos

Para conocer cómo cambian las decisiones del Burócrata en función de las asimetrías de información se definieron dos tratamientos a partir de la información que estuvo disponible para el Burócrata en la Etapa 2 del experimento. Se definieron dos variables: 1) Si el nivel de tecnología era observable por el Burócrata o no, y 2) Si el Burócrata observó la cantidad de dinero que el concursante propuso enviar como soborno o la cantidad de dinero que el concursante propuso enviar a su Cuenta

Personal. Observar o no el soborno permite evidenciar si existe algún tipo de conducta recíproca por parte del Burócrata. Por otro lado, observar el nivel de tecnología brinda la posibilidad de conocer si existe algún tipo de discriminación positiva por parte del Burócrata. Por esta razón, se asume que habría cambios, dado que, al no poder ver el soborno, adoptar una conducta recíproca sería una tarea más compleja. De lo anterior se obtuvo un diseño de tratamientos factorial $2 \times 2$. El tipo de tratamiento ejecutado fue entre sujetos *(between subjecs protocol)*, es decir, que cada participante únicamente pudo participar en un tratamiento del experimento. Los participantes tomaron decisiones a lo largo de diez periodos repetidos y no hubo *re-matching*.

**CUADRO 1**
**DISEÑO EXPERIMENTAL**

| **Información disponible para el Burócrata** | | **Observa Soborno** | **Observa Cuenta Personal** |
|---|---|---|---|
| **Tecnología** | **Observa** | 1 | 3 |
| | **No observa** | 2 | 4 |

NOTA: Información adaptada del diseño experimental
**FUENTE:** Elaboración propia

Los tratamientos se organizaron en función del tipo de asimetría de información a la que se vio expuesto el Burócrata. En el Tratamiento 1 el Burócrata pudo observar el soborno y el nivel de tecnología de los concursantes. En el Tratamiento 2 únicamente visualizó el soborno. En el Tratamiento 3 accedió al nivel de tecnología y la cantidad apropiada por los concursantes, y en el Tratamiento 4 únicamente conoció la cantidad apropiada (Cuadro 1).

## 5. Resultados experimentales

### *5.1. Sobornos*

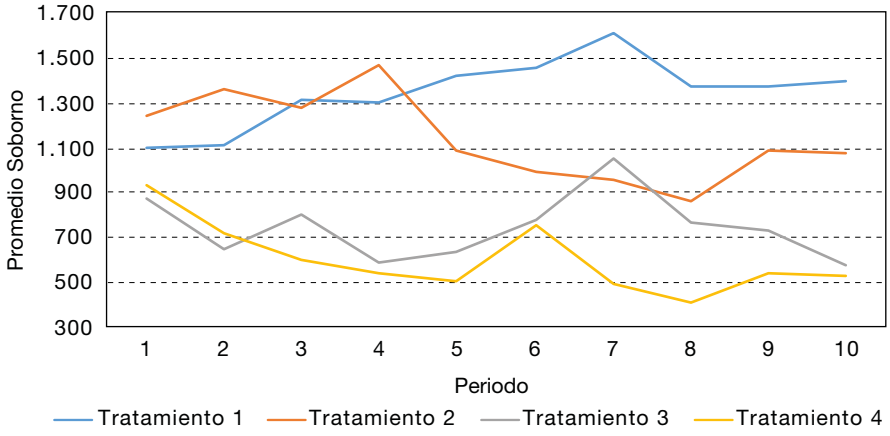
Se encontraron diferencias significativas en el promedio de sobornos enviados por parte de los concursantes al Burócrata. En el Cuadro 2 se muestra que el promedio de los sobornos se ubicó en 1.348, 1.143, 746 y 602 pesos respectivamente desde el primer al cuarto tratamiento. En la Figura 2 se evidencia que a pesar de que en algunos periodos no se observaron diferencias entre tratamientos en el promedio de sobornos enviados, sí se muestra que a pesar de que la decisión de sobornar recaía exclusivamente en los concursantes, entre más alta fue la asimetría de información a la que se sometió el Burócrata (Tratamiento 4), menor fue el incentivo por parte de los concursantes para sobornarlo.

## CUADRO 2
## DIFERENCIAS DE MEDIAS T-STUDENT DE SOBORNO ENTRE TRATAMIENTOS

| | **Grupo 1** | | **Grupo 2** | | **Grupo 3** | |
|---|---|---|---|---|---|---|
| | **T1** | **T2** | **T1** | **T3** | **T1** | **T4** |
| *Media* | 1.348,22 | 1.143,29 | 1.348,22 | 746,14 | 1.348,22 | 602,76 |
| *Error estándar* | 52,98 | 61,72 | 52,98 | 64,86 | 52,98 | 39,88 |
| *Desviación estándar* | 710,74 | 828,03 | 710,74 | 870,25 | 710,74 | 535,05 |
| *Diferencia de media* | 204,93* | | 602,08*** | | 745,46*** | |
| *Int. conf. 95 % inf.* | 1.243,69 | 1.021,51 | 1.243,69 | 618,15 | 1.243,69 | 524,06 |
| *Int. conf. 95 % sup.* | 1.452,76 | 1.265,08 | 1.452,76 | 874,14 | 1.452,76 | 681,46 |
| | **Grupo 4** | | **Grupo 5** | | **Grupo 6** | |
| | T2 | T3 | T2 | T4 | T3 | T4 |
| *Media* | 1.143,29 | 746,14 | 1143,29 | 602,76 | 746,14 | 602,76 |
| *Error estándar* | 61,72 | 64,86 | 61,72 | 39,88 | 64,86 | 39,88 |
| *Desviación estándar* | 828,03 | 870,25 | 828,03 | 535,05 | 870,25 | 535,05 |
| *Diferencia de media* | 397,15*** | | 540,53*** | | 143,38 | |
| *Int. conf. 95 % inf.* | 1.021,51 | 618,15 | 1.021,51 | 524,06 | 618,15 | 524,06 |
| *Int. conf. 95 % sup.* | 1.265,08 | 874,14 | 1.265,08 | 681,46 | 874,14 | 681,46 |

NOTA: * $p < .05$; ** $p < .01$; *** $p < .001$.
**FUENTE**: Elaboración propia.

## FIGURA 2
## SOBORNOS POR TRATAMIENTO




**FUENTE**: Elaboración propia.

De lo anterior puede inferirse que las asimetrías de información juegan un papel crucial a la hora de entregar sobornos, pues a medida que aumentan se desestabiliza el vínculo existente entre los agentes, razón por la cual se dificulta la negociación bilateral. Se puede asumir que parte de los efectos de las asimetrías de información están relacionados con la imposibilidad por parte del Burócrata de asumir una actitud recíproca. Esta evidencia experimental coincide con los hallazgos de Huang y Li (2015), quienes manifiestan que los agentes solo estarán dispuestos a sobornar si pueden sacar algún provecho de ello. En esta línea, el Tratamiento 1 fue el que más se acercó al Equilibrio de Nash y el Tratamiento 4 al Óptimo de Pareto. Además, es necesario destacar que entre los Tratamientos 1 y 4 se observan resultados claramente diferenciados a lo largo de los diez periodos jugados, lo que muestra un evidente efecto tratamiento.

En el Cuadro 2 se evidencian las diferencias de medias de los sobornos entregados en cada uno de los cuatro tratamientos. Se realizaron comparaciones iteradas entre los tratamientos para verificar si hubo cambios estadísticamente significativos a través de una prueba T-Student en la que se comparó el promedio de los diez periodos jugados. Se encontró que en todos los grupos excepto el Grupo 6 hubo diferencias de media significativas. Por lo que se evidencia que en lo referente al soborno sí existió un efecto tratamiento. Además, se observó que respecto al Tratamiento 1, el aumento de las asimetrías de información impactó negativamente en la entrega promedio de sobornos, disminuyendo de 1.348 pesos hasta los 602 pesos en el Tratamiento 4. Los datos experimentales muestran que a pesar de que cada tratamiento introdujo una combinación de asimetrías de información diferentes, estas tuvieron un impacto directo y significativo en la reducción de la entrega de sobornos.

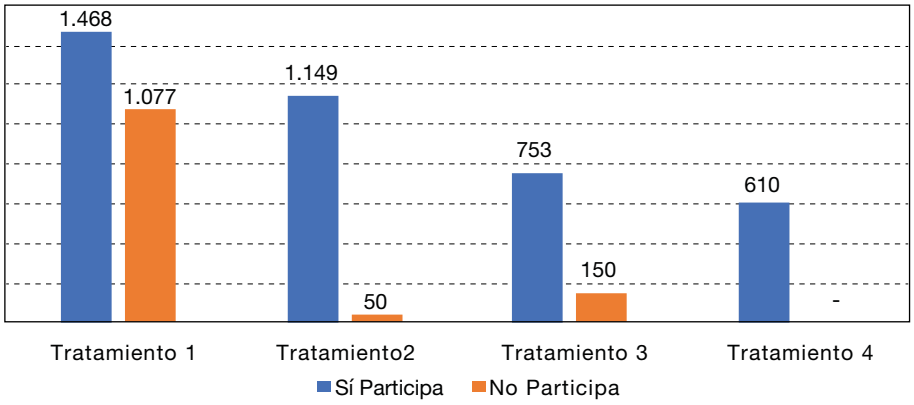
La Figura 2 muestra que en la medida en que el Burócrata se vio sometido a mayores asimetrías de información los concursantes tuvieron menos incentivos para enviar sobornos. No obstante, también es necesario conocer cuáles fueron las decisiones que fueron tomadas por parte del Burócrata. En este caso en particular, se desea conocer si se observó algún tipo de comportamiento recíproco hacia los concursantes. En la Figura 3 se evidencia el comportamiento de las decisiones tomadas por el Burócrata (Sí Participa/No Participa) en función de los sobornos enviados por los concursantes.

La evidencia muestra que sí hubo reciprocidad por parte del Burócrata en todos los tratamientos. De hecho, en cada uno de ellos, en promedio, los concursantes a los que el Burócrata les permitió participar fueron aquellos que enviaron mayores sobornos. En particular, es necesario hacer énfasis en los Tratamientos 1 y 2 en los cuales el Burócrata tuvo la posibilidad de observar el soborno directamente. Llama la atención que en el Tratamiento 4 en la mayoría de los periodos el Burócrata eligió «Sí Participa» a la totalidad de los concursantes, por lo que no hubo jugadores no participantes. En la Figura 4 se muestra que en el Tratamiento 1 únicamente participaron en la subasta el 69,44 % de los concursantes, en contraste con casi la totalidad que se observó en los tratamientos restantes.

**FIGURA 3**

**PROMEDIO DE PROPUESTAS DE SOBORNO ENVIADAS POR TRATAMIENTO Y DISCRIMINADAS POR LA DECISIÓN DEL BURÓCRATA EN LOS DIEZ PERIODOS**

**(En $)**

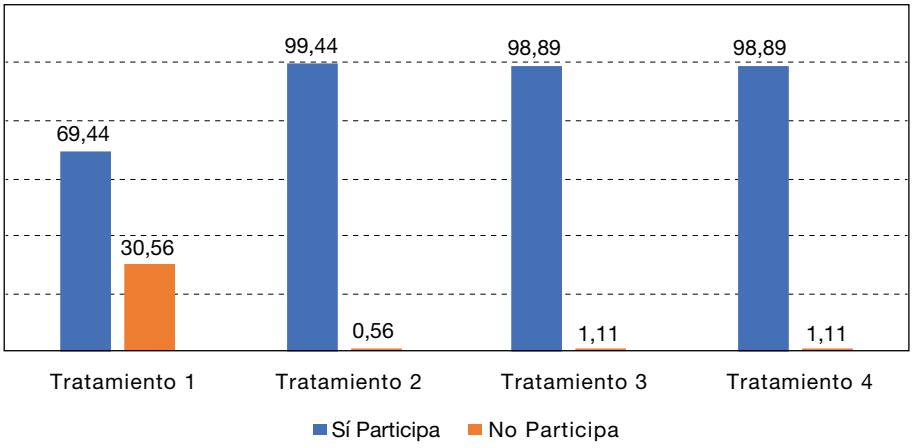



**FUENTE**: Elaboración propia

**FIGURA 4**

**CONCURSANTES HABILITADOS POR TRATAMIENTO EN LOS DIEZ PERIODOS**

**(En %)**

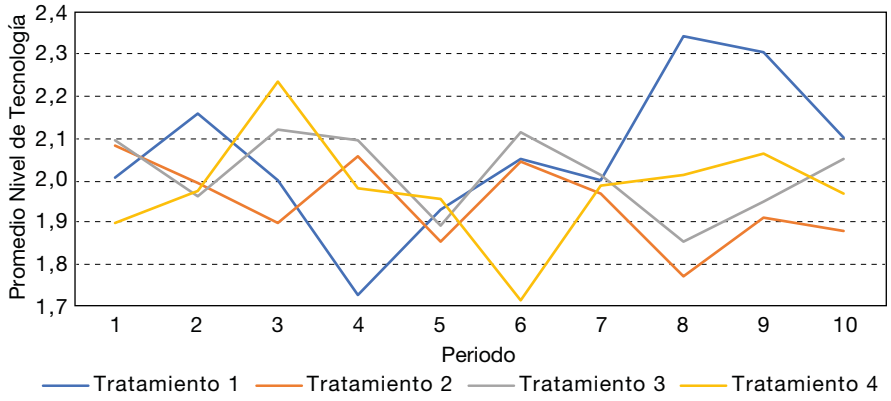



**FUENTE**: Elaboración propia

Por otro lado, la Figura 5 muestra el nivel de tecnología de los concursantes a quienes el Burócrata permitió que participaran en la subasta en cada uno de los tratamientos a lo largo de los diez periodos de juego. En este caso, no se observa el efecto tratamiento respecto a esta variable, lo cual se debe a que la información del nivel de tecnología no fue una variable decisiva a la hora de dejar participar concursantes en la subasta. En efecto, la evidencia permite asumir que no se observó discriminación positiva por parte del Burócrata, porque no se ve una tendencia clara y diferencial entre los tratamientos. Por ende, se asume que a pesar de que el nivel de tecnología era una variable que causaba interdependencia en las utilidades de los jugadores (porque a mayor nivel de tecnología del concursante ganador, mayor sería el pago esperado), no se observa una tendencia preferente del Burócrata para permitir participar en la subasta a los concursantes de mayor tecnología.

**FIGURA 5**

**PROMEDIO DE NIVEL DE TECNOLOGÍA DE LOS CONCURSANTES A QUIENES EL BURÓCRATA PERMITIÓ PARTICIPAR POR TRATAMIENTO**



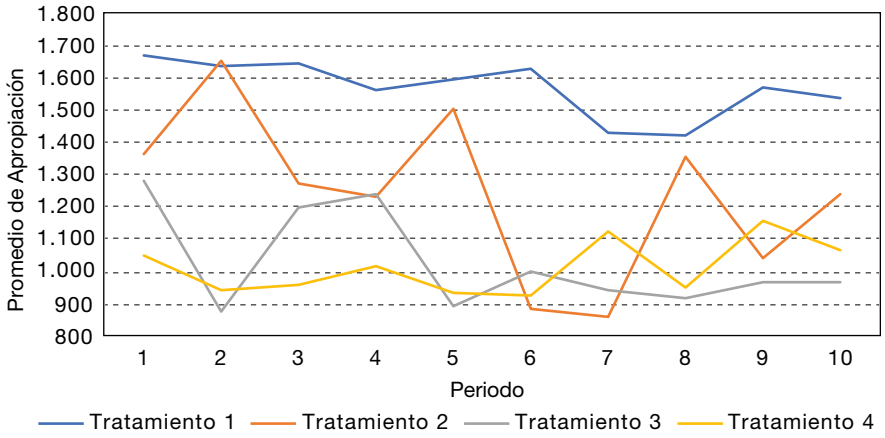

**FUENTE**: Elaboración propia

Con los resultados obtenidos es posible deducir que la discriminación llevada a cabo por el Burócrata sobre la entrada de concursantes a la subasta se basó exclusivamente en la posibilidad de observar el soborno de forma directa. La evidencia experimental muestra que ver el nivel de tecnología de los concursantes no fue un criterio de los Burócratas para su participación. Asimismo, se encontró que imperó la discriminación negativa, puesto que el criterio de selección imperante fue el soborno.

### *5.2. Apropiación*

La apropiación de recursos por parte de los concursantes es entendida como la cantidad de dinero que proponen enviar a su Cuenta Personal. Esta apropiación es un tipo de corrupción relacionada con la malversación. Al igual que en el apartado anterior con el soborno, apropiarse de dinero que es de propiedad colectiva representa un costo de oportunidad en las ganancias de todo el grupo, porque inexorablemente hará que el remanente en la Cuenta Pública sea menor. En la Figura 6 se observa que en el Tratamiento 1, a lo largo de los diez periodos, el promedio de apropiación se mantuvo relativamente estable, oscilando entre los 1.400 y 1.600 pesos. En cuanto a los Tratamientos 3 y 4 se evidencia que el nivel de apropiación de recursos fue menor, oscilando entre los 800 y 1.200 pesos. El único tratamiento que no tuvo una tendencia estable fue el Tratamiento 2, en el cual se observa una amplia volatilidad situándose en rangos que estuvieron entre los 800 y 1.600 pesos, por lo que el resultado en este tratamiento no es concluyente en lo referente a la apropiación.

**FIGURA 6**
**PROMEDIO DE APROPIACIÓN POR PARTE DE LOS CONCURSANTES**




**FUENTE**: Elaboración propia

La Figura 6 muestra que en los tratamientos en los que existieron mayores sobornos, la cantidad apropiada en promedio por parte de los concursantes fue mayor. En el Tratamiento 1 en donde el soborno promedio se ubicó en 1.348 pesos, la cantidad promedio apropiada por los concursantes fue de 1.667 pesos. En cambio, en el Tratamiento 4 en el que el tamaño promedio del soborno fue de 603 pesos, la apropiación ascendió únicamente a 1.053 pesos. Este es un resultado interesante debido a que por racionalidad individual se podría asumir que los concursantes se apropiarían de más recursos en la medida que decidieran enviar menos sobornos al Burócrata, pero

la evidencia experimental muestra que ocurrió un fenómeno totalmente contrario. Al parecer, existe un relativo sentimiento de justicia a la hora de repartirse el botín, en donde de manera implícita se toman decisiones más correctas en conjunto en la medida en que los demás miembros del grupo están imposibilitados para ser corruptos.

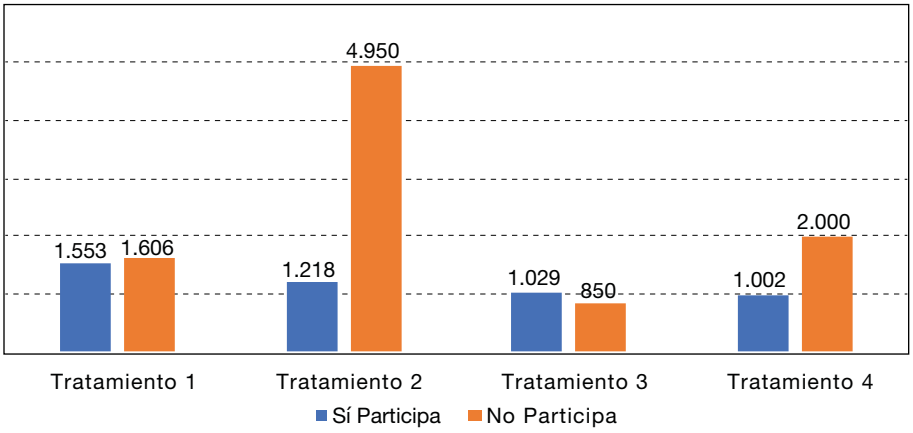
**CUADRO 3**
**DIFERENCIAS DE MEDIAS T-STUDENT DE APROPIACIÓN ENTRE TRATAMIENTOS**

| | **Grupo 1** | | **Grupo 2** | | **Grupo 3** | |
|---|---|---|---|---|---|---|
| | **T1** | **T2** | **T1** | **T3** | **T1** | **T4** |
| *Media* | 1.569,12 | 1.238,71 | 1.569,12 | 1.026,66 | 1.569,12 | 1.013,08 |
| *Error estándar* | 83,80 | 79,69 | 83,80 | 61,25 | 83,80 | 56,49 |
| *Desviación estándar* | 1.124,31 | 1.069,13 | 1.124,31 | 821,69 | 1.124,31 | 757,92 |
| *Diferencia de media* | 330,41** | | 542,46*** | | 556,04*** | |
| *Int. conf. 95% inf.* | 1.403,75 | 1.081,46 | 1.403,75 | 905,80 | 1.403,75 | 901,60 |
| *Int. conf. 95% sup.* | 1.734,48 | 1.395,96 | 1.734,48 | 1.147,51 | 1.734,48 | 1.124,55 |
| | **Grupo 4** | | **Grupo 5** | | **Grupo 6** | |
| | **T2** | **T3** | **T2** | **T4** | **T3** | **T4** |
| *Media* | 1.238,71 | 1.026,66 | 1.238,71 | 1.013,08 | 1.026,66 | 1.013,08 |
| *Error estándar* | 79,69 | 61,25 | 79,69 | 56,49 | 61,25 | 56,49 |
| *Desviación estándar* | 1.069,13 | 821,69 | 1.069,13 | 757,92 | 821,69 | 757,92 |
| *Diferencia de media* | 212,05* | | 225,63* | | 13,58 | |
| *Int. conf. 95% inf.* | 1.081,46 | 905,8 | 1.081,46 | 901,60 | 905,8 | 901,60 |
| *Int. conf. 95% sup.* | 1.395,96 | 1.147,51 | 1.395,96 | 1.124,55 | 1.147,51 | 1.124,55 |

NOTA: * $p < .05$; ** $p < .01$; *** $p < .001$
**FUENTE:** Elaboración propia.

En el Cuadro 3 se muestra que el efecto tratamiento tuvo un impacto significativo en cinco de los seis grupos iterados comparados. Se observa una disminución de la apropiación de recursos en la medida en que aumentaron las asimetrías de información. Al igual que con el soborno, no se observó variación significativa para la apropiación entre los Tratamientos 3 y 4. Sin embargo, es interesante observar que la asimetría de información no afectó únicamente el nivel de soborno entregado, sino que disminuyó las cantidades apropiadas en promedio por los concursantes, causando así un efecto bilateral en las ganancias individuales de los participantes.

**FIGURA 7**
**PROMEDIO DE PROPUESTAS EN LA CUENTA DEL CONCURSANTE (APROPIACIÓN) POR TRATAMIENTO EN LOS DIEZ PERIODOS (En $)**




**FUENTE**: Elaboración propia

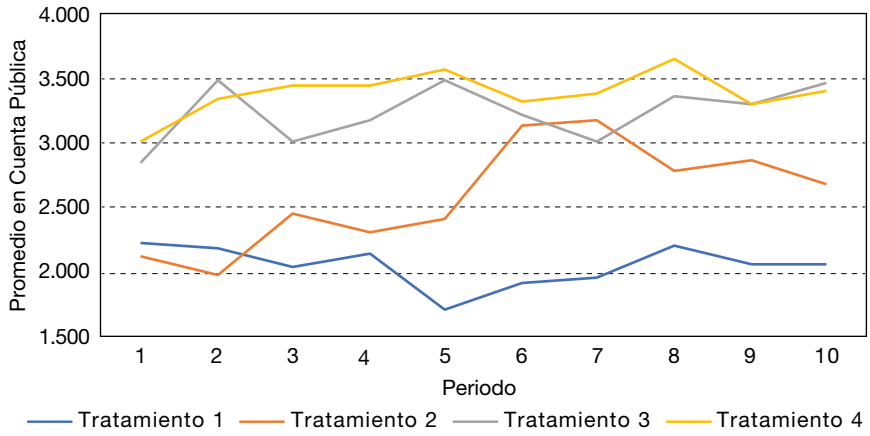
El resultado expuesto en la Figura 7 ofrece evidencia experimental que apoya la hipótesis de que las asimetrías de información, al perturbar la toma de decisiones, distorsionar las expectativas y al añadir incertidumbre en el cumplimiento de los acuerdos, generaron la suficiente inestabilidad para que el tamaño de los sobornos y la apropiación disminuyera en su conjunto. Lo más destacable es que los tratamientos en los que existió mayor asimetría de información, los concursantes a los que el Burócrata permitió participar fueron los que de menos recursos decidieron apropiarse en promedio. Un ejemplo de ello son los tratamientos 2 y 4, en donde se muestra que la media de apropiación de los que no participaron fue superior a lo apropiado por los que sí lo hicieron.

### *5.3. Cuenta Pública*

Teniendo en cuenta el diseño experimental, la estrategia Pareto de los participantes era elegir enviar todo el dinero a la Cuenta Pública, porque elegir otra opción configuraba un ejercicio de desvío de fondos colectivos a cuentas privadas, ya fuese a la Cuenta del Burócrata en forma de soborno o a la Cuenta Personal del Concursante, lo que configuraría un tipo de apropiación. Además, los fondos que se desviaran a cuentas diferentes de la Pública no serían objeto de la externalidad positiva del bien

público al no ser multiplicados por el nivel de tecnología del ganador. En la Figura 8 se observa que en el Tratamiento 1 fue donde, en promedio, menos recursos se enviaron a la Cuenta Pública, como era de esperarse, de acuerdo con los resultados evidenciados en los apartados anteriores. Por su parte, los Tratamientos 3 y 4 fueron los más eficientes en cuanto al envío de recursos a la Cuenta Pública. En resumen, los resultados experimentales muestran que los tratamientos en los que el Burócrata no podía observar directamente el soborno no generaron los incentivos suficientes para sobornarlo.

**FIGURA 8**
**PROMEDIO DE RECURSOS DESTINADO A LA CUENTA PÚBLICA**




**FUENTE**: Elaboración propia

En el Cuadro 4 se evidencia que a través de los tratamientos se observó un aumento progresivo de la cantidad enviada en promedio a la Cuenta Pública, gracias a que a su vez fueron disminuyendo los fondos que eran malversados. También se evidenció que el efecto tratamiento fue significativo en cinco de los seis grupos comparativos. En promedio se aumentó la cantidad en la Cuenta Pública de manera progresiva, pasando de 2.049 pesos en el Tratamiento 1 hasta llegar a los 3.383 pesos en el Tratamiento 4. Es evidente que las asimetrías de información relacionadas con observar el soborno y el nivel de tecnología sí impactaron las decisiones de los concursantes sobre la forma en la que debían distribuir los recursos comunes.

**CUADRO 4**

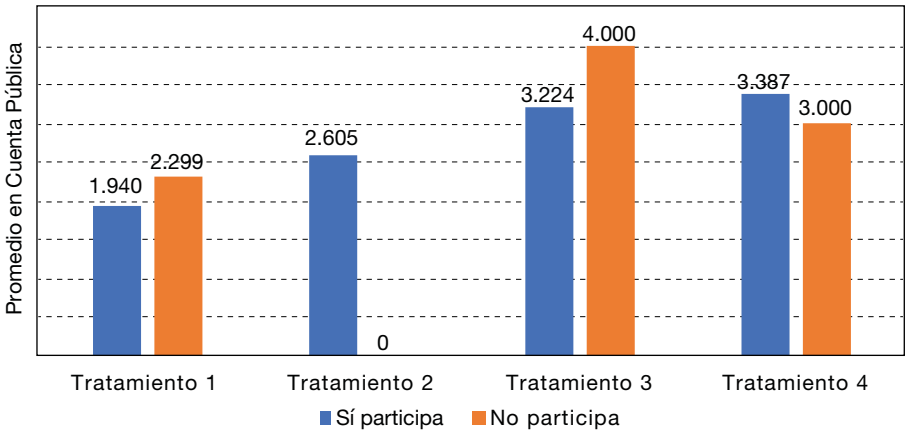
**DIFERENCIAS DE MEDIAS DE DINERO T-STUDENT ENVIADO A LA CUENTA PÚBLICA ENTRE TRATAMIENTOS**

| | **Grupo 1** | | **Grupo 2** | | **Grupo 3** | |
|---|---|---|---|---|---|---|
| | **T1** | **T2** | **T1** | **T3** | **T1** | **T4** |
| *Media* | 2.049,33 | 2.590,23 | 2.049,33 | 3.232,76 | 2.049,33 | 3.383,05 |
| *Error estándar* | 86,72 | 95,33 | 86,72 | 99,81 | 86,72 | 73,35 |
| *Desviación estándar* | 1.163,48 | 1.278,95 | 1.163,48 | 1.339,08 | 1.163,48 | 984,12 |
| *Diferencia de media* | (–540,91***) | | (–1.183,43***) | | (–1.333,72***) | |
| *Int. conf. 95% inf.* | 1.878,20 | 2.402,12 | 1.878,20 | 3.035,80 | 1.878,20 | 3.238,31 |
| *Int. conf. 95% sup.* | 2.220,45 | 2.778,34 | 2.220,45 | 3.429,71 | 2.220,45 | 3.527,80 |
| | **Grupo 4** | | **Grupo 5** | | **Grupo 6** | |
| | **T2** | **T3** | **T2** | **T4** | **T3** | **T4** |
| *Media* | 2.590,23 | 3.232,76 | 2.590,23 | 3.383,05 | 3.232,76 | 3.383,05 |
| *Error estándar* | 95,33 | 99,81 | 95,33 | 73,35 | 99,81 | 73,35 |
| *Desviación estándar* | 1.278,95 | 1.339,08 | 1.278,95 | 984,12 | 1.339,08 | 984,12 |
| *Diferencia de media* | (–642,53***) | | (–792,82***) | | (–150,29) | |
| *Int. conf. 95% inf.* | 2.402,12 | 3.035,80 | 2.402,12 | 3.238,31 | 3.035,80 | 3.238,31 |
| *Int. conf. 95% sup.* | 2.778,34 | 3.429,71 | 2.778,34 | 3527,8 | 3.429,71 | 3.527,80* |

NOTA: $p < .05$; ** $p < .01$; *** $p < .001$.
**FUENTE**: Elaboración propia.

La afirmación anterior se puede complementar con los datos de la Figura 9, en la que se evidencia que, en concordancia con los resultados anteriores, en el Tratamiento 4 se visualiza un tipo de discriminación positiva, pues el Burócrata permitió participar a los concursantes que decidieron enviar en promedio más dinero a la Cuenta Pública. La cantidad enviada por los concursantes que participaron en la subasta pasó en el Tratamiento 1 de los 1.940 a los más de 3.200 pesos enviados en los Tratamientos 3 y 4. Para terminar, se verificó experimentalmente que a medida que aumentan las asimetrías de información para el Burócrata, los incentivos para entregarle sobornos disminuyeron y, por ende, aumentaron los fondos disponibles para el grupo.

**FIGURA 9**
**PROMEDIO DE RECURSOS DESTINADO A LA CUENTA PÚBLICA POR TRATAMIENTO EN LOS DIEZ PERIODOS**
**(En $)**




**FUENTE**: Elaboración propia.

En síntesis, se puede afirmar con los resultados obtenidos que entre menor sea la capacidad de decisión que se entregue a los Burócratas, menores serán los incentivos para corromperlos. Los agentes privados no tendrán interés en sobornar a alguien que poco o nada podrá hacer a cambio de la coima entregada. Si bien en la vida real existen un sinnúmero de variables adicionales que pueden incidir en la entrega de sobornos a funcionarios públicos, la evidencia muestra que en la medida en que el poder esté desconcentrado, y menor sea la capacidad de decisión que se otorgue a cada uno de los Burócratas, más complejo será poder malversar los recursos públicos.

## 6. Conclusiones

Con la ejecución del experimento económico propuesto en el presente trabajo, se encontró que en la medida en que aumentan las asimetrías de información del Burócrata, existen menores incentivos por parte de los concursantes para entregar sobornos. Cuando el concursante conoce que el Burócrata puede observar el soborno, en un posible racionamiento de inducción hacia atrás, decide proponer coimas más altas y se apropia en promedio de más dinero, lo que ocasiona un deterioro de la cantidad de recursos públicos disponibles. Es probable que el fenómeno anterior esté sustentado en la imposibilidad que tiene el Burócrata de asumir una conducta recíproca respecto a los sobornos que envían los concursantes. De tal modo que se

puede concluir que el efecto de la asimetría de información es positivo para disminuir la malversación de recursos, entendiendo que al dificultar la negociación y aumentar la incertidumbre entre los agentes, se hace más complejo el cumplimiento de los acuerdos.

Respecto a la apropiación de recursos, la evidencia mostró que tuvo un comportamiento similar al de la entrega de sobornos a lo largo de los tratamientos. Entonces, es posible concluir que el hecho de que el Burócrata observara la apropiación tuvo un efecto positivo en la reducción de esta por parte de los concursantes. Sin embargo, este comportamiento puede deberse a razones distintas a las que explican el comportamiento de la entrega de sobornos. Por un lado, es comprensible que, si un Burócrata no puede cumplir un acuerdo debido a una asimetría de información, los concursantes no tendrán incentivos a entregarle un soborno; pero, por otro lado, si existe la posibilidad de que el Burócrata observe la cantidad apropiada por el concursante (exposición), este último adoptará una conducta prudente sobre la cantidad de dinero que decide quedarse. Por tanto, cuando el Burócrata vio la cantidad apropiada, los concursantes decidieron disminuir en promedio la cantidad de dinero enviada a la Cuenta Personal como una señal de que estaban repartiendo equitativamente el botín. En este sentido, dicha estrategia buscaba que el Burócrata les permitiera participar deduciendo que ganaría más si entraban en la subasta aquellos que decidieran apropiarse de menos. Por ende, la posibilidad de observar la apropiación por parte del Burócrata llevó a que se disminuyera en promedio la cantidad de sobornos y de apropiaciones aumentando el dinero en la Cuenta Pública.

En cuanto a la discriminación, se evidenció que el nivel de tecnología no fue una variable tenida en cuenta por parte del Burócrata para permitir el acceso de los concursantes a la subasta. Como consecuencia, se concluye que no hubo discriminación positiva en el proceso de entrada. En cambio, se observaron altos niveles de reciprocidad por parte del Burócrata con los concursantes que enviaron los sobornos más altos, pero debido al diseño del experimento, únicamente pudo asumir esta conducta cuando la información disponible se lo permitió. Por consiguiente, es posible afirmar que el poder habilitante del Burócrata no fue utilizado para permitir entrar a la subasta a los mejores concursantes, sino que lo empleó para seleccionar a los que entregaron los sobornos más altos. Lo anterior puede ser objeto de discusión, por cuanto al Burócrata también le beneficiaba elegir al mejor concursante, pero en contraposición, decidió inclinarse hacia el soborno. Por tanto, conocer el soborno lleva a que se discriminen concursantes de forma negativa, pero conocer el nivel de tecnología o la apropiación no llevan a que exista un proceso de discriminación positiva.

Con los resultados experimentales se encontró que el poder de exclusión del Burócrata en el proceso de subasta genera incentivos para la entrega de sobornos por parte de licitantes privados. Finalmente, se pudo evidenciar que la existencia del pliego de condiciones en la contratación pública (endogenizado en el experimento como la decisión del Burócrata), podría terminar siendo un arma de doble filo, debido a que se les otorgan excesivos poderes a los burócratas sobre la definición de los

criterios que deben cumplir los contratistas para entrar en las licitaciones. En este sentido, la discrecionalidad de los funcionarios públicos crea incentivos importantes para que contratistas privados entreguen sobornos en su búsqueda por acceder a los recursos del Estado. Entonces, si se desea disminuir la entrega de sobornos y adjudicar contratos públicos de forma transparente, imparcial y eficiente, no debería ser el Burócrata el encargado de establecer el pliego de condiciones habilitantes, sino que dichos requisitos deberían estar preestablecidos en la ley, de tal forma que no existiese la manera de que pudiera ser recíproco al modificar a la medida de un contratista en particular los criterios necesarios para poder entrar a la subasta.

## Referencias bibliográficas


Abbink, K., Irlenbusch, B., & Renner, E. (2002). An experimental bribery game. *Journal of Law, Economics and Organization, 18*(2), 428-454.

Aidt, T. (2003). Economic analysis of corruption: A survey. *The Economic Journal*, *113*(491), 632-652.

Albano, G., Spagnolo, G., & Zanza, M. (2008). Regulating joint bidding in public procurement. *Journal of Competition Law and Economics*, *5*(2), 335-360.

Andvig, J., & Fjeldstad, O. (2000). *Research on Corruption: A Policy Oriented Survey*. Michelsen Institute (CMI) y Norwegian Institute of International Affairs (NUPI).

Becker, G. (1968). *Crime and punishment*. The University of Chicago Press.

Binmore, K., & Klemperer, P. (2002). The Biggest Auction Ever: The Sale of the British 3G Telecom Licenses. *The Economic Journal*, *112*, 74-96.

Büchner, S., Freytag, A., González, L. G., & Güth, W. (2008). Bribery and public procurement: An experimental study. *Public Choice*, *137*(1-2), 103-117.

Chen, D., 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.

Department of Health Education and Welfare, (1979). *The Belmont Report*.

Hellman, J., Jones, G., & Kaufmann, D. (2000). *Seize the State, Seize the Day: An Empirical Analysis of State Capture and Corruption in Transition* (Mimeo). Banco Mundial.

Huang, H., & Li, Z. (2015). Procurement auctions with ex-ante endogenous bribery. *Economic Modelling*, *47*, 111-117.

Khan, M. (2004). Corruption, governance and economic development. In Jomo, K. S. & Ben Fine (Eds.), *The New Development Economics*. Tulika Press and Zed Press.

Lambsdorff, J. (2007). *Institutional economics and reform: Theory, evidence and policy.* Cambridge University Press.

Makowsky, M., & Wang, S. (2015). *Embezzlement, Whistleblowing and Organizational Architecture: An Experimental Investigation* (GMU Working Paper in Economics No 15-59).

Rose-Ackerman, S. (1975). The economics of corruption. *Journal of public economics, 4*(2), 187-203.

Transparencia Internacional. (2016). *Corruption Perceptions Index*. https://www.transparency.org/news/feature/corruption_perceptions _index_2016.

Vetter, S. (2013). Delegating decision rights for anticipated rewards as an alternative to corruption: An experiment. *European Journal of Political Economy*, *31*, 188-204.

Vickrey, W. (1961). Counterspeculation, auctions, and competitive sealed tenders. *Journal of Finance, 16*(1), 8-37.

CUADERNOS ECONÓMICOS ICE INFORMACIÓN COMERCIAL ESPAÑOLA

# Por qué donar y cuánto donar: determinantes de la generosidad en el juego del dictador

# *Why Donate, and How Much: Determinants of Generosity in the Dictator Game*

Antonio M. Espín*
Teresa García-Muñoz**
Universidad de Granada

Filippos Exadaktylos
ICTA, Universitat Autònoma de Barcelona***

***Resumen***

*Este artículo estudia el papel de las características sociodemográficas y las preferencias sociales sobre el comportamiento generoso. Utilizando una muestra (n = 792) heterogénea representativa de la ciudad de Granada y el juego del dictador se investiga qué mecanismos trabajan sobre la decisión de donar dinero a otras personas anónimas y sobre la decisión de cuánto donar. Nuestros resultados indican que los determinantes que animan a participar (donar) son diferentes de los que llevan a aumentar la participación (donar más). Estos resultados pueden ser de gran interés para las organizaciones y actividades que se nutren de la participación voluntaria de terceros.*

***Palabras clave**: juego del dictador, altruismo, preferencias sociales, donaciones, generosidad, experimento económico.*

***Clasificación JEL**: C93, C91, C70.*

***Abstract***

*This paper studies the role of individuals' sociodemographic characteristics and social preferences on generous behaviour. Using a heterogeneous sample (n = 792), representative of Granada city and the dictator game, we investigate the mechanisms underlying the decision to donate money to other anonymous people and the decision of how much to donate. Our results indicate that the determinants that encourage an individual to participate (i.e. donate) and to increase participation (donate more) are different. These results may be of great interest for organizations and activities that need the voluntary participation of third parties.*

***Keywords**: dictator game, altruism, social preferences, donations, generosity, economic experiments.*

***JEL Classification**: C93, C91, C70.*

* Departamento de Antropología Social, Universidad de Granada.

** Departamento de Métodos Cuantitativos. La autora agradece la financiación recibida por el Ministerio de Ciencia, Innovación y Universidades (PGC18-093506-B-I00).

*** Institut de Ciència i Tecnologia Ambientals–Institute of Environmental Science and Technology (ICTA, UAB).


DOI: https://doi.org/10.32796/cice.2020.99.7010
Recibido: Enero 2020 · Aceptado: Febrero 2020

## 1. Introducción

Uno de los resultados más consistentes que la economía del comportamiento ha puesto de manifiesto es que los individuos no solo se preocupan por su bienestar, también se interesan por el bienestar de los demás. Es decir, las personas, a la hora de tomar decisiones que pueden afectar a terceros, muestran preferencias sociales.

En el juego del dictador, un jugador ha de decidir cómo dividir una cierta cantidad de dinero con otro jugador en condiciones de completa libertad y anonimato. La predicción de la teoría económica «estándar» (donde la función de utilidad de los agentes considera únicamente el beneficio propio) es, obviamente, que el jugador que decide se queda con todo. Sin embargo, los resultados de múltiples experimentos muestran cómo una proporción no despreciable de sujetos es generosa: renuncia a una parte de sus ingresos para beneficiar a otros jugadores, desviándose así del mencionado «equilibrio» del juego. En Brañas-Garza *et al*. (2009) puede encontrarse una revisión de este tipo de experimentos y Engel (2011) presenta un meta-análisis de los resultados hasta la fecha.

Después de años de investigación, se tiene un conocimiento amplio sobre los entornos que hacen que los individuos se desvíen de forma sistemática del equilibrio en el juego del dictador, pero sabemos relativamente poco sobre cómo afectan las características personales de los individuos o sus preferencias sociales a esa desviación. Una de las razones principales de esta laguna tiene que ver con la homogeneidad de las muestras, mayoritariamente formadas por estudiantes, que se han utilizado en la mayoría de los experimentos. En este artículo, trataremos de arrojar luz sobre esta cuestión mediante el análisis del comportamiento de una población marcadamente heterogénea. A finales de 2010 nuestro grupo de investigación llevó a cabo un experimento de campo en el que se recogió una muestra representativa de la población adulta de la ciudad de Granada (España). El experimento consistió en una encuesta en la que se recogían, entre otros, datos socioeconómicos y sobre preferencias sociales y en la que, además, los sujetos encuestados realizaron varios juegos económicos experimentales, entre ellos el juego del dictador. La base de datos resultante de este experimento es mucho más heterogénea que la que se puede obtener con una muestra de estudiantes y nos va a permitir estudiar en profundidad el papel de las características personales sobre la generosidad en el juego del dictador.

Nuestra estrategia estadística se basa en analizar separadamente los determinantes de la decisión sobre si donar o no (es decir, de desviarse del equilibrio) y, una vez se decide donar, de la decisión sobre cuánto donar. Este aspecto novedoso de nuestro trabajo pone el foco en otra característica de las donaciones anónimas que se ha soslayado en la literatura: ¿la decisión de donar algo y la de cuánto donar tienen determinantes diferentes? (Engel, 2011). Esta no es una cuestión baladí y puede tener implicaciones importantes, por ejemplo, para las organizaciones de *crowdfunding* o de voluntariado. Entender qué factores son los que animan a las personas a participar de estas actividades y qué factores son los que les animan a aumentar su participación puede redundar en un aumento de la participación total y

en una mejora de las perspectivas de los procesos que se nutren de las donaciones desinteresadas, de tiempo, dinero, u otros bienes y servicios, de terceros.

En la siguiente sección realizamos una breve revisión bibliográfica de trabajos que se han centrado en el estudio de los determinantes del comportamiento generoso en el juego del dictador. En la sección 3 se describen las variables utilizadas en el estudio. En la sección 4 presentamos el modelo econométrico empleado y los resultados encontrados. Finalizamos presentando las conclusiones en la sección 5.

## 2. El juego del dictador

El juego del dictador lo componen dos jugadores. El jugador 1 (dictador) es invitado a dividir una cantidad específica de dinero entre él y el jugador 2 (receptor). La decisión sobre el reparto la toma exclusivamente el dictador, mientras que el jugador 2 no tiene derecho a reclamar sobre la cantidad obtenida.

Según el principio de maximización del beneficio personal, la solución teórica del juego es simple: «el dictador se queda con todo, el receptor no recibe nada». Sin embargo, en los múltiples experimentos realizados del juego del dictador se observan sistemáticamente valores distintos de cero para la cantidad donada al jugador 2. En un meta-análisis sobre el juego del dictador realizado por Engel (2011) se encontró que el 36.11 % de los dictadores no compartieron nada con los receptores, el 16.74 % eligieron compartir la mitad y el 5.44 % eligieron donar todo a los receptores. En experimentos estándar de laboratorio las donaciones oscilan entre cero y la mitad de la cantidad disponible a repartir. Un dictador típico transfiere un promedio de 20-30 % y muy rara vez supera la mitad de la «tarta» (Camerer, 2003), aunque existen diferencias visibles entre culturas (Henrich *et al*., 2006, 2010).

Pero, ¿qué mueve a los individuos a comportarse de forma generosa? Considerando que cualquier donación positiva se puede interpretar como una muestra de generosidad, el juego del dictador es un mecanismo cuasi-perfecto para estudiar el comportamiento generoso: es un juego simple donde los jugadores no se enfrentan a complejos problemas de maximización, además no hay espacio para la reciprocidad ni para cualquier otra motivación estratégica. A continuación se resumen algunas de las investigaciones realizadas para tratar de comprender por qué los sujetos se desvían sistemáticamente del equilibrio (*donación = 0*).

En primer lugar, resaltamos una serie de trabajos que realizan nuevos diseños del juego incorporando información sobre el receptor. Ejemplos de estos tratamientos son Eckel y Grossman (1998), donde el jugador 2 es la Cruz Roja de la localidad; Burnham (2003), que proporciona fotografías de los receptores a los dictadores; Charness y Gneezy (2001), que proporcionan a los dictadores los apellidos del receptor; Brañas-Garza (2006), donde se les informa que los receptores son africanos y que recibirán el dinero en forma de medicinas; y Frohlich *et al*. (2001), donde los dictadores pueden ver a los receptores un momento antes del experimento. En todos estos trabajos se ha encontrado que la falta de anonimato disminuye el número

de sujetos completamente egoístas (*donación = 0*) e incrementa el porcentaje de individuos generosos con el receptor. Hoffman *et al*. (1996) proponen una serie de tratamientos alternativos que van sistemáticamente aislando al sujeto experimental (aumentando el anonimato), y generando así «distancia social». Esta distancia puede hacer que el sujeto se sienta más libre para llevar a cabo la acción que considere oportuna sin tener que soportar el coste de ser reconocido como egoísta. Como resultado, las transferencias de los dictadores en los experimentos se van reduciendo a medida que el grado de anonimato aumenta.

Otra línea de investigación ha puesto de manifiesto la relación existente entre el comportamiento altruista y el capital social (el grado de integración social) de los individuos. Como muestran Brañas-Garza *et al*. (2010): i) los sujetos menos altruistas (los que no donan nada en el juego del dictador) suelen tener un papel poco importante en la red social (tienen menos conexiones en la red social, son menos centrales) y ii) los sujetos con gran capital social (los más centrales) son más altruistas y, además, más aversos a la desigualdad (Brañas-Garza *et al*., 2012).

Como hemos visto, gran parte de los esfuerzos por comprender por qué los sujetos se desvían sistemáticamente del equilibrio se han centrado en los efectos de diversos tratamientos en lugar de en el análisis de cómo las características personales pueden influir en el comportamiento. Por tanto, no es sorprendente que después de muchos años de intensa investigación sepamos muy poco acerca de la interrelación entre las características personales y las donaciones en el juego del dictador —con pocas excepciones, como es el género, donde existen resultados no concluyentes (Croson y Gneezy, 2009), aunque con tendencia a observarse una mayor generosidad de las mujeres (Brañas-Garza *et al*., 2018)—. Este es el objetivo del trabajo, estudiar qué características personales están relacionadas con el comportamiento generoso. La heterogeneidad y representatividad de los datos utilizados nos ofrecen la oportunidad de obtener resultados robustos y generalizables.

## 3. Variables en el estudio

Como se ha comentado, a finales de 2010 se llevó a cabo la recogida de una muestra representativa de la población adulta de la ciudad de Granada mediante muestreo aleatorio estratificado por distritos con afijación proporcional. Un total de 216 estudiantes universitarios, agrupados en 108 parejas, realizaron la recogida de datos en los domicilios resultantes del proceso de muestreo. Estos estudiantes estaban matriculados en un curso sobre experimentos de campo que se llevó a cabo en otoño de 2010 en la Facultad de Ciencias Económicas y Empresariales de la Universidad de Granada. Con antelación a la recogida de datos los estudiantes tuvieron 10 horas de formación sobre metodología de realización de experimentos de campo, encuestas y procedimientos de muestreo. La recogida de datos formó parte de su formación y evaluación en el curso. Los estudiantes no recibieron ninguna prestación monetaria. El proceso fue controlado en tiempo real por los investigadores principales mediante

(i) una plataforma web creada al efecto donde los encuestadores debían reportar los datos a medida que se iban recogiendo y (ii) llamadas telefónicas a encuestados elegidos aleatoriamente.

La muestra inicial estaba formada por un total de 835 individuos con edades comprendidas entre los 16 y los 91 años de edad (para más información sobre la muestra y la recogida de datos, ver Exadaktylos *et al*., 2013). El tamaño muestral se ha reducido en este estudio por los valores perdidos encontrados en algunas de las variables. De esta forma, la muestra utilizada en el estudio la forman 792 individuos.

Los sujetos encuestados contestaron un amplio cuestionario con preguntas sociodemográficas, de capital social y sobre preferencias sociales. En la última parte de la encuesta realizaron tres juegos experimentales con dinero real; uno de ellos fue el juego del dictador (los otros juegos fueron el ultimátum y el juego de la confianza; ver Exadaktylos *et al*., 2013). Los juegos se realizaron bajo condiciones de anonimato: los participantes respondían marcando sus decisiones en una hoja sin ser vistos por los encuestadores, luego la introducían en un sobre y finalmente lo cerraban. Una vez recogidos todos los datos, los participantes fueron emparejados aleatoriamente de forma anónima para calcular los pagos resultantes de los juegos. Uno de cada 10 participantes fue seleccionado aleatoriamente para recibir el pago correspondiente a uno de los juegos, elegido también aleatoriamente. Como es típico en los experimentos económicos, toda esta información era conocida por los participantes antes de tomar sus decisiones.

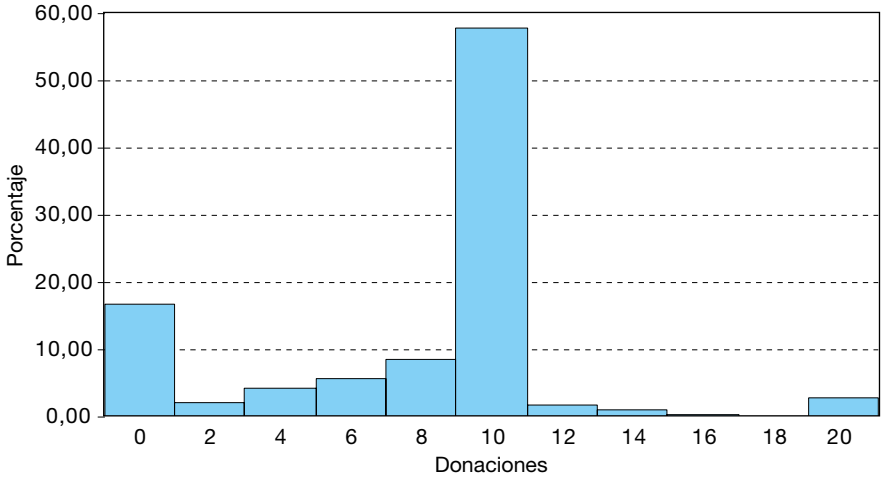
### *3.1. Variable dependiente: donaciones*

El enunciado del juego del dictador era:

> «En esta decisión te damos 20 euros para que los repartas entre la otra persona y tú. De estos 20 euros, tú puedes mandarle a esa persona la parte que tú quieras, es decir, puedes no mandarle nada, todo o mandarle una parte. Obviamente, la parte que no le envíes es directamente para ti. ¿Qué cantidad le envías a esa persona?: tienes que rodear con un círculo el número de euros que le MANDAS a la otra persona en la casilla que está en AZUL, solo pueden ser números pares (0, 2, 4,…,20)».

Como vemos, el participante debía tomar una decisión dentro de un rango de 11 posibilidades (unidad de medida = 2 euros). La donación media de los 792 individuos incluidos en el estudio fue de 7,87 euros (sobre 20, esto es, un 39 %) con una desviación típica de 4,35 euros. El Gráfico 1 muestra la distribución de las donaciones.

**GRÁFICO 1**
**DISTRIBUCIÓN DE DONACIONES EN EL JUEGO DEL DICTADOR**




**FUENTE:** Elaboración propia.

Como podemos observar, la situación más habitual es el reparto igualitario (10 euros; 57,83 % de los participantes), siendo la segunda donación más frecuente la de 0 euros (16,54 %). Comparándolos con los resultados mencionados en la introducción (obtenidos mayoritariamente usando muestras de estudiantes universitarios), nuestros participantes muestran un comportamiento relativamente generoso, lo que es habitual en experimentos de campo como el nuestro (ver, por ejemplo, Henrich *et al*., 2006, 2010).

### *3.2. Variables explicativas*

En esta sección se presentan las variables explicativas incluidas en el estudio. Se han distinguido tres grupos de variables: sociodemográficas, sobre preferencias sociales y de capital social.

#### 3.2.1. Variables sociodemográficas

Las variables sociodemográficas utilizadas son:

– Edad, como variable continua, y su cuadrado para tener en cuenta posibles efectos no lineales.
– Hombre, variable dicotómica que toma el valor 1 si el individuo es hombre, 0 en otro caso.
– Estado civil, variable categórica dividida en: soltero/a (categoría de referencia), casado/a o conviviendo en pareja, divorciado/a o separado/a y viudo/a.

– Hijos, variable dicotómica que toma el valor 1 si el individuo tiene hijos, 0 en otro caso.
– Educación, variable categórica con 4 categorías: sin estudios (categoría de referencia), con estudios primarios, con estudios secundarios, con estudios universitarios.
– Situación laboral, variable categórica con 5 categorías: desempleado/a, trabajador/a por cuenta ajena (categoría de referencia), trabajador/a del sector público, trabajador/a por cuenta ajena, pensionista.
– No creyente, variable dicotómica que toma el valor 1 si el individuo es no creyente (religión), 0 en otro caso.
– Ingresos del hogar, variable categórica con cuatro categorías correspondientes a los cuatro cuartiles en los que se divide la muestra según los ingresos del hogar. La categoría de referencia es el cuarto cuartil, que incluye a las familias con mayores ingresos.
– Contribución alta, variable dicotómica que toma el valor 1 si el individuo contribuye al menos a la mitad del total de ingresos del hogar, 0 en otro caso.

### 3.2.2. Preferencias sociales

Decimos que un sujeto tiene preferencias sociales cuando a la hora de maximizar su bienestar subjetivo (su utilidad), además de sus propios beneficios, tiene en cuenta los beneficios de los demás. Una definición más formal se puede encontrar en Fehr y Fischbacher (2002): «Una persona exhibe preferencias sociales cuando no solo se preocupa por los recursos que le son asignados, sino que también toma en consideración los recursos asignados a agentes de referencia relevantes para él»[1]. Algunas de las preferencias sociales que estos autores destacan, y para las que disponemos de medidas en nuestra encuesta, son:

– *Aversión a la desigualdad:* se presenta cuando el individuo quiere lograr una distribución igualitaria de los recursos. Esto significa que, por un lado, los individuos con aversión a la desigualdad son altruistas y desearán incrementar los beneficios de los otros cuando estos están por debajo de lo que se considera equitativo. Y, por otro lado, también sienten envidia y estarán motivados por reducir los beneficios de los otros cuando consideran que exceden el punto equitativo. Siguiendo el modelo de Fehr-Schmidt (1999) diseñamos dos preguntas sobre aversión a la desigualdad. En ellas se pide al individuo que valore las siguientes afirmaciones en una escala de 1 a 7 donde 1 significa *totalmente en desacuerdo*, 7 *totalmente de acuerdo* y el punto neutral es el 4:
  Aversión a la desigualdad $\alpha$ del modelo de Fehr-Schmidt (Envidia): «No me preocupa cuánto dinero tengo, lo que me preocupa es que otros tienen más que yo».

[1] Traducción propia.

Aversión a la desigualdad β del modelo de Fehr-Schmidt (Compasión): «No me preocupa cuánto dinero tengo, lo que me preocupa es que otros tienen menos que yo». Según nuestro conocimiento, estas preguntas sobre aversión a la desigualdad (basadas en el modelo Fehr-Schmidt) se incluyen en un cuestionario por primera vez en esta encuesta. Espín *et al*. (2018) presentan el primer análisis publicado sobre estas medidas, mostrando que la envidia correlaciona negativamente y la compasión positivamente con la satisfacción con la vida de los individuos. Brañas-Garza *et al*. (2020) presenta un análisis más detallado de estas medidas.

– *Reciprocidad*: un individuo recíproco responde de forma amable ante acciones amables (reciprocidad positiva) y de forma hostil ante acciones hostiles (reciprocidad negativa). Para medir las reciprocidades positiva y negativa utilizamos dos preguntas del cuestionario en las que se pedía responder si se estaba de acuerdo o no con las siguientes afirmaciones en la misma escala de 1 a 7.
  Reciprocidad negativa: «Soy amable y simpático sólo si los demás se comportan bien conmigo, de lo contrario: ojo por ojo».
  Reciprocidad positiva: «Si alguien me hace el favor de prestarme dinero, siento que debería devolverle algo más de lo prestado».
  Estas preguntas sobre reciprocidad se han obtenido del cuestionario *Personal Norm of Reciprocity* desarrollado por Perugini *et al*. (2003).
– *Altruismo:* es una forma incondicional de bondad**.** Se presenta cuando el individuo valora positivamente los recursos de los otros independientemente de «si se lo merecen» o no. Un individuo altruista nunca emprenderá una acción que haga decrecer el pago de otras personas. Para medir el altruismo incondicional se preguntó al individuo que valorase en la misma escala de 1 a 7 la siguiente afirmación: «Ayudaría a un conocido que sé que no lo haría por mí».

### 3.2.3. Variables de capital social

Como medidas del capital social se han elegido dos cuestiones que aparecen en la *World Values Survey*[2] y que representan la opinión que tienen los sujetos entrevistados sobre la gente que les rodea:

– *Confianza*, es una variable dicotómica que responde a la siguiente pregunta: «En general, ¿crees que se puede tener confianza en la mayoría de la gente, o que se debe ser muy prudente al relacionarse con la gente?». La variable toma el valor 1 si la respuesta es «Se puede tener confianza en la mayoría de la gente» y el valor 0 si es «Se debe ser muy prudente al relacionarse con la gente».
– *Justicia*, es una variable dicotómica basada en la siguiente cuestión: «¿Crees que la mayoría de la gente intentaría aprovecharse de ti, si tuvieran la oportunidad, o tratarían de ser justos?». La variable toma el valor 1 si el entrevistado responde «La mayoría de la gente trataría de ser justa» y el valor 0 si contesta «La mayoría de la gente intentaría aprovecharse de ti».

[2] www.worldvaluessurvey.com

Los estadísticos descriptivos de todas las variables utilizadas en el estudio se encuentran en el Cuadro 1.

**CUADRO 1**

**ESTADÍSTICOS DESCRIPTIVOS DE LAS VARIABLES (n = 792)**

| Variable | Media | Desv. típica | Min | Max |
|---|---|---|---|---|
| Donación | 7,871 | 4,354 | 0 | 20 |
| **Variables sociodemográficas** | | | | |
| Edad | 37,841 | 17,158 | 16 | 89 |
| Hombre | 0,458 | 0,499 | 0 | 1 |
| *Estado civil* | | | | |
| Soltero | 0,511 | 0,250 | 0 | 1 |
| Casado o conviviendo | 0,404 | 0,491 | 0 | 1 |
| Divorciado o separado | 0,042 | 0,200 | 0 | 1 |
| Viudo | 0,043 | 0,203 | 0 | 1 |
| Hijos | 0,446 | 0,497 | 0 | 1 |
| *Educación* | | | | |
| Sin educación | 0,038 | 0,037 | 0 | 1 |
| Primaria | 0,109 | 0,311 | 0 | 1 |
| Secundaria | 0,306 | 0,461 | 0 | 1 |
| Universidad | 0,547 | 0,498 | 0 | 1 |
| *Estado laboral* | | | | |
| Trabajador por cuenta ajena | 0,207 | 0,164 | 0 | 1 |
| Desempleado | 0,472 | 0,500 | 0 | 1 |
| Trabajador del sector público | 0,128 | 0,334 | 0 | 1 |
| Trabajador por cuenta propia | 0,081 | 0,273 | 0 | 1 |
| Pensionista | 0,112 | 0,316 | 0 | 1 |
| *Ingresos* | | | | |
| Primer cuartil | 0,330 | 0,470 | 0 | 1 |
| Segundo cuartil | 0,176 | 0,381 | 0 | 1 |
| Tercer cuartil | 0,254 | 0,435 | 0 | 1 |
| Cuarto cuartil | 0,240 | 0,182 | 0 | 1 |
| Contribución alta | 0,426 | 0,495 | 0 | 1 |
| No creyente | 0,312 | 0,464 | 0 | 1 |
| **Variables preferencias sociales** | | | | |
| Envidia | 2,433 | 1,787 | 1 | 7 |
| Compasión | 4,073 | 2,055 | 1 | 7 |
| Reciprocidad positiva | 3,612 | 2,266 | 1 | 7 |
| Reciprocidad negativa | 3,808 | 2,158 | 1 | 7 |
| Altruismo | 4,843 | 1,911 | 1 | 7 |
| **Variable capital social** | | | | |
| Confianza | 0,246 | 0,431 | 0 | 1 |
| Justicia | 0,455 | 0,498 | 0 | 1 |

**FUENTE**: Elaboración propia.

## 4. Modelo econométrico y resultados

A la hora de estudiar el comportamiento generoso a través del juego del dictador se debe tener en cuenta el proceso de decisión al que se enfrenta el sujeto. Tal vez el estudio de los determinantes de la cantidad donada no sea el camino correcto a seguir dado que habrá un alto porcentaje de sujetos que no estén dispuestos a donar nada bajo ninguna circunstancia. Parece más coherente que el proceso de decisión se lleve a cabo en dos pasos: primero los individuos deciden si donar o no, y una vez que han decidido donar eligen qué cantidad de dinero dar. Para explorar esta posibilidad se han utilizado modelos Hurdle (Cragg, 1971; Moffat, 2016). Estos modelos permiten que las decisiones sobre donar y sobre qué cantidad donar se originen por mecanismos diferentes, permitiendo así distintos determinantes. En primer lugar, se estudian los determinantes de que el sujeto decida donar algo mediante una regresión probit y, en una segunda etapa, sabiendo que el individuo ha decidido donar se estudian qué características determinan la cantidad donada mediante un modelo de regresión normal truncado.

Los resultados se muestran en los Cuadros 2 y 3. En el Cuadro 2 se presentan los determinantes que están asociados a la decisión de donar y en el Cuadro 3 los determinantes de la cantidad donada, una vez que el sujeto decide donar. En ambos cuadros aparecen cinco modelos en sendas columnas. En el modelo (1) se consideran solo las variables sociodemográficas, mientras que en los modelos (2-4) se incluyen sucesivamente las tres preferencias sociales consideradas (aversión a la desigualdad, reciprocidad y altruismo) y en el último modelo se analizan también las variables de capital social.

El primer resultado a resaltar es, como se verá a continuación, que los factores asociados a la decisión de donar y a la cantidad donada no son los mismos.

Comenzamos comentando los resultados encontrados para el modelo que describe la decisión de donar en el juego del dictador (Cuadro 2).

En relación a las *variables sociodemográficas*, se observa que los hombres muestran menor probabilidad de decidir donar (aunque solo cuando no se controla el altruismo ni las variables de capital social). Por otro lado, es más probable que los divorciados/as decidan donar en comparación con los solteros. También se observa que los individuos de familias con menores ingresos están más predispuestos a donar; concretamente las familias en los cuartiles 1 y 3 (no significativo para el cuartil 2) tienen mayor probabilidad de realizar una donación positiva en comparación con las familias en el cuartil más alto (familias más ricas). Este resultado es cercano al encontrado por Erkal *et al*. (2011) donde comprobaron que los individuos con ingresos más altos eran más reacios a donar que los individuos con ingresos moderados. Por último, una mayor aportación a los ingresos del hogar es un determinante positivo a la hora de decidir donar, pero solo cuando no se tienen en cuenta ni el altruismo y ni las variables de capital social.

**CUADRO 2**

**PRIMERA ETAPA DEL MODELO HURDLE: DETERMINANTES DE LA DECISIÓN DE DONAR**

| Variables | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Variables sociodemográficas** | | | | | |
| Edad | –0,024<br>(0,024) | –0,024<br>(0,024) | –0,030<br>(0,024) | –0,029<br>(0,024) | –0,031<br>(0,024) |
| Edad al cuadrado | 0,0002<br>(0,0003) | 0,0002<br>(0,0003) | 0,0003<br>(0,0002) | 0,0003<br>(0,0002) | 0,0003<br>(0,0002) |
| Hombre | –0,261**<br>(0,123) | –0,218*<br>(0,122) | –0,187<br>(0,122) | –0,170<br>(0,122) | –0,181<br>(0,121) |
| *Estado civil* | | | | | |
| Soltero | ref. | ref. | ref. | ref. | ref. |
| Casado o conviviendo | 0,057<br>(0,221) | 0,037<br>(0,222) | 0,045<br>(0,228) | 0,061<br>(0,224) | 0,052<br>(0,227) |
| Divorciado o separado | 0,842**<br>(0,428) | 0,832*<br>(0,439) | 0,857**<br>(0,423) | 0,819*<br>(0,422) | 0,766*<br>(0,420) |
| Viudo | –0,295<br>(0,353) | –0,324<br>(0,370) | –0,344<br>(0,372) | –0,374<br>(0,373) | –0,396<br>(0,375) |
| Hijos | 0,216<br>(0,180) | 0,247<br>(0,183) | 0,256<br>(0,179) | 0,252<br>(0,182) | 0,252<br>(0,183) |
| *Educación* | | | | | |
| Sin educación | ref. | ref. | ref. | ref. | ref. |
| Primaria | 0,597<br>(0,373) | 0,588<br>(0,374) | 0,597<br>(0,391) | 0,546<br>(0,380) | 0,546<br>(0,376) |
| Secundaria | 0,444<br>(0,342) | 0,407<br>(0,349) | 0,368<br>(0,362) | 0,326<br>(0,360) | 0,300<br>(0,366) |
| Universidad | 0,401<br>(0,343) | 0,373<br>(0,351) | 0,306<br>(0,362) | 0,267<br>(0,355) | 0,222<br>(0,357) |
| *Estado laboral* | | | | | |
| Trabajador por cuenta ajena | ref. | ref. | ref. | ref. | ref. |
| Desempleado | 0,045<br>(0,146) | 0,044<br>(0,146) | 0,035<br>(0,146) | –0,005<br>(0,152) | 0,006<br>(0,154) |
| Trabajador del sector público | 0,256<br>(0,193) | 0,226<br>(0,199) | 0,181<br>(0,209) | 0,164<br>(0,209) | 0,175<br>(0,207) |
| Trabajador por cuenta propia | 0,279<br>(0,232) | 0,219<br>(0,226) | 0,203<br>(0,229) | 0,183<br>(0,227) | 0,170<br>(0,229) |

**CUADRO 2** ***(Continuación)***

**PRIMERA ETAPA DEL MODELO HURDLE: DETERMINANTES DE LA DECISIÓN DE DONAR**

| Variables | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Pensionista | 0,116<br>(0,356) | 0,061<br>(0,354) | 0,014<br>(0,343) | 0,003<br>(0,337) | 0,019<br>(0,324) |
| *Ingresos* | | | | | |
| Primer cuartil | 0,435***<br>(0,166) | 0,450***<br>(0,167) | 0,465***<br>(0,165) | 0,502***<br>(0,169) | 0,506***<br>(0,170) |
| Segundo cuartil | 0,216<br>(0,195) | 0,232<br>(0,196) | 0,242<br>(0,195) | 0,241<br>(0,195) | 0,221<br>(0,196) |
| Tercer cuartil | 0,386**<br>(0,178) | 0,409**<br>(0,178) | 0,424**<br>(0,177) | 0,430**<br>(0,179) | 0,430**<br>(0,179) |
| Cuarto cuartil | ref. | ref. | ref. | ref. | ref. |
| Contribución alta | 0,219*<br>(0,124) | 0,207*<br>(0,122) | 0,201*<br>(0,120) | 0,167<br>(0,120) | 0,177<br>(0,123) |
| No creyente | –0,180<br>(0,129) | –0,182<br>(0,134) | –0,194<br>(0,131) | –0,168<br>(0,131) | –0,197<br>(0,130) |
| **Variables preferencias sociales** | | | | | |
| Envidia | | –0,047*<br>(0,027) | –0,039<br>(0,027) | –0,034<br>(0,027) | –0,039<br>(0,027) |
| Compasión | | 0,095***<br>(0,035) | 0,090***<br>(0,035) | 0,077**<br>(0,035) | 0,073**<br>(0,036) |
| Reciprocidad positiva | | | 0,038<br>(0,028) | 0,026<br>(0,029) | 0,029<br>(0,029) |
| Reciprocidad negativa | | | –0,081***<br>(0,027) | –0,068**<br>(0,027) | –0,062**<br>(0,028) |
| Altruismo | | | | 0,083***<br>(0,031) | 0,080**<br>(0,031) |
| **Variable de capital social** | | | | | |
| Confianza | | | | | 0,036<br>(0,133) |
| Justicia | | | | | 0,269**<br>(0,117) |
| Constante | 0,682<br>(0,518) | 0,489<br>(0,569) | 0,863<br>(0,605) | 0,536<br>(0,627) | 0,505<br>(0,634) |
| Observaciones | 792 | 792 | 792 | 792 | 792 |

NOTA: Errores robustos entre paréntesis. *** $p < 0.01$. ** $p < 0.05$. * $p < 0.1$.
**FUENTE:** Elaboración propia.

**CUADRO 3**

**SEGUNDA ETAPA DEL MODELO HURDLE: DETERMINANTES DE LA CANTIDAD DONADA (ENTRE QUIENES DONAN)**

| Variables | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| **Variables sociodemográficas** | | | | | |
| Edad | 0,096** (0,044) | 0,094** (0,044) | 0,085* (0,044) | 0,085* (0,044) | 0,089** (0,042) |
| Edad al cuadrado | –0,0009** (0,0004) | –0,0009** (0,0004) | –0,0008** (0,0004) | –0,0008** (0,0004) | –0,0009** (0,0004) |
| Hombre | 0,382* (0,218) | 0,418* (0,218) | 0,459** (0,218) | 0,460** (0,218) | 0,459** (0,220) |
| *Estado civil* | | | | | |
| Soltero | ref. | ref. | ref. | ref. | ref. |
| Casado o conviviendo | –0,031 (0,459) | –0,075 (0,460) | –0,079 (0,464) | –0,076 (0,466) | –0,065 (0,469) |
| Divorciado o separado | 0,202 (0,589) | 0,208 (0,578) | 0,167 (0,580) | 0,172 (0,577) | 0,206 (0,592) |
| Viudo | –0,357 (0,519) | –0,413 (0,511) | –0,486 (0,517) | –0,485 (0,517) | –0,510 (0,510) |
| Hijos | 0,301 (0,384) | 0,309 (0,387) | 0,298 (0,390) | 0,298 (0,389) | 0,246 (0,388) |
| *Educación* | | | | | |
| Sin educación | ref. | ref. | ref. | ref. | ref. |
| Primaria | 0,689 (0,699) | 0,586 (0,685) | 0,592 (0,673) | 0,591 (0,672) | 0,563 (0,668) |
| Secundaria | 0,859 (0,687) | 0,804 (0,675) | 0,789 (0,660) | 0,790 (0,659) | 0,708 (0,644) |
| Universidad | 0,734 (0,648) | 0,699 (0,629) | 0,649 (0,610) | 0,649 (0,609) | 0,599 (0,620) |
| *Estatus laboral* | | | | | |
| Trabajador por cuenta ajena | ref. | ref. | ref. | ref. | ref. |
| Desempleado | –0,787** (0,332) | –0,757** (0,328) | –0,778** (0,325) | –0,779** (0,327) | –0,789** (0,329) |
| Trabajador del sector público | –0,291 (0,352) | –0,323 (0,349) | –0,333 (0,345) | –0,333 (0,346) | –0,314 (0,361) |
| Trabajador por cuenta propia | –0,775* (0,454) | –0,778* (0,450) | –0,768* (0,455) | –0,767* (0,455) | –0,724 (0,464) |

**CUADRO 3** ***(Continuación)***

**SEGUNDA ETAPA DEL MODELO HURDLE: DETERMINANTES DE LA CANTIDAD DONADA (ENTRE QUIENES DONAN)**

| Variables | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| Pensionista | –0,148<br>(0,548) | –0,205<br>(0,547) | –0,247<br>(0,560) | –0,246<br>(0,558) | –0,202<br>(0,586) |
| *Ingresos* | | | | | |
| Primer cuartil | 0,196<br>(0,335) | 0,171<br>(0,332) | 0,165<br>(0,334) | 0,169<br>(0,333) | 0,153<br>(0,338) |
| Segundo cuartil | –0,172<br>(0,329) | –0,162<br>(0,320) | –0,153<br>(0,328) | –0,153<br>(0,327) | –0,150<br>(0,334) |
| Tercer cuartil | –0,101<br>(0,295) | –0,069<br>(0,286) | –0,059<br>(0,285) | –0,058<br>(0,285) | –0,078<br>(0,289) |
| Cuarto cuartil | ref. | ref. | ref. | ref. | ref. |
| Contribución alta | –0,507*<br>(0,296) | –0,487*<br>(0,293) | –0,502*<br>(0,290) | –0,505*<br>(0,292) | –0,522*<br>(0,290) |
| No creyente | –0,273<br>(0,255) | –0,237<br>(0,250) | –0,257<br>(0,249) | –0,256<br>(0,248) | –0,242<br>(0,250) |
| **Variables preferencias sociales** | | | | | |
| Envidia | | 0,031<br>(0,062) | 0,048<br>(0,064) | 0,049<br>(0,064) | 0,050<br>(0,066) |
| Compasión | | 0,120**<br>(0,050) | 0,117**<br>(0,051) | 0,116**<br>(0,050) | 0,111**<br>(0,049) |
| Reciprocidad positiva | | | 0,027<br>(0,047) | 0,026<br>(0,049) | 0,028<br>(0,049) |
| Reciprocidad negativa | | | –0,099<br>(0,061) | –0,098<br>(0,063) | –0,104<br>(0,068) |
| Altruismo | | | | 0,009<br>(0,063) | 0,008<br>(0,062) |
| **Variable capital social** | | | | | |
| Confianza | | | | | –0,312<br>(0,317) |
| Justicia | | | | | 0,143<br>(0,227) |
| Constante | 7,133***<br>(1,041) | 6,634***<br>(1,044) | 7,099***<br>(0,991) | 7,057***<br>(1,009) | 7,110***<br>(1,015) |
| Observaciones | 792 | 792 | 792 | 792 | 792 |

NOTA: Errores robustos entre paréntesis. *** $p < 0.01$. ** $p < 0.05$. * $p < 0.1$.
**FUENTE:** Elaboración propia.

Respecto a las *preferencias sociales*, observamos que el componente β (Compasión) del modelo de Fehr-Schmidt está positivamente relacionado con la decisión de donar. Aunque el diseño del juego del dictador elimina cualquier elemento de reciprocidad de la interacción, encontramos que los sujetos con mayor reciprocidad negativa tienen menor probabilidad de donar. Sería interesante desentrañar en futuras investigaciones por qué y a través de qué canales la reciprocidad puede tener efecto sobre las donaciones anónimas en el juego del dictador. Finalmente, las personas más altruistas tienen mayor probabilidad de donar en el juego.

Por último, respecto a las variables de *capital social*, las personas que consideran que el sentido de la justicia está extendido entre la sociedad son más proclives a donar.

A continuación se comentan los resultados encontrados para las regresiones que modelizan la cantidad donada, una vez se ha decidido donar (Cuadro 3).

En relación a las *variables sociodemográficas*, la variable edad es significativa y tiene un efecto cuadrático (forma de U invertida) a la hora de explicar la cantidad donada. Según los coeficientes estimados, los 49 años es aproximadamente la edad a la que más cantidad se dona. Este resultado es coherente con lo encontrado en Engel (2011): el valor modal entre personas de mediana edad es el reparto equitativo. Los hombres, una vez que deciden donar (a lo que son ligeramente menos proclives, según los resultados anteriores), donan más cantidad que las mujeres. El estatus laboral también afecta a la cantidad donada: desempleados y trabajadores por cuenta propia donan menos cantidad que los trabajadores por cuenta ajena. Y, por último, la contribución del individuo a los ingresos del hogar está asociada negativamente con la cantidad donada; si la contribución del individuo es alta menos cantidad donará.

Respecto a los dos bloques restantes, *preferencias sociales* y *capital social*, la única variable que aparece positivamente asociada a la cantidad donada es el componente β del modelo de Fehr-Schmidt (Compasión).

En resumen, como se ha visto, los factores asociados a la decisión de donar y la cantidad donada no son los mismos. Incluso hay variables que afectan significativamente a ambas decisiones pero en direcciones opuestas. Mientras que los hombres tienen ligeramente menos probabilidad de realizar una donación, cuando deciden hacerla donan mayores cantidades que las mujeres. Este efecto contradictorio del género en las donaciones del juego del dictador también fue resaltado por Andreoni y Vesterlund (2001) y podría explicar por qué los estudios anteriores, que usan la donación total en el juego como medida de generosidad (sin diferenciar las dos etapas analizadas aquí), obtienen resultados mixtos. La responsabilidad sobre los ingresos del hogar también se asocia con distinto signo en ambas decisiones. Mientras que una contribución alta a los ingresos está asociada positivamente con la probabilidad de donar, su relación con la cantidad donada es negativa.

Finalmente, solo la compasión del modelo Fehr-Schmidt predice significativamente ambas decisiones en la misma dirección: las personas más compasivas son más proclives a donar y, además, donan cantidades mayores. Esto concuerda con los resultados de Espín *et al*. (2018).

## 5. Conclusiones

El juego del dictador es un entorno perfecto para estudiar el comportamiento generoso. La gran mayoría de los estudios publicados sobre este tema se han realizado utilizando muestras formadas por estudiantes universitarios. En nuestro trabajo, la heterogeneidad y representatividad de la muestra de la que disponemos, formada no solo por estudiantes, nos permite aportar a la literatura resultados sobre el papel de las características sociodemográficas, aparte de las preferencias sociales, en el comportamiento generoso. Otra característica poco habitual en los estudios sobre el comportamiento generoso es la metodología econométrica que se ha utilizado. Mediante modelos Hurdle se ha puesto de manifiesto que los mecanismos que trabajan sobre la decisión de donar y sobre la decisión de qué cantidad donar no siempre están asociados a los mismos determinantes. En esta línea, dado que la literatura anterior no ha analizado en profundidad las donaciones como un proceso de decisión en dos etapas (Engel, 2011), los resultados que se presentan aquí son novedosos, aunque en algunos casos sean coherentes con lo que se ha encontrado en estudios anteriores.

La decisión de donar está positivamente asociada a las siguientes características sociodemográficas: ser mujer, ser una persona divorciada (en comparación con soltera), pertenecer a familias con ingresos en el primer o el tercer cuartil (en comparación con las familias más ricas) y una alta contribución (mitad o más) a los ingresos del hogar.

Respecto a las preferencias sociales, los sujetos con mayor aversión a la desigualdad, en su componente de compasión (aversión a la desigualdad «ventajosa»), tienen mayor propensión a decidir donar. Es decir, se espera que individuos preocupados por el bienestar de los que se encuentran en peor situación que ellos no se queden con la totalidad del dinero. De igual forma, el altruismo está asociado positivamente a la decisión de donar. Por el contrario, la reciprocidad negativa está relacionada con una menor probabilidad de donar. Parece que las personas propensas a reaccionar negativamente cuando son tratados mal tienden a comportarse de manera poco generosa, aunque en principio la reciprocidad debiera, por diseño, no tener impacto en este juego. Este parece un resultado digno de ser estudiado en mayor profundidad en el futuro.

Respecto a las variables de capital social, las personas que creen que se hallan en una sociedad justa tienen mayor probabilidad de donar, les es más fácil ser generosos con extraños.

Los determinantes asociados a la cantidad donada una vez que se decide donar pertenecen principalmente a las características sociodemográficas de los individuos: edad (con un efecto cuadrático en forma de U invertida), ser hombre (con relación positiva), ser desempleado o trabajador por cuenta propia (que donan menos cantidad que los trabajadores por cuenta ajena), y tener una alta contribución a los ingresos del hogar (relación negativa).

Aparte de las características sociodemográficas, los individuos con mayor aversión a la desigualdad en su componente de compasión donan más, una vez que deciden hacerlo. Esta es la única variable que afecta a ambas partes del proceso de decisión (donar y cuánto donar) de la misma forma. No en vano, quien toma la decisión en el juego del dictador parte de una situación relativa ventajosa respecto al receptor, por lo que parece razonable que las personas compasivas que sufren por obtener mayores beneficios que otras personas se muestren más generosas en el juego.

En conclusión, si se pretende entender y aprovechar las razones que llevan a las personas a ser generosas, nuestros resultados indican que debe tenerse en cuenta que los mecanismos que les pueden llevar a participar no siempre son los mismos que los que les pueden llevar a aumentar su participación. Estos resultados, por tanto, pueden ser de gran interés para las organizaciones y actividades que se nutren de la participación voluntaria de terceros.

## Referencias bibliográficas


Andreoni, J., & Vesterlund, L. (2001). Which is the fair sex? Gender differences in altruism. *Quartely Journal of Economics, 116*(1), 293-312.

Brañas-Garza, P. (2006). Poverty in Dictator Games: Awakening Solidarity. *Journal of Economic Behavior & Organization, 60*(3), 306-320.

Brañas-Garza, P., Espinosa, M. P., & García-Muñoz, T. (2009). Expectativas sobre comportamiento egoísta. *Cuadernos Económicos de ICE* (77), 33-43.

Brañas-Garza, P., Cobo-Reyes, R., Espinosa, M. P., Jiménez, N., Kovarik, J., & Ponti, G. (2010). Altruism and Social Integration. *Games and Economic Behavior, 69*(2), 249-257.

Brañas-Garza, P., Cobo-Reyes, R., Espinosa, M. P., Jiménez, N., Kovarik, J., & Ponti, G. (2012). Prosocial norms and degree heterogeneity in social networks. *Physica A: Statistical Mechanics and its Applications, 391*(3), 849-853.

Brañas-Garza, P., Capraro, V., & Rascon-Ramirez, E. (2018). Gender differences in altruism on Mechanical Turk: Expectations and actual behaviour. *Economics Letters, 170*, 19-23.

Brañas-Garza, P., Jorrat, D., & Espín, A. M. (2020). *Envy, guilt and giving* (Mimeo).

Burnham, T. C. (2003). Engineering altruism: a theoretical and experimental investigation of anonymity and gift giving. *Journal of Economic Behavior & Organization, 50*(1), 133-144.

Camerer, C. (2003). *Behavioral Game Theory: Experiments on Strategic Interaction*. Princenton University Press.

Cragg, J. (1971). Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica, 39*(5), 829-844.

Croson, R., & Gneezy, U. (2009). Gender Differences in Preferences. *Journal of Economic Literature, 47*(2), 1-27.

Charness, G., & Gneezy, U. (2001). What's in a name? Anonymity and social distance in dictator and ultimatum games. *Journal of Economist Behavior & Organization, 68*(1), 29-35.

Eckel, C., & Grossman, P. (1998). Are women less selfish than men? Evidence from dictator experiments. *The Economic Journal, 108*(448), 726-736.

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics 14*(4), 583-610.

Erkal, N., Gangadharan, L., & Nikiforakis, N. (2011). Relative earnings and giving in a real-effort experiment. *American Economic Review, 101*, 3330-3348.

Espín, A.M., Moreno-Herrero, D., Sánchez-Campillo, J., & Martín, J.A.R. (2018). Do envy and compassion pave the way to unhappiness? Social preferences and life satisfaction in a Spanish city. *Journal of Happiness Studies, 19*(2), 443-469.

Exadaktylos, F., Espín, A.M., & Brañas-Garza, P. (2013). Experimental subjects are not different. *Scientific Reports, 3*, 1213.

Fehr, E., & Fischbacher, U. (2002). Why social preferences matter: the impact of non-selfish motives on competition, cooperation and incentives. *The Economic Journal, 112*(478), C1-C33.

Fehr, E., & Schmidt, K.M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics, 114*(3), 817-868.

Frohlich, N., Oppenheimer, J., & Moore, J.B. (2001). Some doubts about measuring self-interest using dictator games: the cost of anonymity. *Journal of Economic Behavior & Organization, 46*(3), 271-290.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cárdenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer D., & Ziker, J. (2006). Costly Punishment across Human Societies. *Science, 312*(5781),1767-1770.

Henrich, J., Ensimger J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cárdenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2010). Markets, Religion, Community Size, and the Evolution of Fairness and Punishment. *Science, 327*(5972), 1480-1484.

Hoffman, E., McCabe, K., & Smith, V. (1996). Social distance and other-regarding behavior in dictator games. *American Economic Review, 86*(3), 653-660.

Moffat, P.G. (2016). *Experimetrics. Econometrics for experimental economics*. Palgrave.

Perugini, M., Galluci, M., Presaghi, F., & Ercolani, A.P. (2003). The personal norm of reciprocity. *European Journal of Personality, 17*(4), 251-283.

CUADERNOS ECONÓMICOS ICE INFORMACIÓN COMERCIAL ESPAÑOLA

# Reshuffling Students to Fight Discrimination and Foster Pro-Social Behavior: A Field Experiment*

# *Cómo luchar contra la discriminación y fomentar el comportamiento pro-social: una intervención en la composición de clases en la escuela*

Marina Bosque Mercader
CEMFI

Andrés Gago
Universidad Torcuato Di Tella

## *Abstract*

*There is substantial evidence that group identity shapes discrimination and pro-social behavior. Subjects tend to be more pro-social towards in-group members and discriminate out-group members. In this paper, we run a field experiment in a school in Barcelona to study whether belonging to the same class induces in-group favoritism. Moreover, we explore whether reshuffling students between classes is a good strategy to expand pro-social behavior within the school and fight discrimination. According to our results, participants allocate more resources, are more generous, retaliate less and give higher rewards to students who have always belonged to their same class relative to those who have never been part of it. Moreover, after varying the composition of classes, we find that new students in class start being considered group members and treated as so, while those that are not any longer in the same class do not lose their in-group status.*

***Keywords**: group identity, dictator game, reciprocity, school, discrimination.*
***JEL Classification**: I28, I29, D63, C90, C72.*

## *Resumen*

*La identidad de grupo es un factor importante que incide en los comportamientos discriminatorios así como en las actitudes pro-sociales. Las personas tienden a comportarse de manera más pro-social con otros integrantes de su grupo y discriminan a aquellos que no pertenecen a su comunidad. En este artículo, a través de un experimento de campo en un colegio en Barcelona, estudiamos si pertenecer a la misma clase genera este tipo de identidad de grupo. Además, exploramos si variar la composición de las clases es una estrategia efectiva*

* The authors are particularly grateful to Caterina Calsamiglia, Guillermo Caruana, Yan Chen, Sherry Xin Li, Laia Bosque and Manuel Arellano for their helpful comments and suggestions. We thank Laia Pemán, Laura Ruiz, Rosa Falgà, Oscar Alba and the rest of teachers and students from Aula Escola Europea for their collaboration to make this project possible. The authors acknowledge research funding support from Centro de Estudios Monetarios y Financieros (CEMFI). This project is based on CEMFI's Master Thesis *Group Identity inside the School: Evidence from a field experiment in Barcelona* (Marina Bosque Mercader, 2019).


DOI: https://doi.org/10.32796/cice.2020.99.7011
Recibido: Febrero 2020 · Aceptado: Abril 2020

*para diseminar el comportamiento pro-social y luchar contra la discriminación en la escuela. De acuerdo con nuestros resultados, los estudiantes discriminan positivamente a aquellos con los que han compartido clase en detrimento de aquellos que nunca han pertenecido al grupo, son más generosos con los primeros, menos vengativos, y premian más sus actitudes pro-sociales. De manera interesante, tras variar la composición de las clases, observamos que mientras los nuevos compañeros empiezan a ser tratados como miembros del grupo, aquellos que pertenecieron a la misma clase y ahora pertenecen a otra no pierden su estatus de miembros de la comunidad.*

***Palabras clave**: identidad de grupo, juego del dictador, reciprocidad, discriminación en la escuela.*

***Clasificación JEL**: I28, I29, D63, C90, C72.*

## 1. Introduction

Would you prefer to share a beer with a supporter of your football team or with a supporter of the rival football team? Why would an economist trust more the evidence coming from another economist than a study conducted by a psychologist or a sociologist? Have you ever said: "My country is the best in the world!"? These are some examples of situations in which group identity shapes the way in which we think and act.

Social or group identity refers to a person's sense of belonging to a particular group. The feeling of attachment makes individuals follow the behavioral pattern of the group (Shih et al., 1999; Benjamin et al., 2010), prefer the aggregate benefit of the group (Halevy et al., 2008; Zdaniuk & Levine, 2001; Van Vugt & Hart, 2004) and exhibit more cooperation with in-group rather than out-group members (Goette et al., 2006). Group identity is a notion that is familiar to all of us and it is a central factor to understand social interactions in psychology, biology, sociology, anthropology, political science, and economics. Group membership leads to out-group discrimination when adults distribute resources among third parties (Chen & Li, 2009; Bernhard et al., 2006; Bernhard et al., 2006) and shapes their pro-social behavior (Chen & Li, 2009). In this paper, we aim to contribute to this literature in two ways. First, we want to establish whether belonging to the same class in school generates group identification. To test this, we study whether children discriminate students who have never belonged to their class in favor of their classmates, and whether they are more pro-social when they interact with the later. Second, we want to explore whether reshuffling students between classes is an effective way to foster pro-social behavior in school and fight discrimination.

In order to shed light on these conjectures, we conduct an experiment among students of fourth and fifth of primary in a private school in Barcelona. Based on the experimental design developed by Chen and Li (2009) and redefined by Movahedi (2018), participants play three games that measure out-group discrimination and pro-social attitudes: an other-other allocation game, a dictator game, and a response game. We compare interactions with classmates and with students that have never belonged to the same class to study the influence of group identity in the preferences of participants. Moreover, we take advantage of a school policy which randomly

reshuffles students into three newborn classes to understand if students consider new and former classmates as part of their group and behave accordingly.

In a population of adults, Chen and Li (2009) find that group membership shapes social preferences and determines how participants allocate resources. When we compare interactions between students who have never been in the same class with students that have always been classmates, results are aligned with those in Chen and Li (2009). We observe that when students have to allocate money between an in-group and an out-group member, they discriminate negatively towards the out-groups. Moreover, we observe that when students have to allocate money between themselves and another student, they are more generous with the in-groups. Finally, we observe that when students have the possibility to reward or punish others for their actions paying a cost, they reward more and punish less when they interact with other in-group students. Interestingly, we observe that after reshuffling students, participants treat new students in class as in-groups, while former classmates do not lose their in-group status. This suggests that reshuffling students between classes could be a good policy to expand pro-social behavior and fight discrimination in school.[1]

From an academic perspective, these results contribute to the literature studying the implications of social identity. Social identity theory was first developed by the psychologists Tajfel and Turner (1979) in order to understand the reasons behind inter-group conflicts. According to their theory, group identity is based on three mental processes in which individuals divide the world between people considered *us* (in-group) and *them* (out-group): categorization, identification, and comparison. First, categorization is the process of assigning people into categories according to particular characteristics (gender, ethnicity, religion, politics, etc.). Second, identification is the process of associating ourselves with certain groups. Particularly, in-group members are defined as people who are associated with us, while out-group members are people who are not considered part of our group. Finally, comparison is the process of comparing the in-group members with another group, creating a positive bias toward the group one belongs to. The findings of Turner and Tajfel (1986) conclude that even trivial inter-group categorization leads to in-group favoritism and discrimination against the out-group. Since then, further evidence have been gathered both in social psychology (see for example Mullen et al., 1992; Brewer, 1999) and, more recently, in economics (e.g. Croson et al., 2008; Chen & Li, 2009; Hargreaves Heap & Zizzo, 2009; Sutter, 2009) pointing in the same direction.[2]

---

[1] Even if participants are always more generous, reward more often and are more lenient with students that at some point have been classmates, in some situations there are differences depending on whether they have always been in the same class or not. We will comment on these differences later. Regarding the allocation task, the only group that suffers negative discrimination is the one including students that have never been classmates, while the other groups are treated equally.

[2] It is worth mentioning that even if most of the literature has found in-group favoritism, some papers have found that in-group defectors are judged more negatively both among adults (e.g. Shinada et al., 2004; Marques et al., 1988) and young children (Schmidt et al., 2012; Abrams et al., 2014). This is known as the *black sheep effect* and it has been argued to help group maintenance. We find no evidence of such an effect, provided that in-group students in our study receive softer (not harsher) punishments when they behave selfishly.

Unlike experiments that use trivial cues to group participants, in this paper we do not assign individuals to different groups using the minimal group paradigm (Tajfel, 1970). Instead, we take advantage of the random assignment mechanisms used by the school to establish the initial classes when students join school at the age of three, and the newborn classes in fifth grade. We think that this fact strengthens the external validity of the paper relative to a lab intervention with minimal groups, as it allows us to have exogenous variation in group formation without missing the social ties that typically arise in real groups. Moreover, Hartstone and Augoustinos (1995) claim that conclusions in the minimal group paradigm artificially rely on a dichotomous classification of subjects, and they show that results disappear once more than two groups are created. In our experiment, students can be classified in four groups depending on whether they share class until fourth grade, and whether they do it in fifth grade. This allows us to show that in real group environments individuals discriminate also in settings in which individuals are classified in several groups.

A complete discussion of the differences between real groups and minimal groups can be found in Goette et al. (2012). Interestingly, unlike Goette et al. (2006) and Goette et al. (2012), we do replicate the finding by Chen and Li (2009) that participants are more lenient towards in-groups. This casts doubt on the belief that once we move from minimal groups to real groups, leniency towards in-groups disappears (Goette et al., 2012). Relative to studies that use endogenously formed groups (see for example Bernhard, Fehr & Fischbacher, 2006; Bernhard, Fischbacher & Fehr, 2006) our study allows us to capture the effect of group membership while we compare individuals that are rather similar and abstract from the confounding effect of self-selection into groups.

An important characteristic of our study is that our pool of subjects are children of 9-10 years old. Previous studies in social psychology have found that children as young as 6 years old show similar patterns of out-group discrimination in third party allocation games as those found in the adult population (Jordan et al., 2014). Similarly, Fehr et al. (2008) show that children at the age of 7-8 have already developed other-regarding preferences, i.e. they are willing to sacrifice resources to enhance the welfare of other individuals. Using laboratory and field experiments, Rao (2019) finds that after integrating poor students in elite private schools in India, rich students become more pro-social, generous and egalitarian. He also finds that it becomes less likely that they discriminate against other poor students and more likely that they socialize with them. We contribute to the literature by showing that group membership shapes pro-social preferences in 9 and 10-year-old children in a similar way as it does in adults. More importantly, we show that the pool of children that are considered as in-groups can be easily extended by reshuffling students between classes, as new classmates earn in-group status while former classmates do not lose it.

In the remainder of the paper, Section 2 describes the experimental design in detail, Section 3 presents the results and Section 4 concludes.

## 2. The Experiment

This experiment is conducted with students of fourth and fifth of primary school in *Aula Escola Europea*, a private school in Barcelona. This school divides each cohort into three groups of approximately 34 students. The composition of these groups remains the same since the students start going to the school at the age of three until the students reach the age of ten (which corresponds to fifth grade). At fifth grade, the school implements a policy that randomly reshuffles students into three newborn classes that remain unaltered until secondary school.[3]

The unit of study considered are pairs of students who can be classified in different categories depending on the course. Any pair of students from fourth grade can be identified in one of these two categories: students who have always belonged to the same class (*In-group*) or students who have never belonged to the same class (*Out-group*). Likewise, any pair of students from fifth grade can be classified in one of the following four categories: students who belong to the same class before and after the policy (*In-group*); students who do not belong to the same class before the policy and remain in different classes after the policy (*Out-group*); students who do not belong to the same class before the policy and start being in the same class after the policy (*New-In-group*); or students who used to belong to the same class before the policy and are not anymore in the same class after the policy (*New-Out-group*). The above-mentioned categories are summarized in Figure 1.

Participants play three games that measure both how they allocate resources among third parties and different aspects of pro-social behavior: an other-other allocation game, a dictator game, and a response game.[4] This games are based on Chen and Li (2009) and Movahedi (2018). We compare different pairings to explore how group identity affects participants' decisions.

[3] In particular, the school employs the following criteria: (1) divide the cohort between males and females and put them in alphabetical order, (2) use a random mechanism (that changes from year to year) to allocate students to the new groups and only if it is necessary (3) do adjustments to avoid bad dynamics in the new groups or to balance the groups in terms of ability. According to the school, there is typically no need to do any adjustments.

[4] The games were presented in hypothetical terms. Given that the interest of the project is to find differences in behavior that depend on group status, there is no obvious reason why using an hypothetical approach would change the conclusions qualitatively. Following an educational criteria and only after the experiment was finished, each student was rewarded with an equal amount of sweets. This criteria was jointly agreed with the Ethics Committee of CEMFI and the School Board. When students joined the experiment, they were not aware of any future reward.

**FIGURE 1**
**CATEGORIES OF PAIRS OF STUDENTS**

**Fourth of Primary**

| | |
|---|---|
| **In-group** | Always together |
| **Out-group** | Always separated |

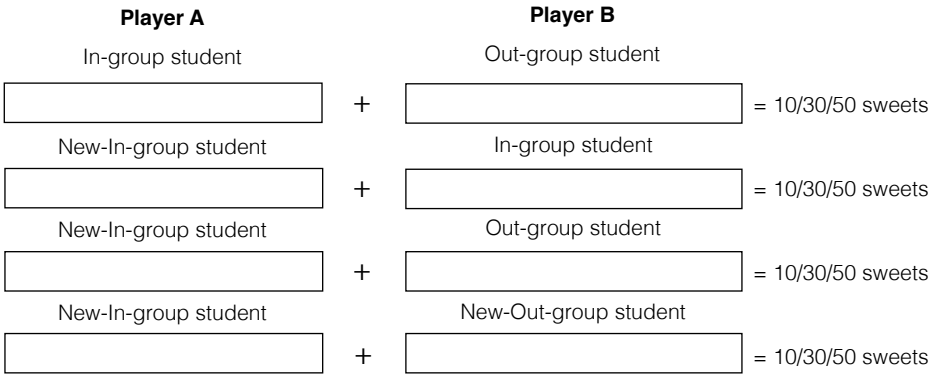
**Fifth of Primary**

| | **Before policy** | **After policy** |
|---|---|---|
| **In-group** | Together | Together |
| **Out-group** | Separated | Separated |
| **New-In-group** | Separated | Together |
| **New-Out-group** | Together | Separated |

### *2.1. First Game: Other-Other Allocation Game*

In this game, participants are asked to allocate a given number of sweets between two other students from their cohort denoted as Player A and Player B. Participants cannot allocate any sweets to themselves. Depending on the relation between Player A, Player B and the allocator, this game considers four different identity combinations (see Figure 2). In order to observe if decisions are sensible to the size of the pie, we also consider three different amounts to divide: 10, 30 and 50 sweets. Provided that in fourth grade there are not *New-In-group* or *New-Out-group* categories, we have 3 observations of the other allocation game for each fourth grade student, whereas we have 12 (4*3) observations for each fifth grade student.

**FIGURE 2**
**OTHER-OTHER ALLOCATION GAME**



### *2.2. Second Game: Dictator Game*

In this game, participants (the dictators) are asked to choose between two options to share a certain amount of sweets with another student of their cohort (the receiver). An illustration of this game is shown in Figure 3. Table 1 summarizes the four versions of this game that consider different allocation choices.

Participants in fourth grade play each version of the game both with an *In-group* and with an *Out-group* student, while participants in fifth grade play with an *In-group*, an *Out-group*, a *New-In-group* and a *New-Out-group* student.

**FIGURE 3**
**DICTATOR GAME**

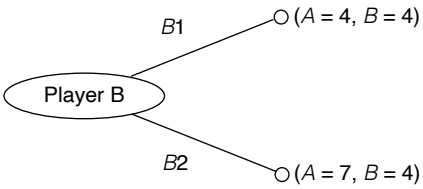



**TABLE 1**
**DICTATOR GAMES**

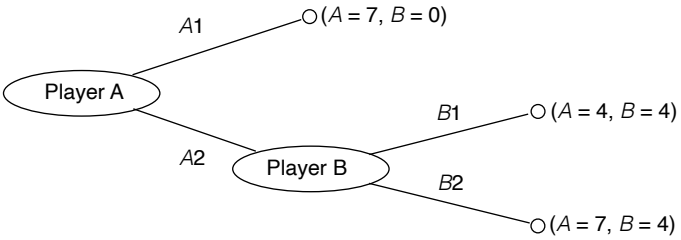
| | **Option B1** | **Option B2** |
|---|---|---|
| DG1 | (4,4) | (7,3) |
| DG2 | (3,6) | (7,5) |
| DG3 | (2,7) | (6,6) |
| DG4 | (0,8) | (4,4) |

**SOURCE:** Based on Chen and Li (2009) and Movahedi (2018).

### *2.3. Third Game: Response Game*

The Response Game is a two-person sequential move game. The first mover, Player A (the proposer), has two options: A1 or A2. If she chooses A1, the game ends with a particular allocation of sweets between herself and Player B. If she chooses A2, Player B (the respondent) decides how to allocate the sweets between options B1 and B2. An illustration of this game can be observed in Figure 4. In this game, participants are asked to play the role of the respondent and consider that the proposer always chooses the option A2. As we will see below, the choice of this option sometimes signals kindness and sometimes unkindness.

**FIGURE 4**
**RESPONSE GAME**




There are nine different versions of this game that consider different allocation choices (see Table 2).[5] We divide the games in three categories: (a), (b) and (c). In categories (a) and (c), when Player A enters the game (i.e., chooses option A2), the payoffs for Player B are lower than if she stays out (i.e., chooses option A1). Conversely, in category (b), when Player A enters the game, payoffs for Player B are larger than if she stays out.

**TABLE 2**
**RESPONSE GAMES**

| | | | | |
|---|---|---|---|---|
| ***(a) Player B punishes Player A at no cost*** | | | | |
| | **A stays out** | **A enters and B chooses:** | | **Associated with:** |
| | | **Option B1** | **Option B2** | |
| RG1 | (5,5) | (4,4) | (7,4) | negative reciprocity |
| RG2 | (4,9) | (2,4) | (4,4) | negative reciprocity |
| ***(b) Player B sacrifices her payoff to reward Player A*** | | | | |
| | **A stays out** | **A enters and B chooses:** | | **Associated with:** |
| | | **Option B1** | **Option B2** | |
| RG3 | (7,0) | (4,4) | (7,3) | positive reciprocity |
| RG4 | (7,1) | (3,6) | (7,5) | positive reciprocity |
| RG5 | (7,2) | (2,7) | (6,6) | positive reciprocity |
| RG6 | (7,0) | (4,4) | (7,3) | positive reciprocity |
| RG7 | (4,0) | (3,4) | (4,3) | positive reciprocity |
| ***(c) Player B sacrifices her payoff to punish Player A*** | | | | |
| | **A stays out** | **A enters and B chooses:** | | **Associated with:** |
| | | **Option B1** | **Option B2** | |
| RG8 | (4,12) | (4,2) | (0,0) | negative reciprocity |
| RG9 | (3,10) | (4,4) | (2,3) | negative reciprocity |

**SOURCE:** Based on Chen and Li (2009) and Movahedi (2018).

[5] In this experiment, following Movahedi (2018), we decided to select those games whose results show the highest difference between in-group and out-group in Chen and Li (2009)'s experiment.

We consider different reactions from Player B to Player A's choice. In the games included in category (a), Player B can decide to punish Player A at no cost. Conversely, in the games included in category (c), Player B needs to incur in a cost if she wants to penalize Player A's behavior. Analogously, in the games included in category (b), Player B needs to sacrifice her own self-interest to reward Player A.

Participants in fourth grade play each version of the game both with an *In-group* and an *Out-group* student, while participants in fifth grade play each game with an *In-group*, an *Out-group*, a *New-In-group* and a *New-Out-group* student.

### *2.4. Experimental Procedure*

We conducted the experiment the 9th and 10th of May 2019 using z-Tree (Fischbacher, 2007). Forty-three students of fourth grade and eighty-eight students of fifth grade participated in the experiment. In order to maximize the number of observations, we used a within-subject design. All decisions taken by participants were anonymous. When participants took decisions, they were informed about the identity of the student they were playing with (name and surname). This conveyed the group categorization in an indirect way.

Before the experiment, all participants took a survey which asked them to name who their five closest friends are, their gender, their age and any extra-curricular activities they are involved in. This information complements the classification of couples of students into one of the four categories and takes into consideration that there might exist a close friendship between two students that do not belong to the same group. This allows us to reduce the variance and obtain more accurate estimators. Instructions were written to make sure that even children of 9-10 years old could understand the implications of the games. This forced us to be explicit, specially, in the reciprocity game. One could argue that this design choice of the instructions could generate demand effects for rewards and punishments. Provided that the focus of our paper is not on the existence of pro-social attitudes among young children (shown in Fehr et al., 2008), but on how pro-social attitudes are affected by group membership, we do not consider this to be a concern. An English version of the instructions and the survey questions can be found in Sections A3 and A4.

## 3. Results

### *3.1. Other-Other Allocation Game*

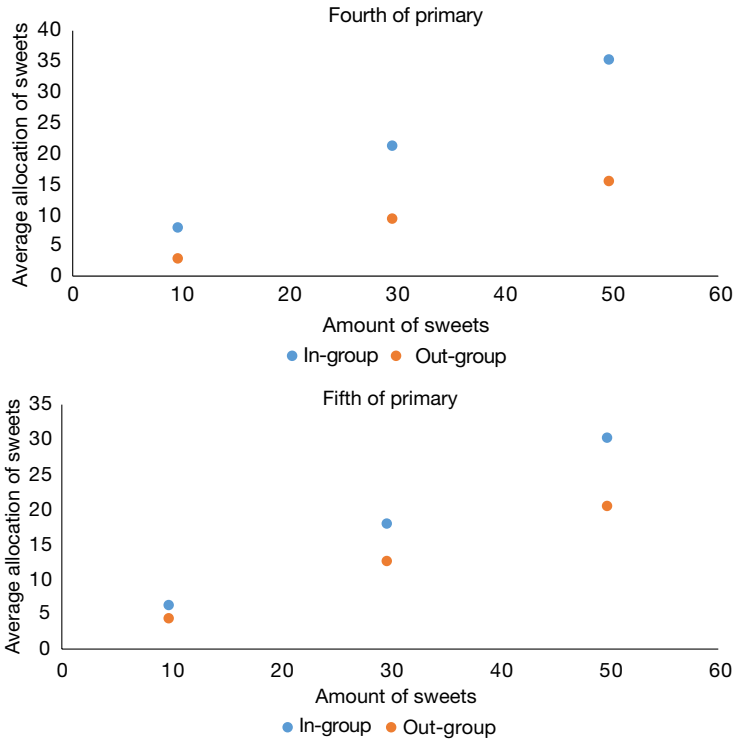
In the other-other allocation game, we want to test two hypotheses. First, we want to understand whether the results in Chen and Li (2009) can be replicated in a school environment. In particular, we test whether being in the same class generates out-group discrimination.[6] Second, we want to evaluate the policy intervention and

[6] Notice that in this experiment In-group favoritism implies Out-group discrimination and viceversa.

understand if we can reduce discrimination by reshuffling students into different classes. Figure 5 plots the average allocation of sweets between *In-group and Out-group students* (y-axis), considering the three different amounts of sweets to divide (x-axis). The top panel shows the results in fourth grade, while the bottom panel shows the results in fifth grade. Results are consistent with class membership generating out-group discrimination. On average, participants allocate significantly more sweets to students who have always belonged to their same class.

Figure 6 evaluates the effect of the policy intervention. These figures compare the way participants allocate sweets between *New-In-group* members and students in the rest of the categories. The top panel presents the average allocation of sweets between a *New-In-group* and an *In-group* student; the panel in the middle presents the average allocation of sweets between a *New-In-group* and an *Out-group* student; and

**FIGURE 5**
**RESULTS OTHER-OTHER ALLOCATION GAME**

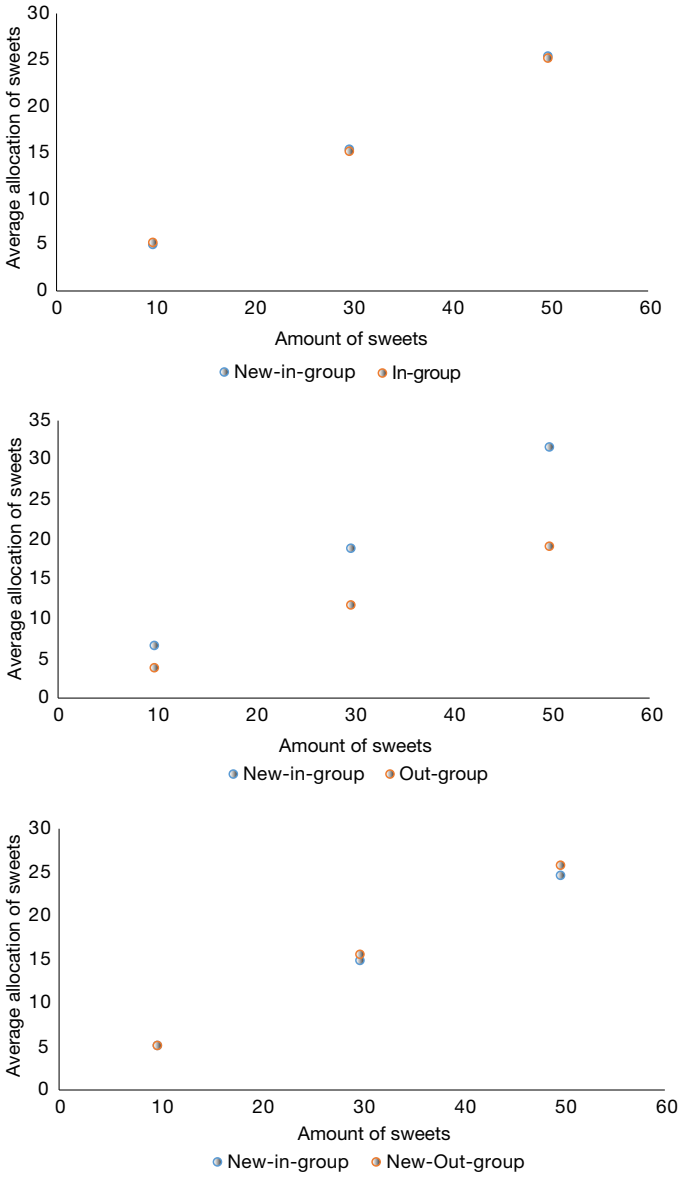



NOTE: The figure plots the average allocation of sweets between students of different categories (y-axis) and the different total amount of sweets to be allocated (x-axis). Blue circles represent the average amount of sweets allocated to In-group students and orange circles represent the average amount of sweets allocated to Out-group students. Symbols labeled as In-group refer to students who have always belonged to the same class as the player; and symbols labeled as Out-group refer to students who have never belonged to the same class as the player.

**SOURCE:** Own elaboration based on experimental data.

**FIGURE 6**

**RESULTS OTHER-OTHER ALLOCATION GAME (FIFTH GRADE)**




NOTE: The figure plots the average allocation of sweets between students of different categories (y-axis) and the different total amount of sweets to be allocated (x-axis). Blue circles represent the average amount of sweets allocated to New-In-group students and orange circles represent the average amount of sweets allocated to In-group, Out-group or New-Out-group students. Symbols labeled as In-group refer to students who have always belonged to the same class as the player; symbols labeled as Out-group refer to students who have never belonged to the same class as the player; symbols labeled as New-In-group refer to students who have never belonged to the same class as the player, but after the policy started being in the same class; and symbols labeled as New-Out-group refer to students who have always belonged to the same class as the player, but after the policy stopped being in the same class.

**SOURCE:** Own elaboration based on experimental data.

the bottom panel presents the average allocation of sweets between a *New-In-group* and a *New-Out-group* student. The results show no difference neither between *New-In-group* and *In-group*, nor between *New-In-group* and *New-Out-group*, while they show a positive discrimination towards *New-In-group* relative to *Out-group* students. Hence, the graphical analysis suggests that one year after reshuffling students, (1) participants consider *New-In-group* students part of their group, while (2) they still treat *New-Out-group* as part of their group even if they are no longer in the same class.

Table 3 presents a formal analysis of the difference in the allocation of sweets between categories. This table displays the average number of sweets allocated to students depending on their category and the category of their pair relative to the allocator. When participants allocate sweets between an *In-group/New-In-group* and an *Out-group* student, the t-tests show they allocate more than half of the pie to the *In-group/New-In-group* student. On the contrary, when they allocate sweets between a *New-In-group* and an *In-group/New-Out-Group*, on average they equally split the number of sweets.

**TABLE 3**

**T-TESTS OTHER-OTHER ALLOCATION GAMES**

| | **Fourth Grade** | | | | **Fifth Grade** | | | |
|---|---|---|---|---|---|---|---|---|
| **Amount of sweets** | **10** | **30** | **50** | **Pooled** | **10** | **30** | **50** | **Pooled** |
| *Categories:* | | | | | | | | |
| In-group vs. Out-group | 7.44*** (0.289) | 20.88*** (0.902) | 34.91*** (1.908) | 0.71*** (0.016) | 6.01*** (0.287) | 17.64*** (0.798) | 29.80*** (1.286) | 0.59*** (0.014) |
| New-In-group vs. In-group | | | | | 4.89 (0.276) | 15.09 (0.836) | 25.13 (1.412) | 0.50 (0.015) |
| New-In-group vs. Out-group | | | | | 6.34*** (0.300) | 18.53*** (0.842) | 31.25*** (1.393) | 0.63*** (0.016) |
| New-In-group vs. New-Out-group | | | | | 5.02 (0.293) | 14.74 (0.841) | 24.42 (1.352) | 0.49 (0.016) |
| Observations | 43 | 43 | 43 | 43 | 88 | 88 | 88 | 88 |

NOTE: The table displays the t-tests from the average amount of sweets which the players decide to give to a student from a specific category (corresponding to the first category mentioned in the column titled Category). In-group refers to games in which the player is interacting with a student who has always belonged to the same class as the player; Out-group refers to games in which the player is interacting with a student who has never belonged to the same class as the player; New-In-group refers to games in which the player is interacting with a student who has never belonged to the same class as the player, but after the policy started being in the same class; and New-Out-group refers to games in which the player is interacting with a student who has always belonged to the same class as the player, but after the policy stopped being in the same class. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**SOURCE:** Own elaboration based on experimental data.

Therefore, after analyzing the results in the other-other allocation game, we can conclude that there is discrimination within school on the grounds of class membership and that reshuffling students between classes is an effective way to fight against it.

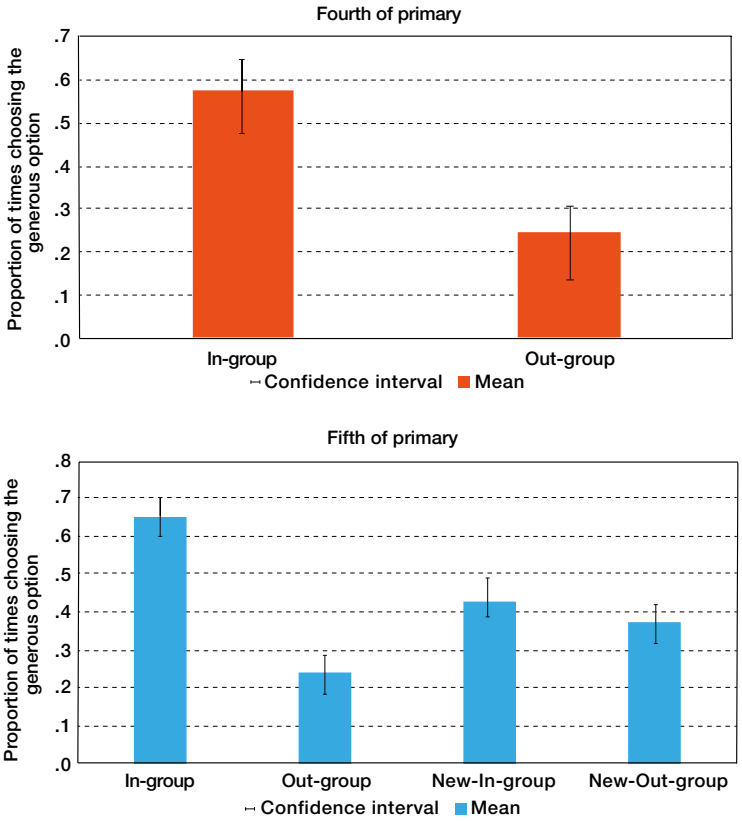
### *3.2. Dictator Game*

The dictator game measures whether subjects are willing to reduce their own wealth to favor others. In this game, we want to test whether students are more

generous when they interact with another student who has always belonged to their same class than when they interact with a student who has never been part of it. Moreover, we want to understand whether reshuffling students into new classes is an effective way to expand the number of children affected by in-group generosity.

The first graph of Figure 7 shows the results of the dictator game from participants in fourth grade and the second graph shows the ones from participants in fifth grade. In particular, the figures present the proportion of times in which the participants choose the generous alternative (option B2 in Table 1) when they are interacting with a student who belongs to a particular category.[7]

**FIGURE 7**
**RESULTS DICTATOR GAME**




NOTE: The figure plots the average amount of times in which players choose option B2 (y-axis) and the category of the student who they are interacting with. The black brackets correspond to the confidence interval of the mean in each category. In-group refers to games in which the player is interacting with a student who has always belonged to the same class as the player; Out-group refers to games in which the player is interacting with a student who has never belonged to the same class as the player; New-In-group refers to games in which the player is interacting with a student who has never belonged to the same class as the player, but after the policy started being in the same class; and New-Out-group refers to games in which the player is interacting with a student who has always belonged to the same class as the player, but after the policy stopped being in the same class.

**SOURCE:** Own elaboration based on experimental data.

We observe that more than 50 % of students choose the generous option when they interact with an *In-Group* student. However, this proportion falls to around 40 % when they interact with a *New-In-group* or a *New-Out-group* student. Finally, when participants interact with an *Out-group* student, this proportion barely exceeds 20 %. This provides evidence that group identity has an effect on generosity. Participants are almost three times more generous towards students that have always belonged to their same class than towards those with whom they have never shared class. Nevertheless, the effects of reshuffling students to foster generosity are somewhat mixed. Changes in class composition increases generosity towards students that used to be in the other class and are now in the same class, but this happens at the expense of students that used to be in the same class and are not anymore. In any case, generosity towards *New-Out-group* students is still sizable and significantly higher than towards *Out-group* students.

The regressions in Table 4 confirm previous evidence. Moreover, we see that for fifth grade, where we have more observations, the probability of choosing the generous option is positively and significantly affected by being a close friend or a central student in the class network.

**TABLE 4**
**REGRESSION DICTATOR GAME**

| | **Probability of choosing the generous action** | |
|---|---|---|
| | **Fourth Grade** | **Fifth Grade** |
| In-group | 0.340***<br>(0.0836) | 0.251***<br>(0.0656) |
| New-In-group | | 0.206***<br>(0.0586) |
| New-Out-group | | 0.141**<br>(0.0553) |
| Degree Centrality | 0.00062<br>(0.0237) | 0.0375***<br>(0.00897) |
| Friend | –0.0227<br>(0.151) | 0.389***<br>(0.0645) |
| Observations | 344 | 1408 |
| $R^2$ | 0.0831 | 0.1282 |

NOTE: The results included in this table are marginal effects evaluated at means from a logit specification. The first column considers results for fourth grade and the second column for fifth grade. Variable Out-group is the reference category and refers to students who have never belonged to the same class as the player; variable In-group refers to students who have always belonged to the same class as the player; variable New-In- group refers to students who have never belonged to the same class as the player, but after the policy started being in the same class; and variable New-Out-group refers to students who have always belonged to the same class as the player, but after the policy stopped being in the same class. Standard errors are in parentheses and clustered at the individual level.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**SOURCE:** Own elaboration based on experimental data.

[7] For more details about the t-tests from the difference in means in each ictator game, look at Table A1 and A2 of Section A1 in the Appendix.

### *3.3. Response Game*

The response game looks at the respondent's reactions to capture negative and positive reciprocity. Again, we first analyze if results in Chen and Li (2009) are also applicable to students. Namely, we test whether participants reward more and punish less when they interact with a student that has always belonged to their class relative to a student that has never been part of it. Afterwards, as in previous sections, we explore if reshuffling students is an effective measure to foster rewards and mitigate punishments.

The following logit regression investigates the effects of group identity on positive reciprocity:

$$Prob\ (B\ rewards\ A) = \frac{e^{(X\alpha)}}{1 + e^{(X\alpha)}} \Lambda[X\alpha]$$

where

$$Fourth\ Grade: X\alpha = \alpha_0 + \alpha_1\ Ingroup + \alpha_3\ BenefitA + \alpha_4\ Bbehind + \\ + \alpha_5\ Centrality + \alpha_6\ Friend$$

$$Fifth\ Grade: X\alpha = \alpha_0 + \alpha_1\ Ingroup + \alpha_2\ NewIngroup + \alpha_3\ NewOutgroup + \\ + \alpha_5\ BenefitA + \alpha_6\ Bbehind + \alpha_7\ Centrality + \alpha_8\ Friend$$

and *Ingroup*, *NewIngroup* and *NewOutgroup* are dummies that indicate the respondent's relation with the proposer; *BenefitA* measures the gain in the proposer's payoff if the respondent chooses to reward; *Bbehind* measures how much the respondent's payoff falls behind the proposer's when the respondent chooses to reward; *Centrality* measures the influence of the proposer in the course network[8]; and *Friend* is a dummy that takes value 1 if the proposer is a close friend.

Analogously, for negative reciprocity:

$$Prob\ (B\ punishes\ A) = \frac{e^{(X\beta)}}{1 + e^{(X\beta)}} \Lambda[X\beta]$$

where

$$Fourth\ Grade: X\beta = \beta_0 + \beta_1 Ingroup + \beta_2 Cost + \beta_3 DamageA + \beta_4 Bahead + \\ + \beta_5 Centrality + \beta_6 Friend$$

$$Fifth\ Grade: X\beta = \beta_0 + \beta_1 Ingroup + \beta_2 NewIngroup + \beta_3 NewOutgroup + \\ + \beta_4 Cost + \beta_5 DamageA + \beta_6 Bahead + \beta_7 Centrality + \beta_8 Friend$$

[8] For more details about the way to construct these variables, look at the Section A2 in the Appendix.

where *DamageA* measures the loss in proposer's payoff if the respondent chooses to punish, *Bahead* measures how much the respondent's payoff gets ahead the proposer's when the respondent chooses to punish, and *Cost* measures the difference in payoffs between the punishing action and the alternative.

Table 5 presents the results from the regressions for fourth and fifth grade. Panel A shows the results of the logit specifications for the determinants of positive reciprocity, whereas Panel B presents the results for the determinants of negative reciprocity. The coefficients are marginal effects evaluated at means.[9]

First, let us consider the determinants of positive reciprocity (see Panel A of Table 5). The analysis focuses on the main specification: Logit with centrality and friends. On average, an *In-group* match increases participants' likelihood to reward by 26.5 percentage points (fourth grade) and 24.3 percentage points (fifth grade) with respect to an *Out-group* member and this effect is significant at 1 percent level. The effect is similar in size and significance when the match is a *New-In-group* (23.3 percentage points) and positive, but slightly smaller, when it is a *New-Out-group* (11.3 percentage points). Moreover, rewards are more likely the more they benefit the proposer, if participants are interacting with a friend, or if they are interacting with someone who is very central in the network. Conversely, rewards are less likely the further behind the respondent lies with respect to the proposer after she rewards. Due to limitations in sample size, some of these effects are only significant for students in fifth grade.

Regarding negative reciprocity, we find the following (see Panel B of Table 5). An *In-group* match decreases participants' likelihood to punish by 17.1 (fourth grade) and 13.8 (fifth grade) percentage points with respect to an *Out-group* member. This effect is even more pronounced for a *New-In-group* match (22.3 percentage points) or a *New-Out-group* match (18 percentage points). All these coefficients are significant at 1 percent level. Additionally, the higher the cost of punishments or the further ahead the respondent ends up after punishing, the lower the probability to punish. Similarly, the probability to punish also decreases when students are interacting with a friend or with someone that is very central in the network. Like in the case of positive reciprocity, due to limitations in sample size, some of these effects are only significant in fifth grade. Furthermore, we find no significant effect of the size of the damage caused to the proposer on the probability to punish.

All these evidence suggests that when participants interact with someone they consider from their own group, the probability that they show positive reciprocity increases, whereas the probability they show negative reciprocity decreases. This phenomenon is reinforced when students interact with a close friend or with a popular student. After evaluating the policy intervention, we conclude that reshuffling students is an effective intervention to increase the probability to reward and mitigate the probability to punish. While participants start treating *New-In-group* as group members, *New-Out-group* students do not lose this status.

[9] For more details about the results in the response game, look at Table A3 and A4 of Section A1 in the Appendix.

## TABLE 5
## DETERMINANTS OF RECIPROCITY

| *Panel A* | **Probability (B rewards A)** | | | |
|---|---|---|---|---|
| | **Fourth Grade** | | **Fifth Grade** | |
| | **Logit** | **Logit with centrality and friends** | **Logit** | **Logit with centrality and friends** |
| In-group | 0.282*** (0.0645) | 0.265* (0.0694) | 0.411* (0.0517) | 0.243* (0.0508) |
| New-In-group | | | 0.261*** (0.0498) | 0.233* (0.0496) |
| New-Out-group | | | 0.148*** (0.0542) | 0.113** (0.0526) |
| Benefit to A if B rewards | 0.0252 (0.0183) | 0.0254 (0.0184) | 0.0393*** (0.00866) | 0.0417*** (0.00919) |
| B's payoff behind of A if B rewards | –0.0736*** (0.0163) | –0.0743*** (0.0164) | –0.0660*** (0.00833) | –0.0711*** (0.00899) |
| Degree centrality | | –0.00868 (0.0308) | | 0.0188** (0.00893) |
| Friend | | 0.306* (0.173) | | 0.519*** (0.0862) |
| Observations | 430 | 430 | 1760 | 1760 |
| $R^2$ | 0.089 | 0.096 | 0.089 | 0.146 |
| *Panel B* | **Probability (B punishes A)** | | | |
| | **Fourth Grade** | | **Fifth Grade** | |
| | **Logit** | **Logit with centrality and friends** | **Logit** | **Logit with centrality and friends** |
| In-group | –0.170*** (0.0581) | –0.171*** –0.0608 | –0.338*** (0.0542) | –0.138*** (0.0534) |
| New-In-group | | | –0.257*** (0.0517) | –0.223*** (0.0492) |
| New-Out-group | | | –0.218*** (0.0537) | –0.180*** (0.0513) |
| Cost to punish A | –0.0482* (0.0284) | –0.0482* (0.0283) | –0.0667*** (0.0206) | –0.0770*** (0.0235) |
| Damage to A if B punishes | 0.0559 (0.0507) | 0.0559 (0.0507) | 0.0104 (0.0293) | 0.0118 (0.0328) |
| B's payoff ahead of A if B punishes | –0.0272 (0.0408) | –0.0272 (0.0408) | –0.0889*** (0.0238) | –0.102*** (0.0269) |
| Degree Centrality | | 0.0157 (0.0217) | | –0.0144** (0.00724) |
| Friend | | –0.0144 (0.134) | | –0.486*** (0.0656) |
| Observations | 344 | 344 | 1408 | 1408 |
| $R^2$ | 0.047 | 0.049 | 0.063 | 0.146 |

NOTE: The results included in this table are marginal effects evaluated at means from different logit specifications. Panel A considers RG3, RG4, RG5, RG6, and RG7 response games and Panel B considers RG1, RG2, RG8, and RG9 response games. Variable Out-group is the reference category and refers to students who have never belonged to the same class as the player; variable In-group refers to students who have always belonged to the same class as the player; variable New-In-group refers to students who have never belonged to the same class as the player, but after the policy started being in the same class; and variable New-Out-group refers to students who have always belonged to the same class as the player, but after the policy stopped being in the same class. Standard errors are in parentheses and clustered at the individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**SOURCE:** Own elaboration based on experimental data.

Interestingly, we find that interacting with an *In-group* relative to interacting with a *New-Out-group* or a *New-In-group* seems to be more effective to foster rewards, but less effective to deter punishments. Further research is necessary to understand this fact. Finally, we also find that the cost of punishing, how much the proposer benefits from rewards and how unequal the payoffs are in case of reward/punishment also have a significant impact on the probability to reciprocate.

## 4. Discussion

This paper studies whether sharing class in school generates out-group discrimination and modulates pro-social behavior. Consistent with findings in Chen and Li (2009) for adults, we find that when students interact with other students that have always been part of their same class, they are willing to favor them in detriment of those that have never been classmates. Moreover, when we measure their pro-social attitudes, we find that they are also more prone to share their own resources with in-groups, more willing to reward their kind actions, and more reluctant to punish their hostile behavior.

In light of these results, we also study whether reshuffling students into different groups is an effective way of fighting discrimination and enlarging the pool of children that benefit from pro-social behavior. We discover that new students in class are no longer discriminated relative to students that have always belonged to the same class. Moreover, we find that the only group that suffers discrimination is the group of students that have never been classmates, while those that were classmates in the past and are not any longer are still treated as part of the group. When we measure pro-sociality, we find that new students in class benefit from being considered part of the group, while there is small or no negative effect on those that do not belong to the group any longer. Hence, taking everything into consideration, our results suggest that schools might prevent internal ghettos and foster pro-social behaviors by modifying the class composition from time to time. To the extent that our results for fourth grade resemble those in Chen and Li (2009) for adults, it seems reasonable to think that these conclusions could be extended to other settings in which homogeneous individuals are divided into groups by superior authorities. We can find examples of such situations in other stages of education and in several work environments.

This paper should be taken as a first approximation to these phenomena and has some limitations that should be addressed in future studies. First and foremost, we cannot form a judgement about efficiency. Even though a higher degree of pro-sociality and lower discrimination might be desirable *per se*, it remains unclear what would be optimal in terms of welfare. If punishments discipline proposers to achieve more efficient outcomes, then lower average punishments towards group members might lead to inefficiencies. Moreover, we find that in some instances having shared class throughout all the courses generates a slightly different pro-social behavior

–sometimes more intense, sometimes less. Further research is necessary to understand these differences.

Before making general policy recommendations, there are also two aspects that should be considered: How reshuffling students between classes affects their academic performance and how it affects their risk habits. There is substantial evidence that peer composition has a significant effect in drinking habits, drug use and criminal behavior, and a modest effect on academic performance (for a survey on this literature, see Sacerdote, 2011). In our study, we could not measure any effects on academic performance, as all classes were reshuffled in fifth grade. Moreover, due to the age of participants, we could not measure either any impact on risk habits. Therefore, complementary research is needed to take an informed policy decision. Finally, we also consider it could be interesting to see how these results extrapolate to other schools, as the experiment is focused on a single intervention in a private school in Barcelona.

## References

Abrams, D., Palmer, S. B., Rutland, A., Cameron, L., & Van de Vyver, J. (2014). Evaluations of and reasoning about normative and deviant ingroup and outgroup members: Development of the black sheep effect. *Developmental Psychology*, *50*(1), 258.

Benjamin, D. J., Choi, J. J., & Strickland, A. J. (2010). Social identity and preferences. *American Economic Review*, *100*(4), 1913-28.

Bernhard, H., Fehr, E., & Fischbacher, U. (2006). Group affiliation and altruistic norm enforcement. *American Economic Review*, *96*(2), 217-221.

Bernhard, H., Fischbacher, U., & Fehr, E. (2006). Parochial altruism in humans. *Nature*, *442*(7105), 912.

Brewer, M. B. (1999). The psychology of prejudice: Ingroup love and outgroup hate? *Journal of social issues*, *55*(3), 429-444.

Chen, Y., & Li, S. X. (2009). Group identity and social preferences. *American Economic Review*, *99*(1), 431-57.

Croson, R., Marks, M., & Snyder, J. (2008). Groups work for women: Gender and group identity in social dilemmas. *Negotiation Journal*, *24*(4), 411-427.

Fehr, E., Bernhard, H., & Rockenbach, B. (2008). Egalitarianism in young children. *Nature*, *454*(7208), 1079.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171-178.

Goette, L., Huffman, D., & Meier, S. (2006). The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review*, *96*(2), 212-216.

Goette, L., Huffman, D., & Meier, S. (2012). The impact of social ties on group interactions: Evidence from minimal groups and randomly assigned real groups. *American Economic Journal: Microeconomics, 4*(1), 101-15.

Grund, T. U. (2015). Social network analysis using stata. In *United Kingdom Stata Users' Group Meetings 2015*, 21. Stata Users Group.

Halevy, N., Bornstein, G., & Sagiv, L. (2008). "In-group love" and "out-group hate" as motives for individual participation in intergroup conflict: A new game paradigm. *Psychological Science*, *19*(4), 405-411.

Hargreaves Heap, S. P., & Zizzo, D. J. (2009). The value of groups. *American Economic Review*, *99*(1), 295-323.

Hartstone, M., & Augoustinos, M. (1995). The minimal group paradigm: Categorization into two versus three groups. *European Journal of Social Psychology*, *25*(2), 179-193.

Jordan, J. J., McAuliffe, K., & Warneken, F. (2014). Development of in-group favoritism in children's third-party punishment of selfishness. *Proceedings of the National Academy of Sciences*, *111*(35), 12710–12715.

Marques, J. M., Yzerbyt, V. Y., & Leyens, J.-P. (1988). The "black sheep effect": Extremity of judgments towards ingroup members as a function of group identification. *European Journal of Social Psychology*, *18*(1), 1-16.

Movahedi, T. (2018). *Essays on group identity and social preferences* (Unpublished doctoral dissertation). Department of Economics, University of Leicester.

Mullen, B., Brown, R., & Smith, C. (1992). Ingroup bias as a function of salience, relevance, and status: An integration. *European Journal of Social Psychology*, *22*(2), 103-122.

Rao, G. (2019). Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools. *American Economic Review*, *109*(3), 774-809.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? In *Handbook of the Economics of Education* (Vol. 3, pp. 249–277). Elsevier.

Schmidt, M. F., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition*, *124*(3), 325-333.

Shih, M., Pittinsky, T. L., & Ambady, N. (1999). Stereotype susceptibility: Identity salience and shifts in quantitative performance. *Psychological Science*, *10*(1), 80-83.

Shinada, M., Yamagishi, T., & Ohmura, Y. (2004). False friends are worse than bitter enemies: "altruistic" punishment of in-group members. *Evolution and Human Behavior*, *25*(6), 379-393.

Sutter, M. (2009). Individual behavior and group membership: Comment. *American Economic Review*, *99*(5), 2247-57.

Tajfel, H. (1970). Experiments in intergroup discrimination. *Scientific American*, *223*(5), 96-103.

Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In W. G Austin & S. Worchel (Eds.), *The social psychology of intergroup relations* (pp. 33-47). Brooks-Cole Publishing.

Turner, J. C., & Tajfel, H. (1986). The social identity theory of intergroup behavior. In W. G Austin & S.Worchel (Eds.), *Psychology of intergroup relations* (pp 7-24). Nelson-Hall Publishers.

Van Vugt, M., & Hart, C. M. (2004). Social identity as social glue: The origins of group loyalty. *Journal of Personality and Social Psychology*, *86*(4), 585.

Zdaniuk, B., & Levine, J. M. (2001). Group loyalty: Impact of members' identification and contributions. *Journal of Experimental Social Psychology*, *37*(6), 502-509.

# APPENDIX

## A1. Tables

**TABLE A1**

**T-TESTS DIFFERENCES IN MEANS DICTATOR GAME (Fourth Grade)**

| | In-group mean | Out-group mean | Difference |
|---|---|---|---|
| Game: | | | |
| DG1 | 0.23 | 0.12 | 0.12<br>(0.082) |
| DG2 | 0.58 | 0.21 | 0.37***<br>(0.099) |
| DG3 | 0.72 | 0.35 | 0.37***<br>(0.101) |
| DG4 | 0.74 | 0.30 | 0.44***<br>(0.098) |
| Pooled | 0.57 | 0.24 | 0.33***<br>(0.071) |
| Observations | 43 | 43 | |

NOTE: The table displays the t-tests from the difference on the average amount of times in which players choose option B2 when they are interacting with students from two different categories. Results are presented for each Dictator Game separated and for all the games pooled. In-group column refers to games in which the player is interacting with a student who has always belonged to the same class as the player; and Out-group column refers to games in which the player is interacting with a student who has never belonged to the same class as the player. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**SOURCE:** Own elaboration based on experimental data.

## TABLE A2
## T-TESTS DIFFERENCES IN MEANS DICTATOR GAME (Fifth Grade)

| | In-group mean | Out-group mean | Diff | New In-group mean | Out-group mean | Diff | New Out-group mean | Out-group mean | Diff |
|---|---|---|---|---|---|---|---|---|---|
| *Game:* | | | | | | | | | |
| DG1 | 0.38 | 0.06 | 0.32*** (0.058) | 0.22 | 0.06 | 0.16*** (0.051) | 0.22 | 0.06 | 0.16*** (0.051) |
| DG2 | 0.63 | 0.18 | 0.44*** (0.066) | 0.39 | 0.18 | 0.20*** (0.067) | 0.31 | 0.18 | 0.13* (0.064) |
| DG3 | 0.81 | 0.35 | 0.45*** (0.066) | 0.53 | 0.35 | 0.18** (0.074) | 0.48 | 0.35 | 0.13* (0.074) |
| DG4 | 0.80 | 0.36 | 0.43*** (0.067) | 0.57 | 0.36 | 0.20*** (0.074) | 0.50 | 0.36 | 0.14* (0.074) |
| Pooled | 0.65 | 0.24 | 0.41*** (0.051) | 0.43 | 0.24 | 0.19*** (0.054) | 0.37 | 0.24 | 0.14** (0.053) |
| Observations | 88 | 88 | | 88 | 88 | | 88 | 88 | |

NOTE: The table displays the t-tests from the difference on the average amount of times in which players choose option B2 when they are interacting with students from two different categories. Results are presented for each Dictator Game separated and for all the games pooled. In-group column refers to games in which the player is interacting with a student who has always belonged to the same class as the player; Out-group column refers to games in which the player is interacting with a student who has never belonged to the same class as the player; New-In-group column refers to games in which the player is interacting with a student who has never belonged to the same class as the player, but after the policy started being in the same class; and New-Out-group column refers to games in which the player is interacting with a student who has always belonged to the same class as the player, but after the policy stopped being in the same class. Standard errors are in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

**SOURCE:** Own elaboration based on experimental data.

**TABLE A3**

**RESULTS RESPONSE GAMES (Fourth Grade)**

| | In-group | | Out-group | |
|---|---|---|---|---|
| | **B1** | **B2** | **B1** | **B2** |
| RG1 | 0.77 | 0.23 | 0.86 | 0.14 |
| RG2 | 0.56 | 0.44 | 0.86 | 0.14 |
| RG3 | 0.53 | 0.47 | 0.77 | 0.23 |
| RG4 | 0.40 | 0.60 | 0.67 | 0.33 |
| RG5 | 0.23 | 0.77 | 0.51 | 0.49 |
| RG6 | 0.63 | 0.37 | 0.74 | 0.26 |
| RG7 | 0.33 | 0.67 | 0.72 | 0.28 |
| RG8 | 0.23 | 0.77 | 0.21 | 0.79 |
| RG9 | 0.44 | 0.56 | 0.19 | 0.81 |
| Observations | 43 | | 43 | |

NOTE: The table displays the proportion of students who choose each of the two options (B1 or B2) in the different response games. In-group columns refer to games in which the player is interacting with a student who has always belonged to the same class as the player; and Out-group columns refer to games in which the player is interacting with a student who has never belonged to the same class as the player.

**SOURCE:** Own elaboration based on experimental data.

**TABLE A4**

**RESULTS RESPONSE GAMES (Fifth Grade)**

| | In-group | | Out-group | | New In-group | | New Out-group | |
|---|---|---|---|---|---|---|---|---|
| | **B1** | **B2** | **B1** | **B2** | **B1** | **B2** | **B1** | **B2** |
| RG1 | 0.69 | 0.31 | 0.91 | 0.09 | 0.72 | 0.28 | 0.76 | 0.24 |
| RG2 | 0.45 | 0.55 | 0.77 | 0.23 | 0.55 | 0.45 | 0.61 | 0.39 |
| RG3 | 0.44 | 0.56 | 0.82 | 0.18 | 0.58 | 0.41 | 0.64 | 0.36 |
| RG4 | 0.25 | 0.75 | 0.65 | 0.35 | 0.39 | 0.61 | 0.55 | 0.45 |
| RG5 | 0.17 | 0.83 | 0.52 | 0.48 | 0.31 | 0.69 | 0.45 | 0.55 |
| RG6 | 0.45 | 0.55 | 0.81 | 0.19 | 0.61 | 0.39 | 0.69 | 0.31 |
| RG7 | 0.34 | 0.66 | 0.73 | 0.27 | 0.45 | 0.55 | 0.55 | 0.45 |
| RG8 | 0.50 | 0.50 | 0.16 | 0.84 | 0.36 | 0.64 | 0.33 | 0.67 |
| RG9 | 0.52 | 0.48 | 0.17 | 0.83 | 0.42 | 0.58 | 0.41 | 0.59 |
| Observations | 88 | | 88 | | 88 | | 88 | |

NOTE: The table displays the proportion of students who choose each of the two options (B1 or B2) in the different response games. In-group columns refer to games in which the player is interacting with a student who has always belonged to the same class as the player; Out-group columns refer to games in which the player is interacting with a student who has never belonged to the same class as the player; New-In-group columns refer to games in which the player is interacting with a student who has never belonged to the same class as the player, but after the policy started being in the same class; and New-Out-group columns refer to games in which the player is interacting with a student who has always belonged to the same class as the player, but after the policy stopped being in the same class.

**SOURCE:** Own elaboration based on experimental data.

## A2. Degree and Eigenvector Centrality

Based on Grund (2015)'s book and using his package of network commands for Stata, we constructed the variable *degree centrality*. It measures the number of times that a student is mentioned as a close friend in the pre-experiment survey. Let $Y$ be an adjacency matrix of a binary network and $Y_{ij}$ the entries in this matrix that indicate if there is a friend relationship from student $i$ to student $j$ ($Y_{ij} = 1$) or not ($Y_{ij} = 0$). Then, the value of degree centrality assigned to student $i$ is defined as:

$$C(i) = \sum_{j=1}^{n} Yji$$

## A3. Pre-Experiment Survey Template

**Identification number**

1. Which is your class group this year: A, I o E?
2. Were you in the school last year?
3. Which was your class group last year: A, B o E?
4. How old are you?

| Male | Female |
|---|---|
| | |

5. Check with a cross your gender:
6. How many siblings do you have?
7. Do your siblings go to the same school as you? How old are them?

| Sibling | Does he/she go to the school? | How old is he/she? |
|---|---|---|
| 1 | | |
| 2 | | |
| 3 | | |
| 4 | | |
| 5 | | |

8. Do you practice any extracurricular activity in the school? Which ones? Circle them.

a) Dance in English
b) Chess
c) Fencing
d) Music
e) Comic
f) Sport School
g) Plastic arts
h) Theatre in English
i) Theatre in French
j) Gymnastics
k) Tennis
l) Choir
m) Orchestra
n) Basket
o) Football
p) Handball
q) Volleyball

9. Could you tell me who are your five best friends from the school in order? You can only mention students from fifth of primary!

1.
2.
3.
4.
5.

## A4. Experiment Template

This is a translation of the experimental instructions into English. The Spanish version can be provided upon request.

### Before the experiment

***Participants are given the instructions of the games***

*Good morning! My name is Marina and I am here today in order to play a game with you. The only rule of the game is that it is an individual game and you cannot talk with anyone until the end of the game. Of course, if you have any doubt, let me know and I will help you!*

*Now, I will propose you different games and I will ask you several questions that you will have to answer in the most honest way possible. There are not correct or incorrect answers, so you will never be wrong with your answer! Moreover, nobody will see your answers!*

*If everything is clear, click on CONTINUE.*

### First Game: Other-Other Allocation Game

***Participants are asked to introduce their identification number.***

*Please, introduce the number that identifies you. Recall to revise you have introduced it well!*

*Once you have introduced it, click on CONTINUE.*

***Participants are asked to divide a particular amount of sweets (10, 30 or 50 sweets) between two other participants who correspond to one of the four different categories (category A, B, C or D).***

*Let us start the games! Imagine that you have a certain amount of sweets and you are playing with TWO other students from the school.*

*I will ask you to divide these sweets between these two students. Be careful, you cannot keep any of the sweets, you have to allocate all of them to the other two students.*

*Now, I will let you know the name of the two students between whom you have to divide the sweets.*

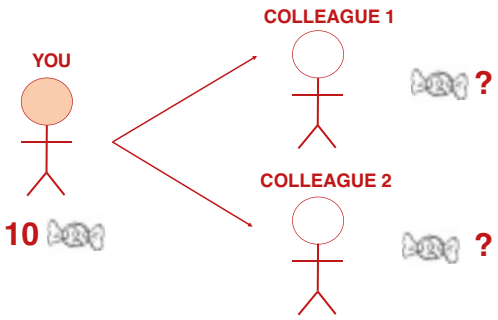
*If everything is clear, click on CONTINUE.*

***In order to illustrate this part of the experiment, let us consider two random names.***

*The two students with whom you will play this round are: Maria (Colleague 1) and Dani (Colleague 2).*

*If you have 10 sweets to divide, how many would you give to Colleague 1?*
*If you have 10 sweets to divide, how many would you give to Colleague 2?*
*Remember that the sum of the two quantities has to be 10!!*
*If you have 30 sweets to divide, how many would you give to Colleague 1?*
*If you have 30 sweets to divide, how many would you give to Colleague 2?*
*Remember that the sum of the two quantities has to be 30!!*
*If you have 50 sweets to divide, how many would you give to Colleague 1?*
*If you have 50 sweets to divide, how many would you give to Colleague 2?*
*Remember that the sum of the two quantities has to be 50!!*
*Once you have taken your decisions, click on CONTINUE.*

***This is the image that the students observe when the game is explained to them.***



***This question will be repeated for the different versions of the game.***
*Game 1 finishes here!*

**Second Game: Dictator Game**

***Participants are asked to introduce their identification number.***

*Please, introduce the number that identifies you. Recall to revise you have introduced it well!*

*Once you have introduced it, click on CONTINUE.*

***Participants are asked to choose between two particular actions that will affect another participant and themselves.***

*Let us start a new game! Imagine you are playing with another student from the school. Thus, consider there are two players in this game, you and another colleague.*

*I will ask you to decide how to share a certain amount of sweets between a colleague and yourself. This time you will not own any amount of sweets, but I will give you two options on how to share the sweets and you will have to choose one of them (OPTION B1 or OPTION B2).*

*Now, I will let you know the name of the student with whom you have to share the sweets and the two options.*

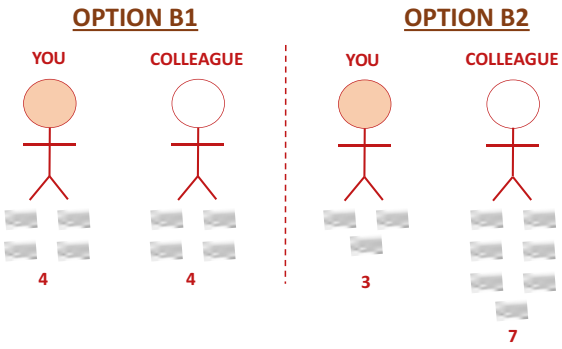
*If everything is clear, click on CONTINUE.*

***In order to illustrate this part of the experiment, let us consider the game DG1 in which the participants, who play the role of Player B, have to choose between option B1 and B2.***

*We are in the DG1 game. The student with whom you will play this round is: Maria.*

*Which of the two option would you choose in order to share the sweets between Maria and you, OPTION B1 or OPTION B2? Once you have taken your decision, click on CONTINUE.*

***This is the image that the students observe when the game is explained to them.***




***This question will be repeated for the different Dictator Games and for the different combinations of student's categories.***

*Game 2 finishes here!*

**Third Game: Response Game**

***Participants are asked to introduce their identification number.***

*Please, introduce the number that identifies you. Recall to revise you have introduced it well!*

*Once you have introduced it, click on CONTINUE.*

***Participants are asked to choose between two particular actions that will affect another participant and themselves. In this case, participants are playing a sequential game with another student.***

*Let us start a new game! Imagine you are playing with another student from the school. Thus, consider there are two players in this game, you and another colleague. This time you will play a sequential game, meaning that your colleague plays first and afterwards it is your turn to play.*

*First, your colleague chooses between two options. If she chooses the first one, the game is over and your colleague decides how to share the sweets. Instead, if she chooses the second option, it will be your turn and you will be responsible to decide how to finish the game.*

*In this game, you will imagine that your colleague always chooses to give you the responsibility to decide how to finish the game. Be careful because being in this situation it is not necessarily good.*

*First, you have to observe the decision of your colleague and understand if that choice was good or bad for you. Then, you can decide how to act in response and choose between two options on how to share the sweets, OPTION B1 or OPTION B2.*

*Now, I will let you know the name of the student with whom you have to share the sweets and the two options. If everything is clear, click on CONTINUE.*

***In order to illustrate this part of the experiment, let us consider the game RG1 in which the participants, who play the role of Player B, have to choose between option B1 and B2, assuming that Player A (their match) has previously decided to continue the game (option A2).***

*We are in the RG1 game. The student with whom you will play this round is: Maria.*

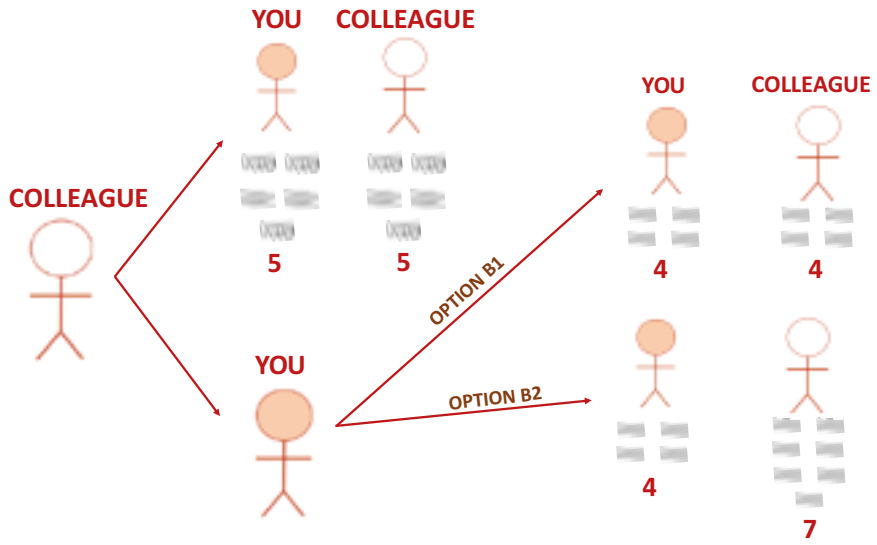
*Imagine that Maria has chosen the second option: you are responsible to decide how to finish the game.*

*Remember, first you have to understand the intentions of your colleague: was she good or bad to you? Notice that the amount of sweets you could have obtained if your colleague had decided to finish the game before is different than the one you can obtain now.*

*Which of the two option would you choose in order to share the sweets between you and your colleague, OPTION B1 or OPTION B2?*

*Once you have taken your decision, click on CONTINUE.*

***This is the image that the students observe when the game is explained to them.***



**This question will be repeated for the different Response Games and for the different combinations of student's categories.**

***Closure***

*This is the end of the games!*
*Thank you very much for your participation!*
*In order to thank you for your help, you will receive 5 sweets!!*

CUADERNOS ECONÓMICOS ICE INFORMACIÓN COMERCIAL ESPAÑOLA

# Country Differences in Ultimatum Wage Bargaining with a Real Task: Evidence from Greece, Spain and the UK*

# *Negociación salarial tipo ultimátum con tarea real: diferencias entre Grecia, España y Reino Unido*

Aurora García-Gallego
Ainhoa Jaramillo-Gutiérrez
Universitat Jaume I

Nikolaos Georgantzís
Universitat Jaume I
Burgundy School of Business

***Abstract***

*From experiments run in Greece, Spain and the UK, this paper studies ultimatum bargaining over the wage paid in order to have a subject perform a given real task. Compared with Greece and Spain, significantly higher wage offers and lower acceptance probabilities are found in the UK. These effects lead to higher wages in the UK. Country differences in both employer and employee behavior have a clear gender component. Specifically, we find that Greek and Spanish females offer lower wages than British females. Contrary to what is obtained for UK subjects, Greek and Spanish females reject more than males.*

***Keywords:*** *ultimatum game, gender differences, experiments.*
***JEL Classification:*** *C91, D03, J16, J3.*

***Resumen***

*Realizamos experimentos en Grecia, España y Reino Unido para estudiar la negociación salarial tipo ultimátum por la realización de una tarea real. Encontramos ofertas salariales significativamente más altas y menor probabilidad de aceptación en Reino Unido que en los otros dos países, lo que conlleva salarios más altos en Reino Unido. Las diferencias en el comportamiento de los sujetos tienen un claro componente de género. Concretamente, encontramos que las mujeres griegas y españolas ofrecen salarios menores que las británicas. Además, en Grecia y España se observa una mayor tasa de rechazo por parte de las mujeres que de los hombres.*

***Palabras clave:*** *juego del ultimátum, diferencias de género, experimentos.*
***Clasificación JEL:*** *C91, D03, J16, J3.*

* The authors acknowledge financial support by the Spanish Ministerio de Ciencia, Innovación y Universidades (grant RTI2018-096927-B-I00), Universitat Jaume I (grant UJI-B2018-76) and ANR (grant ANR-18-CE26-0018, GRICRIS).


DOI: https://doi.org/10.32796/cice.2020.99.7012
Recibido: Noviembre 2019 · Aceptado: Marzo 2020

## 1. Introduction

Behavioral differences across countries are usually attributed to historical, geographical and cultural factors. Furthermore, the overall influence of other social and macroeconomic environmental conditions can also be captured by laboratory data, because norms and beliefs are transferred from the real world to the lab.

Differences in social and cultural environments have been shown to explain behavioral variability in a number of decision making contexts. Among these contexts, the Ultimatum Game (UG) has been often used as a vehicle for assessing behavioral differences across countries. In fact, as Camerer and Fehr (2004) state, the UG is a useful tool for the examination of cultural differences in a negotiation context capturing subjects' monetary expectations and fairness notions presumably dictated at least partly by local social norms. Several contributions in the literature confirm this fact.

In the UG, one proposer offers a share of a fixed amount of money and a responder decides whether to accept or reject the offer. If the offer is accepted, the proposed share is made effective. Otherwise, proposer and responder earn nothing.

Among the earliest, and probably the most well-known study of country differences using the UG, is the one by Roth et al. (1991). The authors compare ultimatum bargaining and multi-person market environments in Israel, Japan, the United States and Yugoslavia. They find differences in bargaining behavior in the four countries, which are tentatively attributed to cultural differences. In Henrich et al. (2001), the authors recruited non-student subjects from 15 small-scale societies in 12 countries and found large group differences in the offers and in the rejection rates. In their follow up studies, Henrich et al. (2006, 2010), the authors confirm that cultural differences explain more variance than individual differences.[1]

Another important contribution is the meta-analysis by Oosterbeek, Sloof and van de Kuilen (2004). They use data obtained from 37 papers which form a total of eight geographical groups. Contrary to Henrich et al. (2001), the authors find that country differences are reflected only on respondents' behavior and not on the behavior of proposers. The authors conclude that these differences cannot be attributed to cultural traits as the ones used by Hofstede (1991) and Inglehart (2000).

In Chuah et al. (2007, 2009), a cross-national UG is used to analyze differences in behavior between Malaysian Chinese and the UK subjects. They find cultural differences in both intra- and inter-national interactions, although they find a small effect of demographic variables on subjects' behavior.

A more general analysis is made by Botelho et al. (2001), whose experimental design controls for demographics. The authors test for cultural differences in Russia

[1] In the three studies by Henrich et al. (2001, 2006, 2010), the responder's willingness to reject is interpreted as a measure of costly punishment. Although we are not against this interpretation, we do not feel our data are in position to disentangle this from a simple non-acceptance to participate in an unfair outcome and we keep this standard notion of rejection throughout the paper.

and the United States. Although there is a country effect, the authors conclude that cultural differences are more complex than the factors captured by either nationality or gender alone.

In this paper, we focus on the role of gender as an explanatory factor of differences across countries in ultimatum wage bargaining behavior. Concerning the role of gender in human decision making, Eckel and Grossman (2008) review the results from two ultimatum experiments, Solnick (2001) and Eckel and Grossman (2001), searching for evidence of systematic differences in the behavior of men and women. Although there are several differences in the design of these two experiments, there are considerable similarities in the results. The differences are mainly two. First, while Solnick (2001) conducts a one-shot UG game in which the responder has to decide about the minimum acceptable offer without being aware of the offer proposed,[2] Eckel and Grossman (2001) run a repeated-play (eight periods) UG using a direct-response method.[3] Second, Solnick (2001) reveals players' gender in the second treatment, while subjects have no information on their partner's identity in Eckel and Grossman (2001). As far as the results are concerned, both studies find small differences in the offers made both by men and women. Moreover, both conclude that offers to women are, on average, lower than offers made to men, independently of the sex of the proposer. The divergence in the results focuses on the behavior of the responder. While the overall rejection rates are similar, Solnick (2001) reports higher rejection rates in offers made by women, while Eckel and Grossman (2001) find higher rejection rates for offers made by men.[4] The difference is conjectured to originate from the fact that risk is absent for the responder in this setting.

There are several studies which confirm the view that women tend to be more risk averse than men (García-Gallego et al., 2012; Filippin & Crosetto, 2016). Consistent with this explanation are the findings of Croson and Buchan (1999) who examine gender differences in the trust game[5] using data from four countries: The United States, China, Japan and Korea. They find that, while no significant effect of gender is found with respect to the proposer, when making the risk-less decision of responder, women return significantly more than their male counterparts. However, in a related paper using the Spanish dataset alone, we show that gender differences in the UG are not due to risk attitude differences (García-Gallego et al., 2012).

[2] This is like an UG played simultaneously. If the proposer's offer equals or exceeds the responder's minimum acceptable offer, the offer is accepted, and the pie divided according the proposer's proposal. Solnick (2001) explicitly names this as the UG played according to the strategy method.

[3] Under the direct-response method, the proposer makes an offer which is presented to the responder, who then decides whether to accept or reject.

[4] In particular, the major difference is found in the rejection rates for offers made by women to women.

[5] In the standard trust game, the proposer and the responder start with the same endowment. The proposer then has the opportunity to give a portion of his endowment to the responder. Whatever amount the proposer decides to give to the responder will be tripled before it is passed on to the responder. The responder then has the option of returning any portion of this tripled amount to the proposer.

Eckel and Grossman (2001) report substantial differences in ultimatum play by African-American subjects. Black subjects are clearly more egalitarian in their proposals. In addition, blacks are more likely to reject an offer of a given size. The authors find a significant difference in the distributions of offers made by black and nonblack subjects.

Saad and Gill (2001) conduct a one-shot UG in which subjects randomly face a subject of the same or contrary gender. Each subject knows the sex of his/her partner. They find that males make more generous offers when pitted against a female. Furthermore, females made equal offers independently of the other's sex.

However, none of these papers addresses the issue of how gender effects depend on the subject pool. In order to answer the question whether gender effects depend on cultural differences, we compare behavior of subjects from three subject pools of university students from Greece, Spain and the UK.

Our analysis is based on a series of experimental sessions which do not *ex ante* control for the composition of pairs in terms of gender. We frame ultimatum bargaining as a situation in which proposers are employers and responders are employees who are offered a salary in order to perform a given task. Among a number of novel design aspects introduced in these experiments, the most prominent feature is that employees accepting a given salary have the obligation to perform a real task.

Our main finding is that gender differences significantly depend on the subject pool, with Greek and Spanish female-responders being more likely to reject a given offer than men-responders, while British females exhibit the contrary tendency.

The paper is organized as follows. In Section 2 we summarize the main design features of the experiment. In Section 3 the main results are discussed. Section 4 concludes. An appendix at the end of the paper includes some extra tables and the instructions given to experimental subjects.

## 2. Experimental design

In the basic UG, two players, proposer and responder, are offered a fixed amount of money which has to be shared between them in the following way: the proposer suggest a split which may then be accepted or rejected by the responder. If accepted, the split takes place as suggested by the proposer, and if not, then both players earn nothing. In the salary negotiation framing used here, employer/employee pairs negotiate over their respective shares from a 10 monetary units (m.u. hereafter) profit earned from a given task performed by the employee. In the experiment, the task is real and a unit of it corresponds to filling 20 numbered envelops with their corresponding numbered single-page letters. The role of this feature on the UG is studied by García-Gallego et al. (2008). As we report there, apart from realism, it makes responders more demanding yielding higher offers and fairer actual splits of the pie.

Three sessions were run, respectively, in three European countries. Specifically, the first session took place in Greece (GR), in the Computer Lab of the Economics Department of the University of Macedonia. The second session took place in Spain (SP), at the Laboratorio de Economía Experimental of the Universitat Jaume I, Castellón. The third session was run in the UK, in the Computing Laboratory of the Economics Department of the University of Aberdeen, Scotland.

Each experimental session was divided into two different sub-sessions: One during which ultimatum salary negotiation took place and another one, in which employees performed their real task obligations. After this, all subjects were paid as explained below. The first sub-session consisted of two parallel procedures. Namely, *i)* randomly formed employee-employer pairs[6] play the salary negotiation UG, which is repeated over a randomly determined number of periods ranging between 30 and 35, and *ii)* subjects respond to a payment-card type of control question designed to elicit subjects' beliefs in form of their valuations (certainty equivalent) of the game. Each subject's role was kept fixed along the whole session.

In five randomly chosen periods of the main experiment, a question is presented to the subjects in an incentive-compatible design to control for variations in a subject's valuation of the game due to learning and/or due to changing from a hypothetical to an incentive compatible design. The hypothetical valuations of the game are denoted as $HGV$[7] and the five incentive compatible ones by $ICGV_1$ to $ICGV_5$. Given the lack of any systematic hypothetical bias (measured as differences between hypothetical and real valuations), in the econometric models, we have used Certainty Equivalent which are constructed as the average of each employer/employees' valuations obtained as responses to the hypothetical and the incentive compatible formats of the control questions.

After the negotiation periods were completed and data were collected for each session, five periods were randomly chosen by the computer to determine each employer-subject's earnings and each employee-subject's salaries and real task obligations. Real rewards were paid to employees at the end of the task-performing sub-session, while employers were rewarded just after the end of the salary negotiation sub-session.

A total of 60 subjects (10 employers and 10 employees per session) participated in this experiment. Subjects were university students studying economics-related topics and they were assigned once to a single session. Each session lasted, on average, one and a half hours. Average per subject earnings were slightly below 25€ in Spain and Greece, and slightly below 25£ in the UK. Specific software was written using z-Tree toolbox (Fischbacher, U., 2007).

---

[6] Subjects in each session were divided into two separate matching groups. Differences across groups within the same session were found not to be statistically significant, and data reported here are the result of pooling across groups.

[7] Each subject responds to a payment-card type of control question designed to elicit subject's valuation (certainty equivalent) of the game. See the Questionnaire subsection in the Appendix C.

Table 1 describes some details of the experimental data obtained: the number of subjects, the males and females numbers and the number of markets in each treatment.

**TABLE 1**

**MAIN CHARACTERISTICS OF THE EXPERIMENT**

| Treatments | N | Male | Female | Markets (Pairs) |
|---|---|---|---|---|
| GR | 20 | 9 | 11 | 10 |
| SP | 20 | 6 | 14 | 10 |
| UK | 20 | 12 | 8 | 10 |
| Total | 60 | 27 | 33 | 30 |

## 3. Results

In this section we present the analysis of our experimental data. Although the number of independent markets is small, the repeated nature of our game adds statistical robustness to the analysis. Specifically, we have taken into consideration the first 33 periods, the maximum number of periods for which we have data for all treatments.

### *3.1. General results and descriptive statistics*

The first step of our analysis is to assess whether behavioral differences exist among subjects from Greece, Spain and the UK.[8] In this subsection we present some descriptive statistics. In Table 2, we provide offer averages by country and by gender. The table also includes hypothetical game valuations (HGV) of subjects as well as incentive compatible game valuations by player type (*ICG*Ver for employers and *ICG*Vee for employees). Finally, this table includes salary averages considering accepted offers as salaries.[9] No statistically significant differences are found in hypothetical game valuations across countries. This means that, before starting the experiment, subjects from the three countries have similar hypothetical valuations of the game. However, incentive compatible game valuations exhibit some differences. Greek and Spanish employee-subjects' valuations in the incentive compatible format are lower than the British employee-subjects' valuations.

[8] For this purpose, we use per subject averages as independent observations.
[9] Zero payoffs due to a rejection are dropped from the sample.

Overall, offers present differences across countries. Mann-Whitney (M-W) tests confirm that offers in the UK are significantly higher (4.97) than in the other two countries (4.09 in Spain and 4.06 in Greece). Moreover, offers in the UK almost coincide with the 50 %-50 % split of the pie.

**TABLE 2**

**DESCRIPTIVE STATISTICS**
**(Mean, Standard Deviation and M-W tests)**

| | **Greece** | | | **Spain** | | | **UK** | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Overall** | **Males** | **Females** | **Overall** | **Males** | **Females** | **Overall** | **Males** | **Females** |
| Offer | 4.06 | 4.11 | 3.98 | 4.09 | 4.53 | 3.80 | 4.97 | 4.91 | 5.06 |
| (St.D.) | (0.22) | (0.20) | (0.26) | (1.18) | (0.34) | (1.47) | (0.40) | (0.40) | (0.44) |
| HGV | 5.13 | 4.89 | 5.32 | 4.65 | 4.75 | 4.61 | 4.45 | 3.92 | 5.25 |
| (St.D.) | (1.55) | (1.24) | (1.81) | (1.36) | (0.76) | (1.57) | (2.44) | (1.68) | (3.91) |
| ICGVer | 5.86 | 5.89 | 5.81 | 5.33 | 5.33 | 5.32 | 5.09 | 5.32 | 4.74 |
| (St.D.) | (1.18) | (1.28) | (1.20) | (0.59) | (0.26) | (0.77) | (0.91) | (0.93) | (0.87) |
| ICGVee | 4.12 | 3.98 | 4.18 | 4.33 | 4.14 | 4.38 | 5.45 | 5.54 | 5.31 |
| (St.D.) | (0.45) | (0.47) | (0.47) | (0.78) | (0.20) | (0.87) | (0.70) | (0.84) | (0.47) |
| Salaries | 4.21 | 4.28 | 4.08 | 4.55 | 4.59 | 4.46 | 5.19 | 5.11 | 5.26 |
| (St.D.) | (0.16) | (0.12) | (0.19) | (0.18) | (0.19) | (0.01) | (0.32) | (0.20) | (0.36) |
| **M-W Tests between countries (*p*-values)** | | | | | | | | | |
| | **Greece-UK** | | | **Greece-Spain** | | | **Spain-UK** | | |
| | **Overall** | **Males** | **Females** | **Overall** | **Males** | **Females** | **Overall** | **Males** | **Females** |
| Offer | 0.0002 | 0.0039 | 0.0209 | 0.1735 | 0.0871 | 0.6698 | 0.0052 | 0.2008 | 0.0105 |
| HGV | 0.2303 | 0.1604 | 0.9338 | 0.1381 | 0.6701 | 0.1341 | 0.9782 | 0.3424 | 0.2843 |
| ICGVer | 0.1978 | 0.6304 | 0.1489 | 0.2256 | 0.3923 | 0.3938 | 0.2501 | 0.1869 | 0.4542 |
| ICGVee | 0.0004 | 0.0196 | 0.0140 | 0.1400 | 0.5637 | 0.2022 | 0.0025 | 0.0442 | 0.0415 |
| Salaries | 0.0002 | 0.0201 | 0.0082 | 0.0007 | 0.0833 | 0.0055 | 0.0002 | 0.0455 | 0.0066 |

Figure 1 shows a positive time trend of offers which is a common feature for the three countries. From early periods, British subjects make offers which, on average, exceed 4.5, while Spanish and Greek subjects start their offers from below 3.5 and raise them over time.

Referred to Table 2, offers made by British females are significantly higher than offers made by females in Spain and Greece. In fact, British females make offers which are, on average, even higher than the 50 % of the pie, while Spanish and Greek females offer their employees a share which lies below 40 %.

As far as males are concerned, in the three countries males' offers are higher than 4 m.u., although the offers by male subjects in the UK are significantly higher than those offered by male subjects in Greece.

**FIGURE 1**
**EVOLUTION OF OFFERS BY COUNTRY**
**(Average per period)**

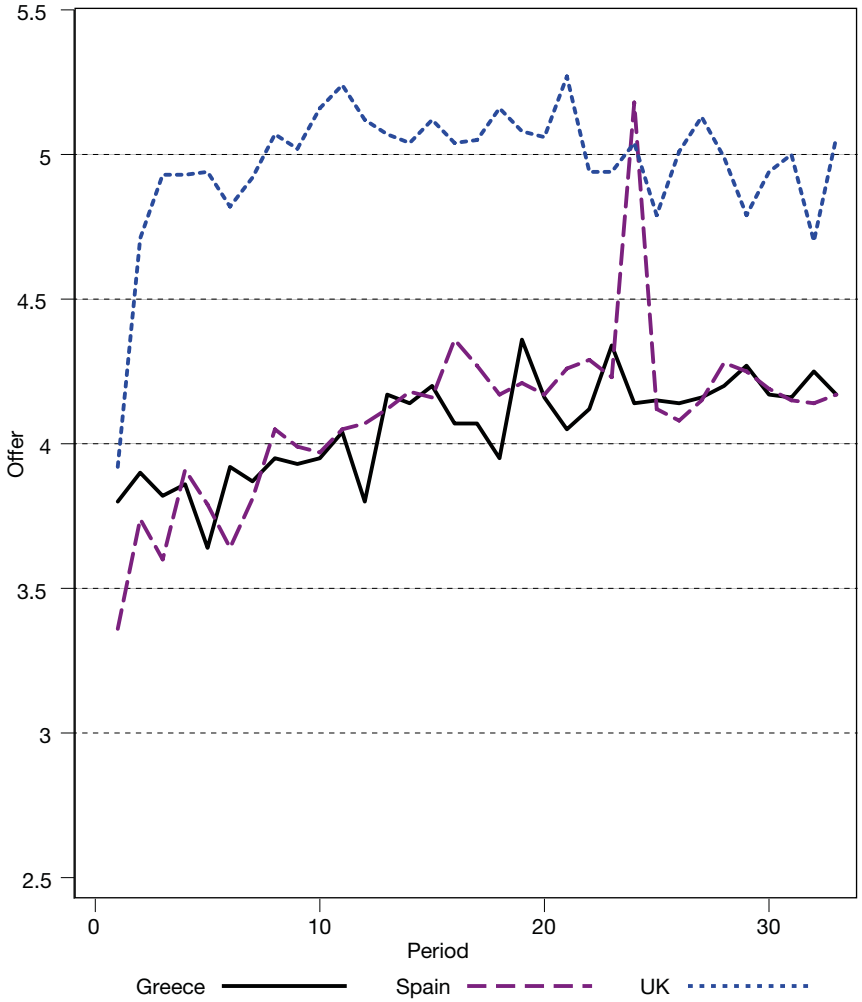
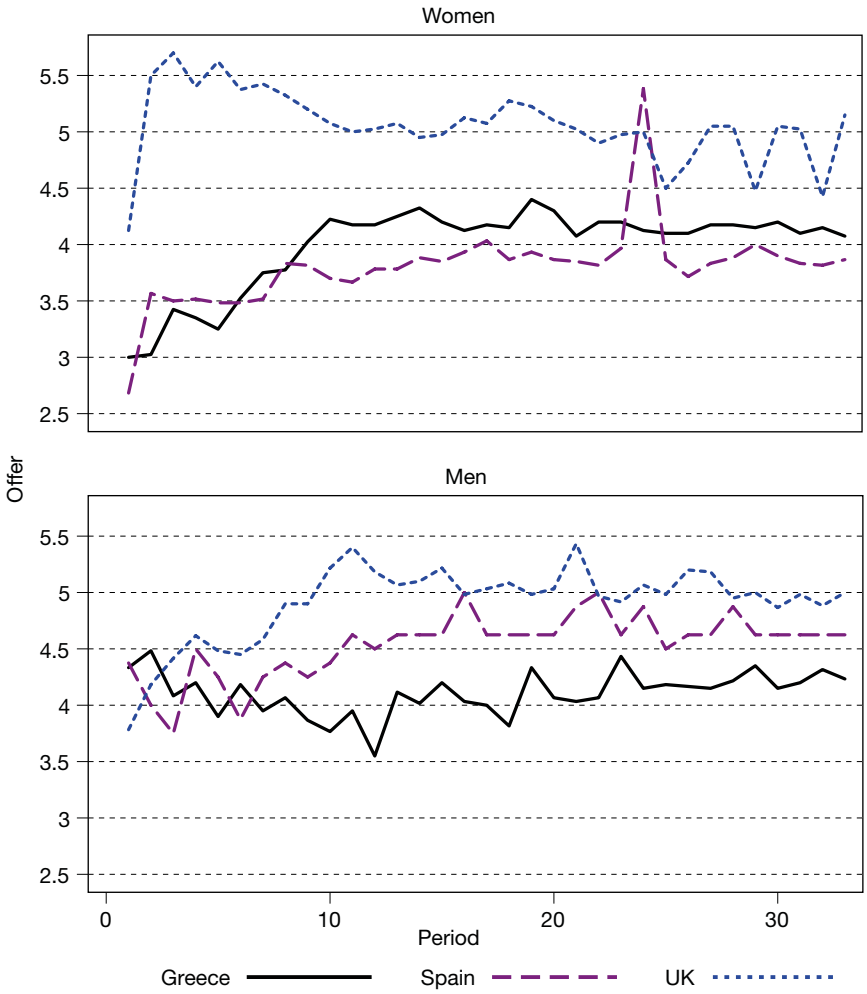



Figure 2 shows the evolution of offers by country and gender. Note that females' offers in the UK are always above the offers made by Greek and Spanish females. The trend is positive for males in the three countries as well as for females in GR and SP. We confirm that Greek and Spanish females offer lower wages than females in the UK. Spanish and Greek females seem to behave in similar ways over time and differently than British females, who systematically make higher offers.

**FIGURE 2**
**EVOLUTION OF OFFERS BY EMPLOYER'S GENDER**
**(Average per period)**




In Table 3 we present information on offers percentage, by country, gender and offers' interval. Approximately 91 % of all offers (301/330) by British subjects are above 4.2 m.u. Offers in Spain and Greece exhibit a distribution with two peaks. Concerning Spain, one peak is observed on the 3.7-4.2 interval, which corresponds to the 30 % (100/330) of all offers. A proportion of 40 % (127/330) of all offers falls within the 4.7-5.2 interval. The remaining 50 offers fall between these two modes. In the case of Greek subjects, the two peaks correspond to offers in the intervals 3.7-4.2 m.u. (37 % of all offers, 122/330) and 4.2-4.7 (43 % of all offers, 143/330). The data by gender follow the same pattern.

**TABLE 3**
**PERCENTAGE OF OFFERS, BY INTERVALS**

| Interval of offers | Greece | | | Spain | | | UK | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Males** | **Females** | **Total** | **Males** | **Females** | **Total** | **Males** | **Females** | **Total** |
| $x < 3.2$ | 4.55 | 2.08 | 3.39 | 0.76 | 14.85 | 13.74 | 1.31 | 1.89 | 1.57 |
| $3.2 \leq x < 3.7$ | 7.07 | 14.39 | 11.29 | 5.30 | 1.01 | 4.35 | 2.53 | 0.76 | 2.23 |
| $3.7 \leq x < 4.2$ | **42.42** | **28.79** | **38.17** | **25.76** | **33.33** | **30.76** | 5.05 | 3.03 | 4.47 |
| $4.2 \leq x < 4.7$ | **42.42** | **44.70** | **43.36** | 18.94 | 12.63 | 15.78 | 17.17 | 12.12 | 15.56 |
| $4.7 \leq x < 5.2$ | 1.52 | 6.06 | 4.82 | **44.70** | **34.34** | **39.15** | **49.49** | **41.67** | **46.68** |
| $5.2 \leq x < 5.7$ | 1.52 | 0 | 1.52 | 1.52 | 0 | 1.52 | 12.12 | 21.97 | 17.51 |
| $5.7 \leq x$ | 0.51 | 0 | 0.51 | 1.52 | 1.01 | 1.26 | 11.11 | 17.42 | 14.34 |

NOTE: The most frequent offers by interval are in bold characters.

In Table 4 we show the rate of rejections per country by salary offers' interval. We include the percentage of rejected offers distinguishing between male and female subjects in each country.

Regarding responders' behavior, we observe that the rejection rate in the three countries is approximately 30 %. Additionally, in all countries the rejection rate decreases as offers increase. As far as gender differences are concerned, note that in Spain and in Greece the percentage of rejected offers is higher for females than for males. On the contrary, males in the UK reject more than females. This result receives a strong support if we compare offers above the 3.7-4.2 m.u. interval.

**TABLE 4**
**PERCENTAGE OF REJECTED OFFERS, BY INTERVALS**

| Interval of offers | Greece | | Spain | | UK | |
|---|---|---|---|---|---|---|
| | **Males** | **Females** | **Males** | **Females** | **Males** | **Females** |
| $x < 3.2$ | 66.67 | 81.82 | 100 | 88.57 | 83.33 | 66.67 |
| $3.2 \leq x < 3.7$ | 27.27 | 81.82 | 0 | 100 | 100 | 66.67 |
| $3.7 \leq x < 4.2$ | 24.39 | 44.44 | 0 | 55.26 | 100 | 57.14 |
| $4.2 \leq x < 4.7$ | 2.56 | 16.35 | 0 | 32.43 | 75.68 | 38.46 |
| $4.7 \leq x < 5.3$ | 0 | 11.11 | 0 | 10.68 | 34.12 | 14.71 |
| $5.2 \leq x < 5.7$ | 0 | 0 | 0 | 50 | 19.35 | 0 |
| $5.7 \leq x$ | 0 | 0 | 0 | 0 | 13.79 | 0 |

Salaries (accepted offers) present the same pattern as offers (see Table 2). We find significant differences across countries. Overall, salaries are significantly higher in the UK (5.19, on average) than in any of the other two countries (averages of 4.55 and 4.21 for Spain and Greece, respectively). Therefore, in terms of salaries, we observe a ranking with UK in the first, Spain in the second and Greece in the third position. Interestingly, observed salaries reproduce the ranking of the three countries in terms of the macroeconomic indicators like GDP and industrial and services earnings, presented in Table A (see Appendix A). This implies that the observed differences may depend on country differences on the levels of income and/or other macroeconomic determinants of subjects' opportunity costs and target earnings.

Figure 3 presents the evolution of salaries over time, by country. We observe similar patterns as with respect to posted offers. Figure 4 presents salaries by country and by gender. Likewise the case of offers, salaries in the UK are always above the salaries in Greece and Spain.

**FIGURE 3**
**EVOLUTION OF SALARIES BY COUNTRY**
**(Average per period)**

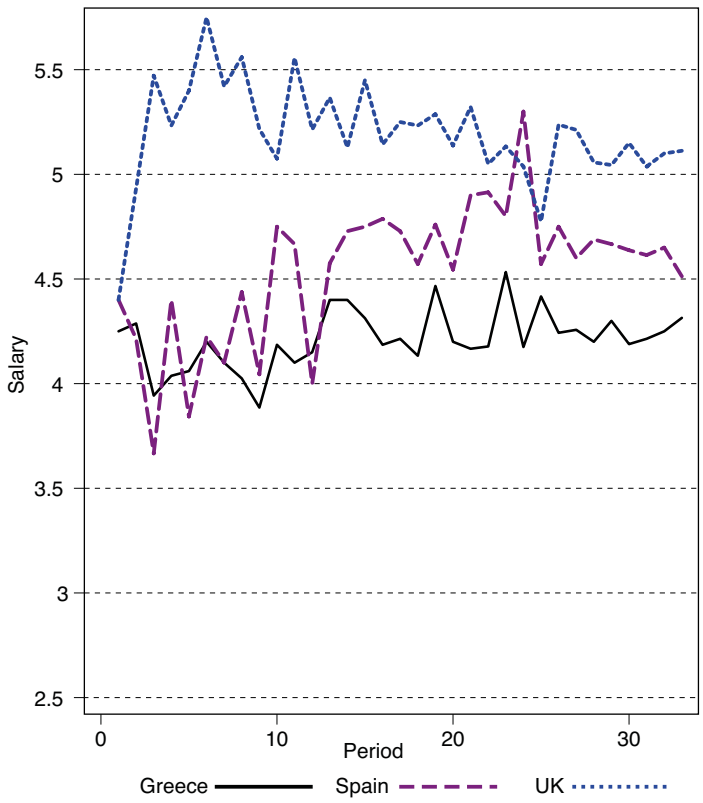
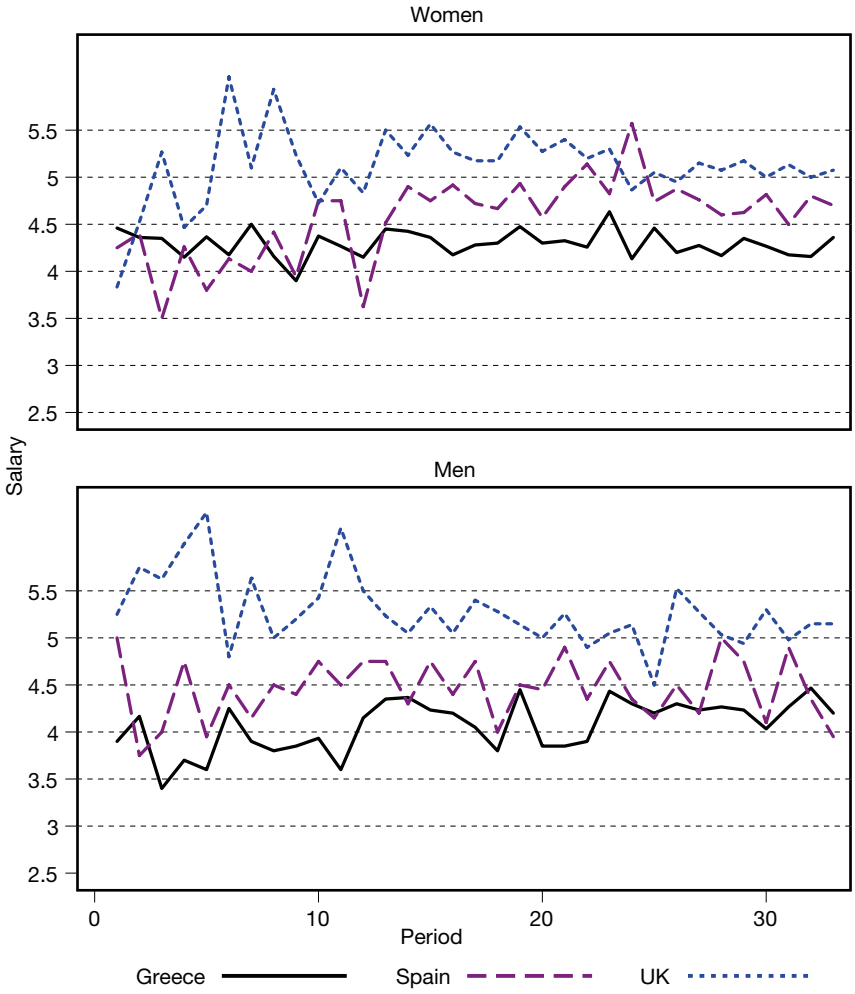


**FIGURE 4**
**EVOLUTION OF SALARIES BY EMPLOYEE'S GENDER**
**(Average per period)**




In the next section we discuss the results in more detail using regression analysis.

### *3.2. Econometric analysis*

Table 5 reports results from the estimation of the models of offers, capturing the main features of employer behavior. We estimate a random effects GLS models to explain the effects of our explanatory variables affects the offers made by subjects. In the first model estimate overall pool and, in the others, we separate our data by

gender. The explanatory variables in all models are: period (T), certainty equivalent and country dummies (Spain and UK). The overall estimation indicates that Greek and Spanish employers have a similar behavior, however British employers' behavior differ from Greeks. In fact, British make higher offers than the others. As far as gender differences are concerned, offers made by Spanish females are not significantly different from the ones made by Greek females. However, offers made by British females are significantly higher than offers made by females in Spain and Greece.

In males' estimation, we observe that the highest wage offers are made by British employers in the first, Spanish in the second and Greek in the third place.

**TABLE 5**

**MODELS OF OFFERS**

| | GLS(Overall) | | GLS(Male) | | GLS(Female) | |
|---|---|---|---|---|---|---|
| Constant | 4.197 | (0.245)*** | 4.448 | (0.348)*** | 3.938 | (0.286)*** |
| T | 0.013 | (0.005)*** | 0.015 | (0.005)*** | 0.012 | (0.008) |
| Certainty equivalent | –0.066 | (0.038)* | –0.104 | (0.057)* | –0.029 | (0.039) |
| Spain | 0.010 | (0.357) | 0.385 | (0.175)** | –0.193 | (0.571) |
| UK | 0.870 | (0.135)*** | 0.762 | (0.162)*** | 1.046 | (0.227)*** |

NOTE: The dependent variable is Offers (Greece baseline). Number of observations: 990, 528 and 462 respectively (30, 16 and14 individuals, 33 periods). Standard error of the estimates between parentheses. Significant estimates at the 10%, 5%, and 1% significance level are marked with *, **, and ***, respectively. Wald test (Spain=UK) *p*-values = 0.024, 0.062 and 0.045, respectively.

Table 6 presents the results from the estimation of a model describing employees' behavior. Specifically, we estimate the effects of the explanatory variables on employees' probability to reject a given wage offer. As we would expect, higher offers entail a lower rejection probability. This may be due to the fact that employees tend to accept with total certainty offers above a given (high) threshold and reject all offers below another (low) threshold. Overall, we observe no differences among

**TABLE 6**

**MODELS FOR THE REJECTION PROBABILITIES BY GENDER**

| | dy/dx(Overall) | | dy/dx(Male) | | dy/dx(Female) | |
|---|---|---|---|---|---|---|
| Offer | –0.324 | (0.042)*** | –0.305 | (0.108)*** | –0.321 | (0.040)*** |
| T | –0.002 | (0.002) | –0.006 | (0.003)** | 0.001 | (0.002) |
| Certainty equivalent | 0.031 | (0.016)* | 0.023 | (0.018) | 0.037 | (0.024) |
| Spain | 0.057 | (0.126) | –0.162 | (0.103) | 0.076 | (0.102) |
| UK | 0.336 | (0.135)** | 0.511 | (0.172)*** | 0.027 | (0.130) |

NOTE: Marginal effects after random effects probit regression (Greece baseline). Number of observations: 990, 528 and 462 respectively (30, 16 and 14 individuals, 33 periods). Standard error of the estimates between parentheses. Significant estimates at the 10%, 5%, and 1% significance level are marked with *, **, and ***, respectively. Wald test (Spain=UK) *p*-values = 0.035, 0.003 and 0.693, respectively.

employees from Greece and Spain. On the contrary, British employees have a higher probability of rejecting a given salary. Specifically, British male-employees are more likely to reject a given offer than other males. However, no differences of rejection probabilities among countries are found for female-employees.

In a similar way as Table 6, Table 7 presents the results from the estimation of a model describing employees' behavior each country. In this estimation, we include a male variable where we can see that Greek and Spanish females reject with a higher probability, whereas in the UK, females reject less than males.

**TABLE 7**

**MODELS FOR THE REJECTION PROBABILITIES BY COUNTRY**

| | dy/dx(Greece) | | dy/dx(Spain) | | dy/dx(UK) | |
|---|---|---|---|---|---|---|
| Offer | –0.587 | (0.179)*** | –0.225 | (0.036)*** | –0.405 | (0.075)*** |
| T | –0.003 | (0.003) | –0.001 | (0.003) | –0.006 | (0.003)* |
| Certainty equivalent | 0.045 | (0.048) | 0.130 | (0.046)*** | 0.027 | (0.019) |
| Male | –0.248 | (0.134)* | –0,350 | (0.057)*** | 0.297 | (0.126)** |

NOTE: Marginal effects after random effects probit regression. Number of observations: 330 (30 individuals and 33 periods). Standard error of the estimates between parentheses. Significant estimates at the 10 %, 5 %, and 1 % significance level are marked with *, **, and ***, respectively.

## 4. Conclusions

In this paper, using an UG with a labor market framing, we study behavior of subjects from Greece, Spain and the UK. Sharp differences are observed between the two Mediterranean countries and the UK that may depend on country differences on the levels of income and/or other macroeconomic determinants of subjects' opportunity costs. Namely, higher offers and lower propensity to accept in the UK could be the result of the country's long history of unionization and a higher standard of living affecting beliefs on both sides of the game. A large part of these differences are due to gender-specific patterns. On one hand, females-employers in the UK are among the most generous proposers ever reported on UG experiments. On the other hand, male-employees from the UK are those most likely to reject a given offer.

Distinguishing by gender, we observe that offers made by Spanish females are not significantly different from the ones made by Greek females. Furthermore, offers made by British females are significantly higher than offers made by Spanish and Greek females. Salaries present the same pattern as offers. With respect to the rejection rate, we find that Greek and Spanish women reject more than men whereas in the UK the contrary holds, namely men reject more than women.

The recent economic crisis in Europe has triggered many discussions on the role of country differences in the functioning of the markets and policy makers should not underestimate this fact.

Our results are rather specific to the labor market. Specifically, we have found that British subjects have behaved in a fairer way than their South-European counterparts. Thus, the British pool has generated more efficiency despite the higher willingness to reject shown by the British employee-subjects. It will remain unanswered here whether the differences discussed above are the cause (as a pre-existing cultural feature) or the result of larger scale differences, but it seems that a clear North-South component exists.

## References

Botelho, A., Harrison, G. W., Hirsch, M. A., & Rutstrom, E. E. (2001). *Bargaining behavior, demographics and nationality: A reconsideration of the experimental evidence* (NIMA Working Paper No. 16). Universidad do Minho.

Camerer, C., & Fehr, E. (2004). Measuring social norms and preferences using experimental games: A guide for social scientists. In J. Henrich, R. Boyd, S. Bowles, C. Camerer, E. Fehr & H. Gintis (Eds.), *Foundations on Human Sociality: Economic Experiments and Ethnographic Evidence from Fifteen Small-Scale Societies* (pp. 55-95). Oxford University Press. http://dx.doi.org/10.2139/ssrn.299143

Chuah, S.-H., Hoffmann, R., Jones, M., & Williams, G. (2007). Do cultures clash? Evidence from cross-national ultimatum game experiments. *Journal of Economic Behavior and Organization, 64*, 35-48. https://doi.org/10.1016/j.jebo.2006.04.006

Chuah, S.-H., Hoffmann, R., Jones, M., & Williams, G. (2009). An economic anatomy of culture: Attitudes and behaviour in inter- and intra-national ultimatum game experiments. *Journal of Economic Psychology*, *30*, 732-744. https://doi.org/10.1016/j.joep.2009.06.004

Croson, R., & Buchan, N. (1999). Gender and culture: International experimental evidence from trust games. *The American Economic Review, Papers and Proceedings, 89*, 386-391. https://www.jstor.org/stable/117141

Eckel, C., & Grossman, P. (2001). Chivalry and solidarity in ultimatum games. *Economic Inquiry, 39*, 171-188. https://doi.org/10.1111/j.1465-7295.2001.tb00059.x

Eckel, C., & Grossman, P. (2008). Differences in the economic decisions of men and women: Experimental evidence. In C. Plott & V. Smith (Eds.) *Handbook of Experimental Economics Results Volume 1* (pp. 509-519). Elsevier. https://doi.org/10.1016/S1574-0722(07)00057-1

Filippin, A., & Crosetto, P. (2016). A reconsideration of gender differences in risk attitudes. *Management Science*, *62*(11), 3138-3160. https://doi.org/10.1287/mnsc.2015.2294

Fischbacher, U. (2007). z-Tree: Zurich toolbox for readymade economic experiments. Experimenter's manual. *Experimental Economics, 10*(2), 171-178.

García-Gallego, A., Georgantzís, N., & Jaramillo-Gutiérrez, A. (2008). Ultimatum salary bargaining with real effort. *Economics Letters, 98*(1), 78-83. https://doi.org/10.1016/j.econlet.2007.04.009

García-Gallego, A., Georgantzís, N., & Jaramillo-Gutiérrez, A. (2012). Gender differences in ultimatum games: Despite rather than due to risk attitudes. *Journal of Economic Behavior and Organization*, *83*(1), 42-49. https://doi.org/10.1016/j.jebo.2011.06.012

Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: Behavioral experiments in 15 small-scale societies. *The American Economic Review, Papers and Proceedings, 91*, 73-78.

Henrich, J., McElreath, R., Barr, A., Ensminger, J., Barrett, C., Bolyanatz, A., Cardenas, J.C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., Ziker, J. (2006). Costly Punishment Across Human Societies. *Science*, *312*(5781), 1767-1770.

Henrich, J., Ensminger, J., McElreath, R., Barr, A., Barrett, C., Bolyanatz, A., Cardenas, J. C., Gurven, M., Gwako, E., Henrich, N., Lesorogol, C., Marlowe, F., Tracer, D., & Ziker, J. (2010). Markets, Religion, Community Size, and the Evolution of Fairness and Punishment. *Science*, *327*(5972), 1480-1484.

Hofstede, G. (1991). *Cultures and organizations: Software of the mind*. McGraw-Hill.

Inglehart, R. (2000). Culture and democracy. In L.E. Harrison & S.P. Huntington (Eds.), *Culture Matters: How Values Shape Human Progress* (pp. 80-97). Basic Books.

Oosterbeek, H., Sloof, R. S., & van de Kuilen, G. (2004). Cultural differences in ultimatum game experiments: Evidence from a meta-analysis. *Experimental Economics, 7*(2), 171-188.

Roth, A. E., Prasnikar, V., Okuno-Fujiwara, M., & Zamir, S. (1991). Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburg and Tokyo: An experimental study. *The American Economic Review, 81*(5), 1068-1095. https://www.jstor.org/stable/2006907

Saad, G., & Gill, T. (2001). Sex differences in the ultimatum game. *Journal of Bioeconomics, 3*(2-3), 171-193.

Solnick, S. J. (2001). Gender differences in the ultimatum game. *Economic Inquiry, 39*(2), 189-200. https://doi.org/10.1111/j.1465-7295.2001.tb00060.x

## APPENDIX A
## Macroeconomics indicators

In Table A we include, for each country, data taken from the Eurostat database, referred to several macroeconomic magnitudes for the year in which our experimental sessions took place. Specifically, the table includes data for the GDP per capita, inflation rate, minimum wage, unemployment rate by gender, the gender pay gap in unadjusted form and the average gross annual earnings in industry and services.

**TABLE A**
**SOME MACROECONOMICS INDICATORS FOR THE THREE COUNTRIES CONSIDERED**

| | | Greece | Spain | UK |
|---|---|---|---|---|
| GDP per capita | | 92.1 | 101 | 121.8 |
| Inflation rate | | 2.4 | 3.1 | 1.4 |
| Minimum wage | | 605 | 526 | 1,106 |
| Unemployment rate by gender | *Total* | 9.7 | 11.1 | 5 |
| | *Males* | 6.2 | 8.2 | 5.5 |
| | *Females* | 15 | 15.3 | 4.3 |
| Gender pay gap in unadjusted form | | 11 | 18 | 22 |
| Average gross annual earnings in industry and services | | 16,738.5 | 19,220 | 38,792.5 |

NOTE: *GDP per capita* in purchasing power standards (PPS) (EU-27=100). *Inflation rate* calculated as the annual average rate of change in Harmonized Indices of Consumer Prices. *Minimum wage* expressed in Euros per month. *Unemployment rate by gender* calculated as a % of the labor force, taking persons aged 15-74 who were: a) without work during the reference week; b) currently available for work; c) actively seeking work. *Gender pay gap in unadjusted form* calculated as the difference between men's and women's average gross hourly earnings as a % of men's average gross hourly earnings. *Average gross annual earnings in industry and services* taken from the full-time employees in firms with 10 or more employees.

**SOURCE:** Eurostat data base. Data referred to year 2003.

## APPENDIX B
## Instructions to Experimental Subjects[10]

You are going to participate in an experiment about individual decision making, that will last approximately 90 minutes. You must follow the instructions carefully and, depending on your performance, you may earn a considerable amount of money. The context in which you will have to take your decisions is described below.

This session will consist of two main parts:

- **First part**: 30-35 rounds of a salary-and-task-negotiation session.
- **Second part**: task-performing and payment.

- **First Part**:

An equal number of employers and employees are in a group of 20 individuals. At the beginning of the session, you will be randomly assigned the role of employer or employee. In each period, you will be randomly matched with a player of the other type (if you are an employer, you will be matched with an employee and vice versa). The experiment will be repeated over 30 to 35 periods (randomly determined by the server).

In each period, each employer-employee pair is faced with the following situation: The employer offers the employee a share from a 10 m.u. profit yielded from the task (filling 20 envelopes numbered, from 1 to 20, with their respective one-page letter, also numbered from 1 to 20) which will be performed (in the second part of the session) by the employee. If the employee accepts, the task will have to be performed by the employee (see "Second Part") and the two players' earnings are determined as proposed by the employer. Otherwise, the task is not performed and both players earn nothing.

If you are an employer, your decision will consist of offering a salary to the employee. Such a salary will be a quantity between 0 and 10 m.u., in multiples of 0.10 m.u.. If you are an employee, your decision will consist of accepting or rejecting the salary offered by the employer.

- **Second part**:

Your payment (and the tasks to perform if you are an employee) will be determined according to the outcome of five periods, which will be randomly chosen among the total number of periods played during this session. A minimum of 90% (at least 18 out of 20 envelopes must contain the correct sheet) reliability will be required for each task unit to be considered successfully performed.

[10] This is the version of the instructions given to subjects in the UK. In Greece and Spain, the subjects were given the corresponding instructions in Greek and Spanish, respectively. The monetary units used were the corresponding of each country (pounds in the UK and euro in Greece and Spain).

## APPENDIX C
## Questionnaire

*Control Question*

Imagine you are assigned the role of an "employee" in the following hypothetical market situation: An equal number of employers and employees are in a group of 20 individuals, forming random employee-employer pairs. You are going to negotiate your share over a total of 10 m.u. earned by one of the employers from the task you will perform (filling 20 envelopes numbered, from 1 to 20, with their respective one-page letter, also numbered from 1 to 20). If you accept the salary, you will perform the task and earnings for both, you and your employer, will be determined as proposed by the employer. If you reject the salary, the task is not performed and you both earn nothing. Alternatively, to your earnings and task-performing obligations, you may prefer a certain payoff, whose value is provided below, under 20 different scenarios. Please mark with an "X" your preferred option in each one of the following scenarios:

- **SCENARIO 1**: You are offered an alternative of a certain payment of 0.5 m.u.
  Do you prefer the certain payoff? ..... ☐
  Or your earnings from the above hypothetical market situation?..... ☐
- **SCENARIO 2**: You are offered an alternative of a certain payment of 1 m.u.
  Do you prefer the certain payoff? ..... ☐
  Or your earnings from the above hypothetical market situation?..... ☐
- **SCENARIO 3**: You are offered an alternative of a certain payment of 1.5 m.u.
  Do you prefer the certain payoff? ..... ☐
  Or your earnings from the above hypothetical market situation?..... ☐

  ...
- **SCENARIO 20**: You are offered an alternative of a certain payment of 10 m.u.
  Do you prefer the certain payoff? ..... ☐
  Or your earnings from the above hypothetical market situation?..... ☐

CUADERNOS ECONÓMICOS ICE INFORMACIÓN COMERCIAL ESPAÑOLA

# Conflicto 2 × 2: una lección de teoría de juegos en el laboratorio*

# *2 × 2 Conflict Game: A Game Theory Lesson in the Laboratory*

Luis Alejandro Palacio García
Alexandra Cortés Aguilar
Universidad Industrial de Santander

***Resumen***

*Los experimentos en el aula son una estrategia pedagógica que complementa las lecciones magistrales y los talleres de ejercicios, lo que los consolida como una alternativa práctica e interactiva que propicia la discusión de los conceptos y el fortalecimiento de habilidades transversales. En particular, en este artículo se presenta un protocolo que permite modelar diferentes niveles de conflicto en el laboratorio, donde dilemas clásicos como cazar ciervo, la gallina y el dilema del prisionero se consideran casos particulares. Las preguntas que se discuten a la luz de la experiencia son: ¿Los estudiantes están dispuestos a cooperar en función del nivel de conflicto al que se enfrenten? ¿Es mayor la cooperación en el juego secuencial en comparación al juego simultáneo? Los resultados muestran que el nivel de conflicto afecta la decisión de cooperar. Además, los incentivos para actuar agresivamente aumentan entre mayor sea el conflicto. Con respecto al protocolo secuencial, no hay evidencia de cambios significativos en el comportamiento con respecto al juego simultáneo.*

***Palabras clave:*** *investigación formativa, negociación, coordinación, dilemas sociales.*
***Clasificación JEL:*** *A22, C91, D91.*

***Abstract***

*Classroom experiments are a pedagogical strategy that complements master lessons and problem sets, which consolidates them as a practical and interactive alternative that encourages the discussion of concepts and the strengthening of transversal skills. In particular, this article presents a protocol that allows modeling different conflict levels in the laboratory, where classic dilemmas games such stag hunt, chicken and the prisoner's dilemma are specific cases. Two questions are discussed based on experience: Are students willing to cooperate given the conflict level they face? Is the cooperation greater in sequential game than in simultaneous game? Results show that conflict affects the decision to cooperate. In addition, incentives to act aggressively increase with greater conflict levels. Respect to the sequential protocol, there is no evidence of significant changes in behavior compare to simultaneous game.*

***Keywords:*** *formative research, bargaining, coordination, social dilemmas.*
***JEL Classification:*** *A22, C91, D91.*

* Este artículo hace parte de los resultados de los proyectos de investigación «Formación en competencias ciudadanas en educación superior» código 8265, financiado por el Instituto Colombiano para la Evaluación de la Educación y la Universidad Industrial de Santander; y «Filosofía moral experimental para la formación de competencias ciudadanas» código 2403, financiado por la Universidad Industrial de Santander.


DOI: https://doi.org/10.32796/cice.2020.99.7013
Recibido: Enero 2020 · Aceptado: Febrero 2020

## 1. Introducción

Los experimentos en el aula surgen como una alternativa didáctica que busca mejorar el aprendizaje de los conceptos propios de la teoría económica, complementando las lecciones tradicionales con la contrastación empírica. Controlar los supuestos y recrear los incentivos permite una comprensión rápida e intuitiva (Holt, 1996). Aunque es válido diseñar actividades lúdicas con lápiz y papel en un aula convencional, es conveniente contar con un espacio físico que facilite los procedimientos logísticos; es decir, un laboratorio. Estos elementos pedagógicos brindan al estudiante la oportunidad de vivir y descubrir las conclusiones de los modelos formales, sistematizando las decisiones con el fin de evidenciar el efecto de diferentes variables de interés (Emerson y Hazlett, 2012).

El objetivo de este artículo es presentar un protocolo experimental diseñado para facilitar la explicación de los conceptos básicos de la teoría de juegos, como el equilibrio de Nash y la perfección en subjuegos. La clase se realiza en el laboratorio, situando al estudiante en un problema de acción colectiva enfocado sobre dos ejes temáticos: la confianza y el conflicto. Bajo este marco, se hace una adaptación del modelo propuesto por Palacio *et al*. (2015b), permitiendo analizar y discutir el efecto del nivel de conflicto y los mensajes vinculantes sobre la decisión de cooperar en juegos $2 \times 2$. Dado este enfoque, los libros de Dixit y Nalebuff (2010) y Brañas-Garza (2011) son la base para la discusión teórica que se complementa con diversos dilemas sociales que permiten analizar las decisiones y discutir sus consecuencias, valorando los argumentos a favor y en contra de diferentes cursos de acción.

De esta forma, los estudiantes conocen que hacen parte de una investigación-acción participativa, donde no se les consideran el «objeto de estudio» sino el «sujeto protagonista». En otras palabras, se busca entender el proceso de enseñanza-aprendizaje al tratar de cambiarlo colaborativa y reflexivamente[1]. Aunque la economía experimental es el referente general, es importante hacer dos aclaraciones conceptuales. Primero, a diferencia de un experimento, la simulación de un dilema social es una actividad formativa que busca garantizar un ambiente democrático, donde el docente ejerce principalmente como mediador. Las decisiones en los ordenadores se acompañan del debate continuo y reflexivo, en el que se cuestiona y discute a medida que se está viviendo la experiencia (Stodder, 1998). Segundo, no se garantiza el valor inducido; en su lugar, se hace explícito un sistema de puntos acorde con las decisiones, sin pagos monetarios, como mecanismo de incentivo a la participación de los estudiantes en las diferentes etapas del juego (Rousu *et al*., 2015).

Con estas aclaraciones, el juego del conflicto $2 \times 2$ se presenta a los estudiantes como un reto interactivo multijugador que simula diferentes niveles de alineación de las preferencias. Esto significa que se modela el nivel de conflicto como variable

[1] La efectividad de la actividad formativa para mejorar la comprensión de los conceptos propios de teoría económica de la elección racional se ha discutido en Emerson y Taylor (2004), Dickie (2006) y Emerson y English (2016), entre otros.

tratamiento, evidenciando un orden lógico en los parámetros. Adicionalmente, los estudiantes se enfrentan primero al juego donde toman decisiones simultáneamente, para luego pasar al protocolo secuencial. De esta forma, en la clase se discuten las siguientes preguntas orientadoras: ¿Los estudiantes están dispuestos a cooperar en función del nivel de conflicto al que se enfrenten? ¿El hecho de jugar de forma secuencial aumenta la cooperación en comparación al juego simultáneo?

El diseño y validación de este juego recoge la experiencia acumulada en el grupo de investigación en Economía Aplicada y Regulación (EMAR)[2] de la Universidad Industrial de Santander, UIS (Palacio *et al*., 2017). A efectos de la contrastación empírica, en este artículo se presentarán los datos recolectados en el taller «Juegos de Negociación», realizado con 98 estudiantes de primer semestre matriculados en la Cátedra Vida y Cultura Universitaria, UIS. Se escogió este espacio porque permitía convocar de una forma organizada a estudiantes de diferentes programas, consolidando una muestra más amplia y diversa, comparada con el curso tradicional de Teoría de Juegos que se imparte únicamente a estudiantes de economía.

Los resultados de la implementación experimental permiten responder las dos preguntas orientadoras que se discuten en la clase, evidenciando si el nivel de conflicto, entendido como la desigualdad en los pagos, efectivamente afecta la decisión de cooperar. Los datos indican que la probabilidad de actuar de forma agresiva aumenta con el nivel de conflicto, sacrificando la eficiencia. Así, la simulación de este dilema social lleva a reflexionar sobre cómo los incentivos externos influyen en la decisión de tomar un curso de acción más agresivo y egoísta. Con respecto al protocolo secuencial, no hay evidencia de cambios significativos en el comportamiento con respecto al juego simultáneo.

Con estos elementos, en la sección 2 se presenta una síntesis de la literatura experimental sobre juegos de coordinación y se realiza la descripción del modelo formal que captura los diferentes niveles de conflicto, resaltando los equilibrios de Nash acorde con cada caso particular. En la sección 3 se describe el protocolo experimental que permite el contraste empírico, con el propósito de validar las observaciones y experiencias particulares que discuten los estudiantes en las clases de teoría de juegos. Por último, en las conclusiones se realizan algunas recomendaciones para llevar a cabo este tipo de actividades pedagógicas.

## 2. Coordinación: conflicto y negociación

En la vida cotidiana existe un sinnúmero de situaciones en las que es necesario coordinar las decisiones de diferentes personas. Aunque algunas veces puede parecer fácil, en otras ocasiones los incentivos contradictorios dificultan la coincidencia. En

[2] El protocolo experimental se diseñó en alianza con Estévez (2015), programado inicialmente en el software Z-Tree (Fischbacher, 2007). En la actualidad el juego del conflicto 2×2 se realiza en línea gracias al software OTree (Chen, *et al*., 2016).

los juegos de coordinación los agentes tienen varias opciones y la ganancia es más alta si todos concuerdan en una acción en particular. Por tanto, es esencial analizar cuándo y cómo se puede mejorar la coordinación, entendiendo el conflicto como una serie de negociaciones, ofertas y contraofertas, con el fin de caracterizar las estrategias óptimas. Esto implica recalcar la idea de que, además de la divergencia de intereses, existe un poderoso interés común en llegar a una solución que evite caer en el punto de desacuerdo[3].

En esta sección se presenta la discusión teórica que se ha dado a partir de la literatura experimental relacionada con juegos de coordinación. Posteriormente, se describe en detalle el modelo de referencia para la clase de teoría de juegos en el laboratorio. El protocolo de juego secuencial toma como inspiración el juego del conflicto 2×2 originalmente propuesto por Baliga y Sjöström (2004), y discutido ampliamente en Palacio *et al*. (2015b). Este juego de referencia es útil porque es un modelo estilizado que captura diferentes niveles de alineación en las preferencias, donde juegos clásicos como cazar ciervo, halcón y paloma y el dilema del prisionero son considerados como casos particulares.

### *2.1. Literatura experimental*

En términos formales, los juegos de coordinación se caracterizan por tener múltiples equilibrios de Nash, lo que dificulta la predicción sobre el comportamiento esperado en el laboratorio (García-Gallego y Miller, 2011). Así, aunque exista un equilibrio de Nash dominante en ganancia, este no siempre es el que se evidencia efectivamente (Van Huyck *et al*., 1990). Ello apoya la hipótesis de que la dominancia en ganancias no es un criterio de selección en juegos de coordinación (Cooper y DeJong, 1990). En este contexto, para lograr la coordinación los agentes necesitan algún mecanismo para elegir y seleccionar uno de los equilibrios de Nash.

En la literatura experimental se ha estudiado la comunicación como mecanismo de coordinación en la negociación. Cooper *et al*. (1992) presentan evidencia experimental sobre la importancia de la comunicación no vinculante en juegos de coordinación. Según los autores, si se permite una fase de comunicación previa al juego que es completamente independiente de las acciones que generan los pagos, estos mensajes pueden influir en el resultado final, afectando las creencias que los agentes tienen sobre las acciones del oponente. Por su parte, Brandts y Charness (2000) muestran que cambiar el orden de la elección entre el mensaje y la acción cambia el comportamiento de los jugadores. La coordinación mejora cuando se escoge primero una señal y luego la acción que generará los pagos.

[3] Estos problemas se discuten en teoría de juegos gracias a los aportes pioneros de Schelling (2006) y Aumann (2006), que en 2005 recibieron el Premio Nobel en Economía. Thomas Schelling propuso que los juegos se deben concebir desde un enfoque táctico para abordar el tema de la negociación bilateral. Por su parte, la contribución de Robert Aumann consiste en entender las transacciones como juegos repetidos con el fin de analizar los incentivos a mantener la cooperación.

En este contexto, el juego de halcón y paloma con información perfecta es el modelo ideal para estudiar los problemas de coordinación y conflicto, porque los pagos asimétricos en equilibrio aumentan la posibilidad de caer en el punto de desacuerdo. La evidencia experimental se puede consultar en los trabajos de Bornstein *et al*. (1997), Duffy y Feltovich (2002) y Neugebauer *et al*. (2008). Los efectos del procedimiento experimental sobre el comportamiento, siguiendo los trabajos de Brandts y Charness (2000; 2011) se estudian en Palacio *et al*. (2015a).

Sin duda un juego que se ha estudiado en profundidad es el dilema del prisionero. Como lo afirman Andreoni y Varian (1999), la evidencia que se ha recopilado es muy interesante porque muestra la divergencia entre equilibrio y eficiencia. La literatura experimental ha profundizado en el dilema a partir del juego de contribución voluntaria a los bienes públicos. En este sentido, Zelmer (2003) realiza un meta-análisis que permite estimar el efecto de una serie de factores, en una muestra de 27 estudios relevantes y representativos sobre el tema. A la par, se han propuesto una serie de modelos teóricos a raíz de los resultados empíricos, como es el caso de Fehr y Schmidt (1999) que exploran las preferencias sobre la distribución de los pagos; Fehr y Gächter (2000) y Rabin (1993) quienes trabajan el concepto de reciprocidad; así como Fehr y Fischbacher (2004) que estudian el papel de las normas sociales. Cooper y Kagel (2016) presentan una revisión de estudios experimentales sobre preferencias sociales.

Con respecto a los experimentos en el aula, Alba-Fernández *et al*. (2006) muestra que el concepto de equilibrio de Nash se puede enseñar por medio del «concurso de belleza», un simple juego de adivinanzas en el que la repetición permite a los estudiantes reaccionar a las elecciones de otros jugadores, acercándose iterativamente a la solución de equilibrio. Por su parte, Rebelein y Turkay (2016) analizan cómo la información al momento de decidir puede afectar dramáticamente las ganancias de la empresa y los resultados del mercado. Recreando un modelo de Stackelberg muestran que existe una ventaja de primer movimiento cuando las empresas eligen cantidades de producción. Por el contrario, hay una ventaja de segundo movimiento cuando las empresas eligen los precios. Por último, Murphy y Cárdenas (2004) se enfocan en enseñar el efecto de la regulación en el cumplimiento para reducir la extracción de un recurso local. En particular, encuentran que la comunicación no vinculante puede ser notablemente eficaz en la implementación de políticas ambientales.

A partir de los elementos de coordinación, comunicación y conflicto que han sido discutidos desde la economía experimental, en las siguientes secciones se describen la estructura del modelo de negociación propuesto para llevar estos conceptos al salón de clase, el protocolo de implementación y los hallazgos más sobresalientes.

### *2.2. Modelo de referencia: Conflicto 2 × 2*

Existen dos jugadores, Azul y Verde, quienes toman su decisión de forma simultánea, escogiendo entre dos alternativas, $p$ o $h$. Las estrategias se deben interpretar

a la luz de la metáfora, donde *p* es comportarse como paloma (negociador cooperador), y *h* es comportarse como halcón (negociador agresivo). La función de utilidad queda definida en la matriz de pagos que aparece en el Cuadro 1. Es importante resaltar que en términos sociales es mejor cuando los dos cooperan (1000,1000), comparado con el punto de desacuerdo ineficiente que se alcanza cuando los dos se comportan agresivamente (250, 250).

**CUADRO 1**
**MATRIZ DE PAGOS**

| | **p** | **h** |
|---|---|---|
| **p** | 1000 , 1000 | X , Y |
| **h** | Y , X | 250 , 250 |

**FUENTE**: Elaboración propia.

Adicionalmente, en el juego existen dos parámetros, X y Y. El valor de X varía en el intervalo de 0 y 500, mientras Y varía en el intervalo de 500 y 1500. Por tanto, Y siempre será mayor o igual a X, lo que significa que, al enfrentarse un halcón con una paloma, quien toma la postura agresiva siempre obtendrá pagos superiores. Es posible interpretar el parámetro Y como la tentación a actuar agresivamente, dado que entre mayor es su valor, más incentivos se tienen a ser halcón. Por su parte, el parámetro X denota qué tan mal lo pasa una paloma cuando se enfrenta con un halcón. Esto lleva a reconocer que la diferencia entre estos parámetros (Y-X) es cuantitativamente una medida del nivel de conflicto que permite caracterizar intervalos, evidenciando un orden lógico como se ilustra en el Cuadro 2.

Dado los supuestos del juego, se deben analizar cuatro casos particulares para calcular el equilibrio de Nash en estrategias puras. En el juego C1 las preferencias de los jugadores están bien alineadas y no hay ningún problema de coordinación porque el equilibrio es único en estrategias dominantes. En este ambiente es muy fácil alcanzar el resultado socialmente deseable. En el juego C2 la mutua cooperación es un equilibrio de Nash, pero no es único en estrategias puras. El problema radica en coordinar entre un equilibrio Pareto dominante ( *p, p*), y un equilibrio dominante en

**CUADRO 2**
**CUATRO NIVELES DE CONFLICTO**

| | **Parámetros** | **Nivel de conflicto. Y-X** |
|---|---|---|
| C1. Estrategia paloma dominante | Y < 1000; X > 250 | [0, 750] |
| C2. Complementos estratégicos | Y < 1000; X < 250 | [250, 1000] |
| C3. Sustitutos estratégicos | Y > 1000; X > 250 | [500, 1250] |
| C4. Estrategia halcón dominante | Y > 1000; X < 250 | [750, 1500] |

**FUENTE**: Elaboración propia.

el riesgo ($h$, $h$). Este es el juego conocido como cazar ciervo, que formaliza la idea que la desconfianza puede causar una escalada de agresiones. También es conocido como complementos estratégicos, dado que la mejor respuesta es hacer lo mismo que el oponente.

El juego C3 es un ambiente más complejo porque ahora existen dos problemas simultáneamente, uno de coordinación y otro distributivo. En equilibrio solo uno debe comportarse agresivamente, poniendo el resultado a su favor; aunque seleccionar quién será halcón y quién paloma es un problema de negociación tácita. En este ambiente no existe un criterio para predecir el comportamiento y, por tanto, el resultado final. Es el juego clásico conocido como la gallina, un modelo de anticipación y disuasión, donde el miedo a las consecuencias hace que una de las partes se rinda. Dado que la mejor respuesta es hacer lo contrario que el oponente, cumple la definición de sustitutos estratégicos.

El último juego es el de mayor conflicto, porque es el clásico dilema social sobre las limitaciones del comportamiento racional para asignar eficientemente los recursos. Se ha clasificado el juego C4 como el más conflictivo dado que los jugadores se enfrentan a un ambiente donde la elección racional predice claramente alcanzar el punto de desacuerdo. Este es el clásico dilema del prisionero, donde los incentivos individuales llevan a una asignación ineficiente de los recursos. Un resumen de estos resultados teóricos se presenta en el Cuadro 3.

**CUADRO 3**
**EQUILIBRIOS DE NASH EN EL JUEGO DEL CONFLICTO 2×2**

| | **Estrategias** (a* , v*) | **Utilidades** (Ua , Uv) | **Eficiente** |
|---|---|---|---|
| C1. Estrategia paloma dominante | (p , p) | (1000 , 1000) | Sí |
| C2. Complementos estratégicos | (p , p)<br>(h , h) | (1000 , 1000)<br>(250 , 250) | Sí<br>No |
| C3. Sustitutos estratégicos | (p , h)<br>(h , p) | (X , Y)<br>(Y , X) | Sí<br>Sí |
| C4. Estrategia halcón dominante | (h , h) | (250 , 250) | No |

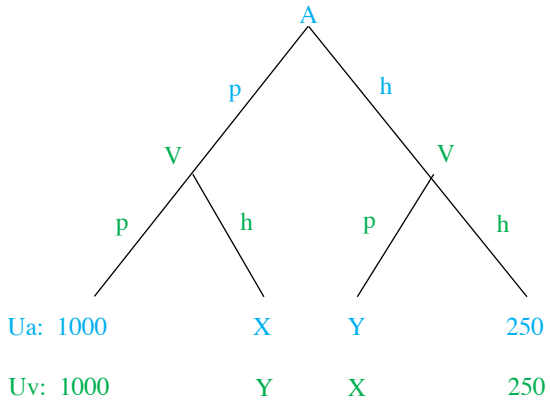
**FUENTE**: Elaboración propia.

### *2.3. El juego del conflicto secuencial*

Considérese ahora una situación donde el jugador Azul mueve primero, luego, el jugador Verde observa la acción escogida y tiene su turno para decidir. En términos formales, esto simplemente es pasar de un juego 2×2 en forma estratégica, a un juego en forma extensiva con información perfecta, como aparece en la Figura 1. Una estrategia del jugador Verde es una especificación de una acción ante cada posible acción del jugador Azul. Por tanto, el conjunto de estrategias del Verde ahora es una

regla de respuesta, de tal forma que el primer componente denota la acción que escoge si observa *p*, y el segundo componente especifica la acción en caso de observar *h*. El conjunto de estrategias del Azul no se modifica porque tiene solo un conjunto de información.

**FIGURA 1**
**JUEGO DEL CONFLICTO CON INFORMACIÓN PERFECTA**




**FUENTE**: Elaboración propia.

Dentro de este marco, el compromiso racional hace referencia a los mecanismos que permiten hacer público que la posición de una de las partes es inamovible, porque con premeditación se han rechazado todas aquellas alternativas que pudieran llevar a un resultado desfavorable. La idea es cambiar las expectativas del rival sobre las respuestas a sus acciones. Esta táctica de negociación es lo que se ha denominado jugada estratégica. En términos formales, el conjunto de estrategias del jugador Verde ahora es más amplio, haciendo que las predicciones basadas en el equilibrio de Nash en estrategias puras sean menos relevantes. Por esta razón, en el juego del conflicto secuencial el concepto de equilibrio aplicable es el Equilibrio de Nash Perfecto en Subjuegos (ENPS), como aparece en el Cuadro 4.

**CUADRO 4**
**ENPS EN EL JUEGO DEL CONFLICTO 2×2**

| | **Estrategias** (a* , v*) | **Utilidades** (Ua , Uv) | **Eficiente** |
|---|---|---|---|
| C1. Estrategia paloma dominante | (p , pp) | (1000 , 1000) | Sí |
| C2. Complementos estratégicos | (p , ph) | (1000 , 1000) | Sí |
| C3. Sustitutos estratégicos | (h , hp) | (Y , X) | Sí |
| C4. Estrategia halcón dominante | (h , hh) | (250 , 250) | No |

**FUENTE**: Elaboración propia.

Como se puede observar, la posibilidad de enviar mensajes vinculantes ayuda a solucionar el problema de coordinación presente en los juegos con múltiples equilibrios. En otras palabras, en los juegos C2 y C3 se espera que el jugador Azul seleccione aquel equilibrio que le sea favorable. Esto puede ser mutuamente ventajoso en la medida en que se evita caer en el resultado ineficiente de mutua agresión. Sin embargo, en los casos extremos C1 y C4 los mensajes no aportan mucho, dado que los incentivos llevan a los mismos resultados que en el protocolo simultáneo.

## 3. Clase de teoría de juegos en el laboratorio

Según Emerson (2014), los experimentos en el aula proporcionan a los estudiantes oportunidades para descubrir y experimentar los conceptos económicos por sí mismos. Por tanto, al diseñar nuevos protocolos se debe proceder con este propósito en mente: transmitir la teoría económica a los estudiantes con eficacia. Para crear una experiencia de aprendizaje valiosa, el profesor elige un juego que promueva la comprensión de un concepto importante, complementando la explicación con un ejemplo práctico. Los experimentos efectivos involucran a los estudiantes en la toma de decisiones, manteniendo su atención sin confundirlos con una complejidad innecesaria.

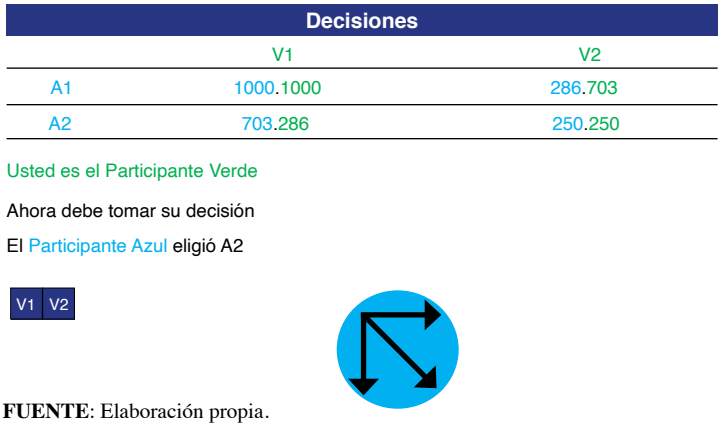
Los estudiantes deben estar motivados para escuchar todas las instrucciones, además de tener la oportunidad de hacer preguntas aclaratorias. Para minimizar la frustración, los estudiantes deben entender la tarea que van a realizar, necesitando en muchas ocasiones de una ronda de experiencia para entender la mecánica del ejercicio y formular preguntas de procedimiento (Emerson y Hazlett, 2012). En la experiencia que se describe en este documento se aclara a los participantes que su objetivo es ganar la mayor cantidad de puntos, que dependen tanto de sus decisiones como de las decisiones de los demás participantes. Se insiste en que no existen respuestas incorrectas pues no se trata de generar un comportamiento concreto. Se pide que se sientan libres de tomar sus decisiones y opinen abiertamente lo que piensan en todo momento.

### *3.1. Protocolo de laboratorio*

El experimento consta de 20 periodos, 10 periodos iniciales en el protocolo simultáneo y 10 periodos finales en el protocolo secuencial. En los periodos finales la única diferencia es que el participante Azul jugará primero, por lo que el participante Verde observa la decisión del Azul antes de decidir. Los estudiantes son emparejados al azar al inicio de cada periodo con alguien que se encuentra dentro de la sala. Luego se explica que cada estudiante debe decidir entre dos acciones, 1) Paloma o 2) Halcón, y los pagos se calculan de acuerdo con la Matriz de pagos presentada en el Cuadro 1. Es importante tener presente que hay dos variables, X e Y. Estas variables

tomarán un valor diferente para cada pareja al inicio de cada periodo, y serán mostradas al azar antes de tomar la decisión. El valor de X siempre será un número al azar entre 0 y 500 y el de Y será un número al azar entre 500 y 1500. A continuación, se presenta un ejemplo de la tarea realizada en el tratamiento secuencial, donde el jugador Verde puede observar la decisión tomada por el Azul (Gráfico 1).

**GRÁFICO 1**
**INTERFAZ DE USUARIO OTree**




**FUENTE**: Elaboración propia.

Para facilitar la comprensión de las instrucciones se utiliza la metáfora de halcón y paloma[4]. Por simplicidad se puede afirmar que existen dos formas de abordar la negociación, ser halcón o paloma. Se entiende que actuar como halcón es abordar las disputas de forma agresiva y egoísta, poniendo los intereses personales por encima de los del grupo. Por el contrario, actuar como paloma es buscar el bien común, procurando la mutua cooperación. Es obvio que las personas no se comportan siempre de la misma forma, en todas las situaciones, planteando la pregunta sobre las condiciones que llevan a favorecer la conducta egoísta, o lo que es lo mismo, qué factores pueden fomentar cooperación.

Como se ha explicado, la simulación del juego del conflicto $2 \times 2$ busca enseñar la interdependencia estratégica. Esto obliga a considerar el problema colectivo sistémico, donde existen interrelaciones entre las decisiones y el bienestar colectivo. Como situación problemática de contexto se les pide a los estudiantes que reflexionen sobre los conflictos que ven en su vida diaria. Claramente no todos son iguales, algunos son sencillos como las discusiones de pareja sobre la distribución de tareas. Otros son mucho más complejos, como los enfrentamientos entre naciones por controlar un territorio.

[4] Las instrucciones se pueden consultar en el blog del Grupo EMAR. https://racionalidadltda.wordpress.com/2018/01/23/conflicto-2x2-instrucciones/. A diferencia de un experimento económico, en un juego de negociación se contextualiza la situación problemática en la Guía del estudiante. https://racionalidadltda.wordpress.com/2018/01/25/conflicto-2x2-guia-para-el-estudiante/

### *3.2. Taller juegos de negociación*

A efectos de realizar la contrastación empírica del modelo, se llevó a cabo una sesión experimental con estudiantes de primer semestre inscritos en la Universidad Industrial de Santander, quienes sin distinción de carrera deben cursar y aprobar de manera obligatoria la Cátedra Vida y Cultura Universitaria. De acuerdo con el objetivo de la Cátedra, se diseñó el taller Juegos de Negociación como una estrategia pedagógica, construida alrededor de la metodología de dilemas morales, en donde se conjugan las competencias cognitivas como el multiperspetivismo, pensamiento sistémico y valoración de argumentos, en tres momentos diferentes: foros en Facebook, debate competitivo y simulación.

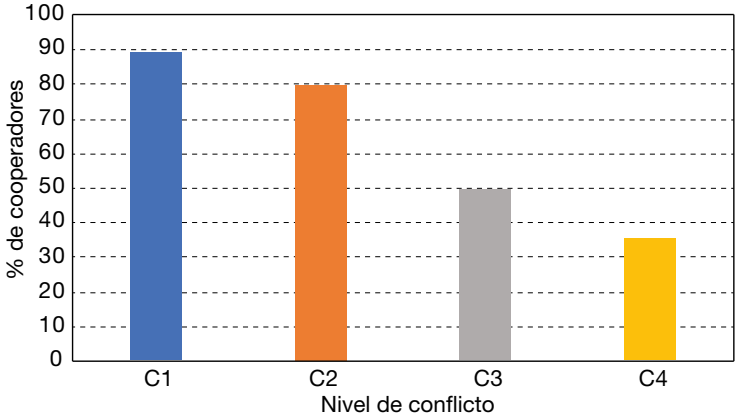
Se realizó una selección aleatoria de 100 estudiantes que fueron invitados por correo electrónico. Una vez recibidos los correos de aceptación se asignaron a tres grupos y horarios definidos, buscando balance por sexo y carrera en la que se encontraban matriculados. El jueves 11 de abril de 2019 se realizó el juego del conflicto en las instalaciones del Centro de Tecnologías de la Información y Comunicación (CENTIC) de la UIS. En total participaron 98 estudiantes distribuidos en tres horarios diferentes, 38 estudiantes a las 2:00 p.m., 28 estudiantes a las 4:00 p.m. y 32 estudiantes a las 6:00 p.m.

### *3.3. Resultados*

Al finalizar la sesión en los ordenadores se discute en voz alta alrededor de dos preguntas orientadoras: 1) ¿Los estudiantes están dispuestos a cooperar en función del nivel de conflicto al que se enfrentan? 2) ¿El hecho de jugar de forma secuencial aumenta la cooperación en comparación al juego simultáneo? En clase usualmente se usan los comentarios de los estudiantes para resaltar y aclarar los conceptos claves del juego con el propósito de estudiar colectivamente si la experiencia individual coincide o no con los resultados teóricos. Sin embargo, para tener una mejor compresión del fenómeno colectivo es necesario indagar con mayor detalle en los datos recolectados en el experimento.

En este sentido, la primera pregunta se puede responder al observar la proporción de estudiantes que deciden jugar cooperativamente como palomas en función del nivel de conflicto. Como se puede observar en el Gráfico 2, la cooperación es mayor en los casos donde existe menos conflicto, mostrando un orden lógico muy concreto. En particular, cuando los parámetros configuran un juego donde la estrategia paloma es dominante, el 90 % de los participantes cooperaron. Al aumentar levemente la diferencia en los parámetros, que para este caso se ha llamado juego de complementos estratégicos, el 80 % decidió cooperar. Cuando la configuración de parámetros representa el tradicional juego de sustitutos estratégicos, el 50 % de los participantes decidió cooperar. Finalmente, en el caso de mayor conflicto, donde actuar agresivamente es una estrategia dominante, solo el 35 % de los participantes cooperó.

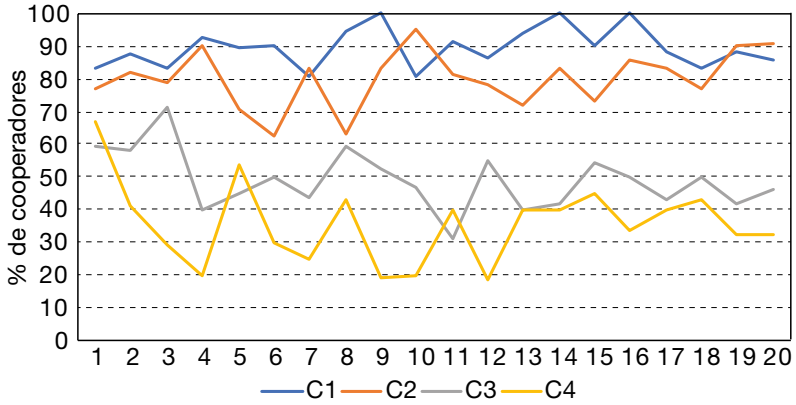
**GRÁFICO 2**
**PROPORCIÓN DE COOPERADORES POR NIVELES DE CONFLICTO**
**(En %)**




**FUENTE:** Elaboración propia.

Para responder la segunda pregunta es posible explorar las diferencias en la proporción de cooperadores en el juego simultáneo y secuencial. En el Gráfico 3 se puede observar que la proporción de palomas se mantiene y es muy estable a través de los periodos. No se ve un cambio de tendencia entre los periodos iniciales donde se juega el protocolo simultáneo y los periodos finales donde se juega secuencialmente. En otras palabras, lo que realmente importa al momento de tomar la decisión es el nivel de conflicto, y el hecho de jugar secuencialmente no hace una gran diferencia. Esto es algo que usualmente los estudiantes discuten en la clase, pues se entiende que el razonamiento estratégico es más complejo cuando se debe hacer inducción hacia atrás.

**GRÁFICO 3**
**PROPORCIÓN DE COOPERADORES POR NIVELES DE CONFLICTO Y PERIODO**
**(En %)**




**FUENTE:** Elaboración propia.

Para confirmar estos hallazgos se realiza una estimación de los efectos de las diferentes variables tratamiento sobre la decisión de cooperar, es decir, elegir la opción paloma. Por medio de un modelo panel logit se puede confirmar que las diferencias son significativas entre los cuatro niveles de conflicto. La estimación también permite calcular el efecto de los parámetros X e Y, así como del nivel de conflicto, medido como la diferencia en los pagos cuando un jugador decide paloma y el otro halcón. Como se observa en el Cuadro 5, en una primera especificación, los parámetros X e Y muestran efectos significativos en la dirección esperada. Se coopera menos cuando la tentación a actuar agresivamente es mayor y cuando las consecuencias de caer en la posición desventajosa son peores.

Sin embargo, resulta interesante analizar la diferencia entre estos parámetros como una medida del nivel de conflicto, independiente de la clasificación de cuatro juegos que hemos realizado. En este sentido, en las especificaciones (2) y (5) se

**CUADRO 5**
**MODELO LOGIT TIPO PANEL SOBRE LA DECISIÓN DE COOPERAR**

| **Prob (Paloma=1)** | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** |
|---|---|---|---|---|---|
| X | 0,0024*** (0,00) | | | –0,0001 (0,00) | |
| Y | –0,0045*** (0,00) | | | –0,0031*** (0,00) | |
| Y-X | | –0,0041*** (0,00) | | | –0,0024*** (0,00) |
| C2 | | | –0,9939*** (0,20) | –1,0768*** (0,29) | –0,4287 (0,22) |
| C3 | | | –2,5713*** (0,20) | –1,1364*** (0,27) | –1,4257*** (0,26) |
| C4 | | | –3,2891*** (0,21) | –1,8990*** (0,35) | –1,5406*** (0,33) |
| Periodo | –0,0304 (0,02) | –0,0333 (0,02) | –0,0213 (0,02) | –0,0241 (0,02) | –0,0248 (0,02) |
| Secuencial | 0,2808 (0,23) | 0,3210 (0,23) | 0,2206 (0,23) | 0,2323 (0,24) | 0,2282 (0,24) |
| Azul | 0,3205** (0,12) | 0,3199** (0,12) | 0,3330** (0,12) | 0,3320** (0,12) | 0,3335** (0,12) |
| Intercepto | 4,7279*** (0,32) | 3,9258*** (0,26) | 2,5238*** (0,24) | 4,9781*** (0,52) | 3,5171*** (0,29) |
| N | 1.960 | 1.960 | 1.960 | 1.960 | 1.960 |
| Chi2 | 349,27 | 336,20 | 345,12 | 362,86 | 361,11 |

NOTA: Errores estándar en paréntesis. Los asteriscos *, **, *** indican significancia al 10%, 5% y 1%, respectivamente.
**FUENTE:** Elaboración propia.

observa que el coeficiente asociado a Y-X es negativo y significativo, lo que confirma que existe un conflicto distributivo que explica el comportamiento agresivo, o menos cooperativo. En otras palabras, aunque siempre la mutua cooperación es un resultado eficiente, es el riesgo de caer en una posición desventajosa lo que lleva a jugar como halcón. Adicionalmente, los resultados del Cuadro 5 confirman que el cambio de protocolo no tiene un impacto significativo, pues las probabilidades de jugar paloma no cambian cuando pasan de un juego simultáneo a uno secuencial. La ganancia adicional, en términos pedagógicos, es generar la reflexión y discutir el razonamiento estratégico que realizan los estudiantes.

## 4. Conclusiones

El conflicto, la competencia y la cooperación son consecuencias de vivir en sociedad. Bajo esta óptica, estudiar el conflicto supone aceptar la idea de que la mayoría de interacciones son esencialmente situaciones de negociación. Aunque tomar decisiones es algo cotidiano, no es muy frecuente que estas se discutan constructivamente en las aulas de clase. Por el contrario, la valoración de las diferentes estrategias suele estar cargada de juicios de autoridad, sanciones o celebraciones. Usualmente en los cursos de economía los contenidos se imponen sobre lo que los estudiantes realmente quieren. Por el contrario, la confrontación permanente de las ideas entre iguales permite explorar cómo piensan otros alrededor de un problema y desentrañar el sentido que le dan a sus decisiones.

En este artículo se describen el protocolo y los hallazgos de la implementación del juego del conflicto 2x2, que conjuga elementos de los experimentos económicos con aspectos pedagógicos que permiten analizar y discutir cuándo y cómo se puede mejorar la coordinación en contextos de negociación. Por medio de la simulación se busca que el estudiante construya los conceptos a través de la experiencia en ambientes digitales controlados. Posteriormente se fomenta un debate sobre las decisiones económicas relacionadas con la temática propia de cada curso. La estrategia permite que los estudiantes tomen decisiones e interactúen entre ellos, viviendo conceptos antes de aprender su definición formal. Esto posibilita que cada estudiante pueda tomar una posición activa y crítica en su propio proceso de aprendizaje.

De acuerdo con Ruiz y Chaux (2005), manejar pacífica y constructivamente los conflictos requiere de conocimientos específicos, algunos de los cuales se han desarrollado en el marco de la teoría de juegos. Bajo este enfoque se entiende el conflicto como una serie de negociaciones, ofertas y contraofertas, que permitan caracterizar estrategias óptimas orientadas por el interés común de llegar a una solución que evite caer en el punto de desacuerdo. El carácter pedagógico esta enriquecido a partir de la teoría constructivista del aprendizaje (Sarnikar, 2015). Esta se basa en la idea de que el individuo no es un simple producto del ambiente, sino una construcción propia que se produce día a día como resultado de la interacción entre los aspectos cognitivos, sociales y afectivos.

En consecuencia, el propósito al realizar un experimento en el aula es suministrar apoyo al estudiante a través de actividades intencionales, planificadas y sistemáticas. Además, se busca garantizar el espacio para desarrollar habilidades cognitivas, emocionales y comunicativas como la autorregulación, la creatividad para generar opciones y soluciones ante una situación conflictiva y la capacidad de transmitir asertivamente intereses particulares. En este contexto, los resultados alcanzados en este estudio muestran que con planeación y armonización de los objetivos trazados en el curso de teoría de juegos se pueden diseñar experimentos pedagógicos como una alternativa ventajosa para el proceso de aprendizaje, complementando los libros de texto y las charlas magistrales tradicionales.

## Referencias bibliográficas


Alba-Fernández, V., Brañas-Garza, P., Jiménez-Jiménez, F., & Rodero-Cosano, J. (2006). Teaching Nash equilibrium and dominance: A classroom experiment on the beauty contest. *Journal of Economic Education*, *37*(3), 305–322. https://doi.org/10.3200/JECE.37.3.305-322

Andreoni, J., & Varian, H. (1999). Preplay contracting in the prisoners' dilemma. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(19), 10933-10938.

Aumann, R. (2006). Guerra y Paz. *Revista Asturiana de Economía*, *36*, 193–204. https://doi.org/10.2307/344005

Baliga, S., & Sjöström, T. (2004). Arms races and negotiations. *Review of Economic Studies*, *71*(2), 351-369. https://doi.org/10.1111/0034-6527.00287

Bornstein, G., Budescu, D., & Zamir, S. (1997). Cooperation in intergroup, N-person, and two-person games of chicken. *Journal of Conflict Resolution*, *41*(3), 384-406. https://doi.org/10.1177/0022002797041003003

Brandts, J., & Charness, G. (2000). Hot vs. cold: Sequential responses and preference stability in experimental games. *Experimental Economics*, *2*(3), 227-238. https://doi.org/10.1007/BF01669197

Brandts, J., & Charness, G. (2011). The strategy versus the direct-response method: a first survey of experimental comparisons. *Experimental Economics*, *14*(3), 375-398. https://doi.org/10.1007/s10683-011-9272-x

Brañas-Garza, P. (Ed.). (2011). *Economía experimental y del comportamiento*. Antoni Bosh Editor.

Chen, D., 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

Cooper, D., & Kagel, J. (2016). Other-Regarding preferences: A selective survey of experimental results. In J. Kagel & A. Roth (Eds.), *Handbook of Experimental Economics*. Princeton University Press.

Cooper, R., & DeJong, D. (1990). Selection criteria in coordination games: Some experimental results. *The American Economic Review*, *80*(1), 218-233.

Cooper, R., DeJong, D., Forsythe, R., & Ross, T. (1992). Communication in coordination games. *Quarterly Journal of Economics*, *107*(2), 739-771.

Dickie, M. (2006). Do classroom experiments increase learning in introductory microeconomics? *Journal of Economic Education*, *37*(3), 267-288. https://doi.org/10.3200/JECE.37.3.267-288

Dixit, A., & Nalebuff, B. (2010). *El arte de la estrategia*. Anthoni Bosh Editor.

Duffy, J., & Feltovich, N. (2002). Do actions speak louder than words? An experimental comparison of observation and cheap talk. *Games and Economic Behavior*, *39*(1), 1-27. https://doi.org/10.1006/game.2001.0892

Emerson, T. (2014). Anyone? Anyone? A guide to submissions on classroom experiments. *The Journal of Economic Education*, *45*(2), 174-179.

Emerson, T., & English, L. (2016). Classroom experiments: Teaching specific topics or promoting the economic way of thinking? *The Journal of Economic Education*, *47*(4), 288-299. https://doi.org/10.1080/00220485.2016.1213684

Emerson, T., & Hazlett, D. (2012). Classroom experiments. In *International Handbook on Teaching and Learning Economics* (pp. 90-98). Edgar Elgar Publishing.

Emerson, T., & Taylor, B. A. (2004). Comparing student achievement across experimental and lecture-oriented sections of a principles of microeconomics course. *Southern Economic Journal*, *70*(3), 672-693. https://doi.org/10.2307/4135338

Estévez, C. (2015). *Juegos en el salón de clase: Juego del conflicto* [Tesis doctoral]. Universidad Industrial de Santander.

Fehr, E., & Fischbacher, U. (2004). Social norms and human cooperation. *Trends in cognitive sciences*, *8*(4), 185-–90.

Fehr, E., & Gächter, S. (2000). Fairness and Retaliation: The economics of reciprocity. *Journal of Economic Perspectives*, *14*(3), 159-182.

Fehr, E., & Schmidt, K. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*(3), 817-868.

Fischbacher, U. (2007). Z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, *10*(2), 171-178.

García-Gallego, A., & Miller, L. (2011). Negociación. En P. Brañas-Garza (Ed.), *Economía experimental y del comportamiento* (pp. 143-153). Antoni Bosh Editor.

Holt, C. (1996). Classroom games: Trading in a pit market. *Journal of Economic Perspectives*, *10*(1), 193-203.

Murphy, J., & Cárdenas, J. (2004). An experiment on enforcement strategies for managing a local environment resource. *The Journal of Economic Education*, *35*(1), 47-61. https://doi.org/10.3200/JECE.35.1.47-61

Neugebauer, T., Poulsen, A., & Schram, A. (2008). Fairness and reciprocity in the hawk–dove game. *Journal of Economic Behavior & Organization*, *66*(2), 243-250. https://doi.org/10.1016/j.jebo.2005.12.008

Palacio, L., Cortés, A., & Muñoz-Herrera, M. (2015a). The bargaining power of commitment: An experiment of the effects of threats in the sequential hawk-dove game. *Rationality and Society*, *27*(3), 283-308.

Palacio, L., Cortés, A., & Muñoz-Herrera, M. (2015b). The strategic role of nonbinding communication. *Applied Mathematics*, 2015.

Palacio, L., Saravia, I., & Vesga, M. (2017). Juegos en el salón de clase: El mercado de los limones. *Revista de Economia Institucional*, *19*(36), 291-311. https://doi.org/10.18601/01245996.v19n36.11

Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review*, *80*(5), 1281-1302.

Rebelein, R., & Turkay, E. (2016). When do first-movers have an advantage? A Stackelberg classroom experiment. *The Journal of Economic Education*, *47*(3), 226-240. https://doi.org/10.1080/00220485.2016.1179144

Rousu, M., Corrigan, J., Harris, D., Hayter, J., Houser, S., Lafrancois, B., Onafowora, O., Colson, G., & Hoffer, A. (2015). Do monetary incentives matter in classroom experiments? Effects on course performance. *Journal of Economic Education*, *46*(4), 341-349. https://doi.org/10.1080/00220485.2015.1071214

Ruiz, A., & Chaux, E. (2005). *La formación de competencias ciudadanas* (p. 150). Ascofade.

Sarnikar, S. (2015). *What can behavioral economic teach us about teaching economics?* Palgrave Macmillan.

Schelling, T. (2006). Sesenta Años Asombrosos: El Legado de Hiroshima. *Revista Asturiana de Economía*, *35*, 183-196.

Stodder, J. (1998). Experimental moralities: Ethics in classroom experiments. *The Journal of Economic Education*, *29*(2), 127-138. https://doi.org/10.1080/00220489809597946

Van Huyck, J., Battalio, R., & Beil, R. (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *American Economic Review*, *80*(1), 234-248.

Zelmer, J. (2003). Linear public goods experiments: A meta-analysis. *Experimental Economics*, *6*, 299-310.

# *TRIBUNA DE ECONOMÍA*

*Esta sección ofrece al lector la oportunidad de conocer los análisis y puntos de vista de economistas y expertos de reconocido prestigio sobre temas de actualidad económica.*

CUADERNOS ECONÓMICOS ICE
INFORMACIÓN COMERCIAL ESPAÑOLA

# Impacto económico de la pandemia en Europa

# *Economic Impact of the Pandemic in Europe*

Emilio Ontiveros Baeza
Presidente de Afi
Catedrático Emérito de Economía de la Empresa
Universidad Autónoma de Madrid

***Resumen***

*El objetivo del trabajo es analizar los rasgos diferenciales de la crisis económica originada por la pandemia derivada del Covid19, con especial atención a su particularización en la Unión Europea y en la economía española. Especial atención se presta a las decisiones de política económica nacionales y a las de carácter paneuropeo por las instituciones comunitarias. Se concluye en la adecuación de las decisiones del Banco Central Europeo (BCE), pero en la necesidad de complementar las adoptadas por la Comisión con un verdadero plan de reconstrucción que además de asentar la recuperación de la recesión afiance la necesaria modernización de las economías de la Unión Europea (UE) en la transición energética y en la inserción en la economía digital.*

***Palabras clave****: crisis económica, Unión Europea, políticas fiscales.*
***Clasificación JEL****: E61.*

***Abstract***

*The objective of the work is to analyze the differential features of the economic crisis caused by the pandemic of Covid19, with special attention to its particularization in the European Union and in the Spanish economy. Special attention is paid to national economic policy decisions and those of a European nature by Community institutions. It concludes in the adaptation of those adopted by the European Central Bank (ECB), but in the need to complement those adopted by the Commission with a true reconstruction plan that, in addition to consolidating the recovery from the recession, strengthens the necessary modernization of the European Union (UE) economies in the energy transition and integration into the digital economy.*

***Keywords:*** *economic crisis, European Union, fiscal policies.*
***JEL Classification:*** *E61.*

## 1. Singularidad. Difíciles analogías con otras crisis

No todo está echado a perder. Vivimos una época de gran enriquecimiento analítico. Los economistas, en mayor medida los académicos, no han dejado en las dos últimas décadas de acumular evidencias, no siempre conformadoras de sus presunciones, acerca de las crisis económicas y financieras, de sus determinantes y mecanismos de propagación. La historia económica ha enriquecido la diversidad de sus registros sobre las crisis económicas, sobre su determinación y alcance, mucho más en estas dos últimas décadas que en todo el siglo pasado.


DOI: https://doi.org/10.32796/cice.2020.99.7028
Recibido: Mayo 2020 · Aceptado: Mayo 2020

Apenas habíamos superado completamente las secuelas de lo que considerábamos la crisis más compleja y la más severa desde la Gran Depresión, la desencadenada en 2008 en el sistema financiero estadounidense[1], cuando la derivada de la epidemia del Covid-19, reclama con más autoridad un registro excepcional.

Desde luego por el hecho de no provenir de ningún desequilibrio económico o financiero específico como las anteriores, pero también por haber propiciado la más rápida propagación y concentrar en menos tiempo la mayor pérdida de bienestar y de riqueza financiera. Precisamente esa ausencia de precedentes suficientemente significativos ha debido propiciar la búsqueda de analogías con fenómenos distintos, como guerras, catástrofes naturales e incluso con invasiones extraterrestres, como sugería Kenett Rogoff (2020). Lawrence Summers (2020) también diferencia la importancia que recibirá en los libros de historia este shock global frente a los originados por los atentados de septiembre del 2001 en Nueva York o la quiebra de Lehman Brothers. Ningún episodio desde la guerra civil americana, considera el profesor de Harvard, ha cambiado tan dramáticamente la vida de tantas familias. Los estudiantes del futuro, estudiarán su impacto con la misma relevancia que el que tuvo el asesinato del Archiduque en 1914, el colapso del mercado de acciones de 1929 o la Conferencia de Múnich de 1938. Más allá de la significación de los mismos, su relevancia histórica es por lo que siguió, por los cambios que indujo, constituyéndose en «una bisagra en la historia». Pero con independencia de esas solemnes presunciones, es muy probable que esta crisis tenga la capacidad de acelerar transformaciones estructurales en algunos sectores productivos, desde las manufacturas ya afectadas en los últimos años, hasta los servicios.

Esta crisis reúne más atributos que la singularizan, pero los iremos viendo en el contexto de este trabajo cuyo objeto, además de esa caracterización diferencial, es destacar esa otra dimensión singular como las decisiones de política económica adoptadas y su particularización en Europa y España. Las más conjeturables circunstancias de la recuperación y de la reconstrucción cuando este trabajo se concluye y, muy especialmente, la más aventurada presunción por intuir la conformación del paisaje después de la batalla, completarán el recorrido que propongo al lector, tratando de cumplir con la invitación de los editores de la revista. Difícil, porque las limitaciones que el análisis económico pueda presentar en circunstancias de una permeabilidad geográfica tan acusada, se erigen sobre una dificultad más compleja de abordar como es el conocimiento de la propia evolución del virus, de las eventuales mutaciones y mecanismos de propagación de la epidemia. Y esta es la que determina esa tesitura entre vidas y economía, entre salud y capacidad de subsistencia, en definitiva.

[1] En Tooze (2019) y en Ontiveros (2019) se analizan los determinantes y rasgos que singularizaron la crisis desencadenada en 2008

### *1.1. Shocks de oferta y de demanda*

La manifestación económica más explícita fue la contracción en la oferta derivada del aislamiento al que se sometió a la población, consecuente con las pretensiones por neutralizar la propagación del virus. Tuvo lugar en algunas de las provincias chinas más activas como partícipes en las cadenas de producción transfronterizas. El cierre de centros de producción afectó de forma inmediata al suministro de bienes intermedios en numerosos sectores al que siguió un descenso en la demanda, el transporte, etc. La extensión a las economías avanzadas, desde luego a Europa y algo más tarde a EE.UU., fue igualmente rápida. En Europa casi coincidieron las primeras señales de extensión del virus con el descenso de la actividad económica.

El, desde años antes, debilitado crecimiento del comercio internacional sufría igualmente una contracción significativa, en gran medida determinada por esas disrupciones en las cadenas de producción globales y el asociado desplome en la demanda global. En la actual fase de globalización nunca habíamos observado una contracción en todas las formas de transporte y en la actividad aeroportuaria como el que siguió de forma inmediata a la decisión de confinamiento chino.

Los mercados financieros, en principio de forma más exagerada los de acciones, se encargaron de escenificar la singularidad de la crisis que se nos venía encima: las pérdidas de riqueza financiera marcaron máximos en apenas un par de semanas. La liquidez empezó a escasear para algunos demandantes, en especial para los que la precisaban para atender el servicio de un endeudamiento público y privado en máximos históricos. La intensificación de la aversión al riesgo también extendió las correcciones a los mercados de bonos, al oro y a otras materias primas, el petróleo incluido.

A diferencia de otras crisis, estos tres shocks son autogenerados, estrechamente entrelazados, provenientes de las propias decisiones tendentes a neutralizar la extensión del virus, en particular la inmovilización de los centros de producción.

### *1.2. Severidad de la recesión*

Esa triple configuración –perturbaciones en la oferta, desplomes en la demanda y destrozos de riqueza financiera–, concedía capacidad suficiente a esa crisis para determinar una recesión en toda regla. Los iniciales registros de empleo y los de crecimiento de las economías correspondientes al primer trimestre confirmaron las peores estimaciones. En todas las economías, la contracción del Producto Interior Bruto (PIB) en esos tres primeros meses disponía de pocos precedentes, aun cuando en países como la mayoría de los de la eurozona, el impacto más severo era solo en las últimas dos semanas de marzo. Suficiente para asentar con algo más de credibilidad unas previsiones para el conjunto del año que convertían en optimistas las suministradas a mediados de abril por el Fondo Monetario Internacional (FMI) o las avanzadas semanas después por la propia Comisión Europea. La recesión del

conjunto de la economía mundial en todo el año acabará siendo en gran medida la consecuencia más directa de la muy pronunciada contracción del PIB que tendrá lugar en el segundo trimestre del año.

El ascenso del desempleo ha sido más intenso y rápido en esta crisis que en la anterior. La Organización Internacional del Trabajo (OIT[2], 2020) anticipaba a finales de mayo una reducción de las horas de trabajo en todo el mundo en el segundo trimestre del año equivalentes a 305 millones de puestos de trabajo a tiempo completo. El impacto en los jóvenes es particularmente severo. Un total de 178 millones de trabajadores jóvenes, más del 40 % de los empleados globalmente estaban trabajando en sectores vulnerables cuando se desencadenó la crisis. En ese mismo informe, la OIT señala que el 77 % de los 388 millones de trabajadores jóvenes tenían trabajos en la economía sumergida o informal, frente al 66 % de los trabajadores adultos, aquellos con edades superiores a 25 años.

Ese deterioro era suficientemente explícito en todas las economías avanzadas, merecedores en los cuatro primeros meses del año de registro histórico. Desde luego en EE.UU., donde las intensas elevaciones del desempleo se comparaban con las ocurridas tras la desmovilización de las tropas y el descenso en la producción militar consecuente con el final de la Segunda Guerra Mundial. Apenas concluido abril, las estimaciones de contracción del PIB para el conjunto del año no bajaban del 5 % y la proyección que hacía la Oficina Presupuestaria del Congreso de la tasa de desempleo se situaba en el 16 %. La contracción de la actividad industrial en abril, del 11,2 %, era la mayor mensual de los registros estadísticos. En ese mismo mes de abril, perdía 20,5 millones de empleos y la tasa de desempleo ascendía al 14,7 %, los peores registros desde la Gran Depresión. A principios de mayo, el número de trabajos perdidos en EE.UU. durante la pandemia era equivalente a la población activa de 25 Estados.

En Europa la situación no era precisamente más tranquilizadora. Un trabajo difundido por McKinsey, del que son autores Chinn, Klier, Stern y Tesfu (2020), estima en 59 millones lo puestos de trabajo que se encuentran en situación de riesgo, un 26 % del empleo en toda la UE, Reino Unido incluido. Estar «en riesgo» significa, para esos autores, reducción de las horas de trabajo, o en la remuneración, permisos temporales o despidos permanentes. La experiencia de la pasada crisis puede volver a ser ilustrativa de algunas consecuencias sociales y de las dificultades para restaurar los daños en los segmentos de población con menor capacidad defensiva. Entonces, para recuperar los niveles de empleo previos a la crisis fueron necesarios ocho años.

Con todo, aunque sea pronto para cuantificar las contracciones en el PIB que tendrán lugar en 2020 y, por tanto, hacer lo propio en términos per cápita, no es difícil anticipar un impacto muy importante en la distribución de la renta y en sus consecuencias sociales. A finales de 2019, el FMI había previsto tasas positivas de variación en el PIB por habitante en más de 160 de sus países miembros para 2020, pero a principios de abril, ya anticipaba que serían más de 170 los que sufrirían un decrecimiento en ese indicador de bienestar en el año en curso, un registro solamente

[2] ILO por sus siglas en inglés.

superado por el generado por la Gran Depresión. A esas consecuencias de fenómenos susceptibles de ser estadísticamente captados, habría que añadir el impacto asociado a la caída de la actividad informal, de la economía sumergida, compensatoria en crisis anteriores de la caída del empleo, que dejará sin beneficios sociales a cantidades significativas de personas, incluso en las economías avanzadas.

Tendremos que esperar varios meses para particularizar el balance de daños por grupos de países con un mínimo de fiabilidad. Aun cuando el conjunto convencionalmente agrupado como economías emergentes contrajeran menos su ritmo de crecimiento del PIB, algunas de ellas sufrían de forma considerable. Las salidas de flujos de capital daban también una idea de esa estigmatización en la que caían algunos países menos avanzados. Las salidas de flujos de cartera de economías emergentes, estimadas por el FMI en más de 100.000 millones de dólares en febrero y marzo, más que triplicaban las registradas en el mismo periodo de la crisis financiera global de 2008. Los productores de materias primas o los que son receptores de remesas de emigrantes, también estaban abocados a descensos de la tasa de crecimiento económico significativamente superiores a las previsiones iniciales. Esa diferencia en el impacto de la recesión entre economías avanzadas y emergentes, otro rasgo común con lo ocurrido en la crisis de 2008, se veía velada por el mejor comportamiento relativo de las dos grandes economías de este último grupo, China e India. Ambas no registrarían en el conjunto de 2020 una recesión; la primera recuperaría rápidamente su ritmo de crecimiento para, según el FMI, superar el 1,5 % de crecimiento anual, e India podría rozar el 1 %. Las otras, América Latina, en concreto, sufrirán de forma particular esos factores antes comentados: caída en los precios de las materias primas, flujos de salida de capitales, importancia de la deuda denominada en dólares, entre otros.

Al término de mayo de 2020, cuando se revisaban estas notas, la generalidad de las previsiones no anticipaban, precisamente, una intensa recuperación para el conjunto de la economía mundial. En el caso de la estadounidense, el presidente de la Reserva Federal emplazaba a finales de mayo la completa recuperación no antes de final de 2021, y siempre condicionada al descubrimiento de una vacuna contra el virus, condición necesaria para que empresas y familias recuperen la confianza[3]. Adicionalmente, eran numerosos los analistas que más allá de las precisiones sobre la caligrafía de la recuperación (descartada esa «V» completamente simétrica), estimaban que la crisis ya estaba generando daños estructurales en algunas economías que limitarían seriamente el crecimiento tendencial. Era el caso, desde luego, de algunas economías europeas.

[3] «Jay Powell warns US recovery could take until end of 2021», *Financial Times*, 18 de mayo 2020. https://www.ft.com/content/2ed602f1-ed11-4221-8d0b-ef85018c96ea

## 2. Europa dañada

Sin menoscabar la singularidad de la crisis, otro de los denominadores comunes con la de 2008 que cabría identificar es la facilidad con la que Europa ha sido contagiada. Pocos días después de la emergencia en el epicentro chino ya se detectaban casos de extensión del virus en Europa. Italia, España y Reino Unido fueron los países más rápida e intensamente contagiados. Las decisiones de confinamiento fueron estrictas, especialmente en los dos primeros, y con ellas la paralización de la actividad económica y la subsiguiente recesión. Dependiendo de las economías, un mes de confinamiento puede llegar a contraer el crecimiento anual del PIB hasta 3 puntos porcentuales; el Secretario de la Organización para la Cooperación y el Desarrollo Económicos (OCDE), en una declaración ante el G20 el 5 de abril, lo cifraba en 2 puntos porcentuales, apuntando al sector turístico como el que sufriría un mayor desplome. En el caso de Europa, la contribución al PIB de ese sector es del 10 %, cinco puntos porcentuales menos que en España.

El conjunto de la UE, después de que en 2019 dos de sus principales economías, Italia y Alemania, estuvieran durante el segundo semestre en recesión, encajaba el impacto de la pandemia con previsiones de recesión generalizada, más intensa que el promedio de las economías avanzadas. La actividad industrial y un sector tan relevante como el del automóvil sufrían las consecuencias de las tensiones comerciales y alteraciones regulatorias, con especial impacto en economías como la alemana y la española, y añadía en marzo las señales de dislocación de las cadenas de producción transfronterizas. El turismo, también de importancia significativa en varias economías de la eurozona, acusaba más que ningún otro sector esas limitaciones a la movilidad física. La Comisión Europea (2020), en su informe económico de primavera, anticipaba una contracción del PIB para el conjunto de 2020 del 7,5 % para crecer al 6 % en 2021. No menos preocupante es el hecho de que sea la inversión empresarial el componente del PIB que seguirá registrando un peor comportamiento, con impacto nada favorable sobre el crecimiento potencial de las economías.

Como ya ocurriera en la crisis de 2008, aunque por razones bien distintas, el impacto de la crisis en la región no será en modo alguno homogéneo. Aun cuando la pandemia es un «shock simétrico» sus impactos son bien distintos entre los países miembros. La propia Comisión Europea (2010) destaca la ampliación de las divergencias económicas en el seno de la UE, en gran medida, debido a la severidad de la infección, a las decisiones de reclusión adoptadas o a la importancia relativa de sectores tan expuestos como el turismo o, no menos importante, al margen de maniobra de la política fiscal para intentar neutralizar el ascenso de la mortalidad empresarial y del desempleo.

En el seno de la eurozona serán algunas de las economías que en mayor medida sufrieron las consecuencias de la crisis de 2008 las que lo hagan con esta, Italia y España de forma destacada. Con la información hoy disponible también puede anticiparse que la pendiente de la recuperación de esas economías será menos pronunciada de lo que está siendo la contracción. En realidad, se estima que transcurran más de dos años para recuperar el PIB perdido en estas economías.

Con datos de Eurostat, en el primer trimestre la contracción ha sido del 3,6 %, pero en el segundo podría llegar al 15 %. Por eso, del rango de contracción del PIB agregado que se manejaba a principios de mayo, entre el 5 y el 12 %, nadie asume que será el punto más bajo. Su cuantía final dependerá en todo caso de la evolución de la pandemia, de la detección de vacunas y medicamentos para su tratamiento y, no menos importante de las medidas compensadoras de la recesión.

## 3. Nuevamente España

Al igual que ocurriera en la crisis de 2008, la economía española, junto al resto de las consideradas periféricas, ha vuelto a ser rápida e intensamente contagiada, hasta el punto de ser una de las economías cuya recesión será más pronunciada en 2020, según todas las previsiones. A diferencia de lo ocurrido en aquella crisis, no son desequilibrios económicos o financieros los que han propiciado esa más severa particularización de la crisis. Algunas de las principales razones de esa facilidad de contagio han ido apareciendo en los comentarios anteriores, pero puede ser útil tratar de sistematizarlas en las cinco siguientes[4].

**a) Severidad del confinamiento**. Como hemos comentado más arriba, la vinculación entre duración de la reclusión e intensidad de la contracción del PIB es muy estrecha: un mes de reclusión puede contraer el crecimiento del PIB anual entre un 2 % y 3 %, dependiendo de la conformación sectorial de las economías. En España, donde el contagio arraigó pronto y de forma intensa, el confinamiento está siendo de los más estrictos y duraderos, con impacto significativo en los sectores que más contribuyen a la determinación del PIB.

**b) Estructura productiva de la economía española**. Todas las actividades económicas han quedado afectadas por la reducción de la movilidad de las personas, pero la vuelta a la normalidad será más lenta en aquellas vinculadas al sector turístico, hostelería y, desde luego, al transporte aéreo.

Este no solo será el año en que se trunque esa larga racha de récords de visitantes e ingresos por turismo extranjero. Concluyen siete años de crecimiento ininterrumpido en ambas variables, que han consolidado la posición de España como una de las grandes potencias turísticas globales. Será también el año en el que probablemente habrá una reconversión sin precedentes en ese sector. Algunas asociaciones sectoriales y empresarios son desde hace tiempo conscientes de la conveniencia de diversificar y mejorar la calidad de la oferta; de reducir algunos elementos de vulnerabilidad asociados a la elevada especialización en ese binomio sol y playa que expone el sector nacional a la competencia de otros destinos con costes más reducidos. Son más de 20.000 empresas turísticas, según el Instituto Nacional de Estadística (INE),

[4] En Ontiveros (2020) se comentaron sucintamente esos factores

las que operan en nuestro país, no todas con la capacidad defensiva suficiente para resistir caídas en los ingresos como las que ya se anticipan para el conjunto del año.

Otros sectores importantes en la industria española, el del automóvil entre ellos, añaden a sus problemas previos a esta crisis los derivados de las perturbaciones en las cadenas de producción transfronterizas. Por eso, tan importante como recuperar la demanda, es hacerlo con el atractivo localizador de inversión extranjera directa en sectores industriales españoles intensivos en empleo y en potencial exportador.

**c) Censo empresarial**. En gran medida consecuente con la estructura sectorial de nuestra economía, el protagonismo de las pequeñas empresas en la determinación del PIB es muy importante. Análisis como el antes comentado de Chinn, Klier, Stern y Tesfu (2020) revelan esa mayor vulnerabilidad de algunas economías europeas, la nuestra entre ellas, destacando que al menos dos de cada tres puestos de trabajo en Europa están en pequeñas y medianas empresas, pero más del 30 % están en microempresas, aquellas con menos de 9 trabajadores. Esa proporción es significativamente más elevada en la economía española, donde los sectores de hostelería, comercio al por menor, o actividad inmobiliaria, señalados como muy expuestos, es también mayor.

De los 3,6 millones de empresas que aporta el último censo empresarial del INE, un 81 % opera en el sector servicios, de los que el comercio representa un 23 % del total; la industria apenas alcanza el 6 %, y la construcción el 13 %. Esa elevada concentración sectorial ayuda a explicar la dimensión media reducida de las empresas: solo el 4 % de todas las empresas españolas tienen 10 o más trabajadores; el 90 % no tienen más de cuatro. La relación también es estrecha entre tamaño y capacidad defensiva frente a choques como el que ahora sufrimos: a pesar de los apoyos públicos, la tasa de mortalidad empresarial está siendo en estos meses muy elevada en nuestro país. Lo es entre empresas de pequeña dimensión, y con ella el aumento del desempleo de trabajadores en ese tipo de empresas, por lo general con contratos temporales y menores grados de cualificación que los empleados en empresas de mayor tamaño. Es decir, en mayor medida abocados al desempleo de larga duración.

**d) Mercado de trabajo**. La estructura del mercado de trabajo español tampoco es ajena a la distribución sectorial de la economía española y a su correspondiente censo empresarial. Su singularidad y disfunciones conocidas, es por sí solo uno de los factores de vulnerabilidad más destacados, pero también uno de los elementos que han acentuado las consecuencias sociales adversas de esta crisis en nuestro país. Recordemos que todavía no se han alcanzado los niveles de empleo previos a la crisis del 2008. La dualidad tradicional, el predominio cada vez más explícito de contratos temporales, justifica esa rápida reacción de exclusión de ese tipo de contratos cuando emergen las primeras señales adversas sobre las empresas. El aumento del desempleo por el dominio de esos contratados temporales ha coexistido con el de esa figura subsidiada de los Expedientes de Regulación Temporal de Empleo (ERTE), cuya importancia ha impedido que las cifras de desempleo alcanzaran niveles

todavía más alarmantes. A inicios de mayo más de medio millón de empresas habían solicitado un ERTE, afectando a más de 4 millones de trabajadores. Es altamente probable que el horizonte al término de la vigencia de los mismos sea, en muchos casos, su mutación hacia los EREs tradicionales, despidos colectivos, en definitiva.

La asunción de ese panorama explica la intensa contracción del consumo, superior a la de las rentas, derivada de ese descenso de la confianza y el correspondiente aumento de la propensión al ahorro por razones precautorias. Un comportamiento tal condiciona, a su vez, las posibilidades de recuperación de forma significativa, como ya tenemos ocasión de observar.

**e) Margen de maniobra presupuestario**. La llegada de la pandemia coincidió con la revisión al alza del déficit público español, hasta el 2,8 % del PIB, por encima del compromiso de las autoridades españolas con la Comisión Europea. La deuda pública, del 95,5 % del PIB, también revelaba una mayor exposición al desplome de los ingresos tributarios y una menor capacidad para utilizar el gasto público para neutralizar la severidad de la recesión. En consecuencia, el Gobierno español era, al menos hasta mediados de mayo, uno de los que menor proporción del PIB había asignado a paliar las consecuencias de la crisis, como se comenta en el siguiente apartado.

Frente a esos rasgos vulnerables, apenas podrán actuar como paliativos algunas de las ventajas que nuestra economía acentuó en los últimos años de recuperación de la anterior crisis, como es la muy significativa reducción del endeudamiento privado, el aumento de la propensión exportadora o la mejora de la solvencia del sistema bancario. La deuda de empresas y hogares se redujo hasta el 129,7 % del PIB en el cuarto trimestre de 2019, situándose las ratios respectivas en el 72,8 % y 56,9 % del PIB. En realidad, es altamente probable que ese saneamiento financiero haya incluso mejorado durante los dos primeros meses de la pandemia, aunque está lejos de trasladarse a decisiones de gasto e inversión por los agentes económicos. La falta de confianza, el ascenso del ahorro precautorio, explica el descenso del consumo de las familias y la caída igualmente importante de la inversión empresarial con la que los agentes privados han reaccionado al deterioro de las expectativas económicas.

El horizonte inmediato no es muy favorable. Las empresas renovarán sus ventajas en costes, pero se toparán con la debilidad de la demanda de nuestros socios comerciales, mayoritariamente europeos, y con un entorno internacional en el que se acentuará el proteccionismo y la desconfianza de las empresas multinacionales en la dispersión geográfica de las cadenas de producción. Los efectos de la competitividad del sector primario y la demanda alcista de sus transformados actualmente solo podrán contribuir parcialmente a la recuperación, dada la escasa participación del sector en la determinación del PIB y del empleo.

En definitiva, no es de extrañar que todas las previsiones anticipen contracciones pronunciadas en el crecimiento español y ascensos en el desempleo, como se aprecian en el Cuadro 1, sin que se pueda asumir una recuperación tan rápida e intensa como la que está teniendo el desplome.

**CUADRO 1**
**PREVISIONES DE CRECIMIENTO DEL PIB**

| % anual | Afi | | | FMI (abril 2020) | | Comisión Europea | |
|---|---|---|---|---|---|---|---|
| | 2019 | 2020 | 2021 | 2020 | 2021 | 2020 | 2021 |
| **Mundo** | **3,0** | **–3,2** | **6,0** | **–3,0** | **5,8** | **–3,5** | **5,2** |
| **Desarrrollada** | **1,7** | **–6,1** | **5,2** | **–6,1** | **4,5** | | |
| Estados Unidos | 2,3 | –5,8 | 6,3 | –5,9 | 4,7 | –6,5 | 4,9 |
| Área Euro | 1,2 | –7,3 | 4,9 | –7,5 | 4,7 | –7,7 | 6,3 |
| Alemania | 0,6 | –5,8 | 4,4 | –7,0 | 5,2 | –6,5 | 5,9 |
| Francia | 1,3 | –7,4 | 5,3 | –7,2 | 4,5 | –8,2 | 7,4 |
| Italia | 0,3 | –9,9 | 6,1 | –9,1 | 4,8 | –9,5 | 6,5 |
| España | 2,0 | –8,9 | 6,1 | –8,0 | 4,3 | –9,4 | 7,0 |
| Japón | 0,7 | –5,5 | 3,7 | –5,2 | 3,0 | –5,0 | 2,7 |
| Reino Unido | 1,4 | –7,1 | 5,2 | –6,5 | 4,0 | –8,3 | 6,0 |
| **Emergentes** | **3,9** | **–1,3** | **6,5** | **–1,0** | **6,6** | | |
| Brasil | 1,1 | –5,1 | 3,3 | –5,3 | 2,9 | | |
| México | –0,1 | –7,0 | 2,8 | –6,6 | 3,0 | | |
| Rusia | 1,3 | –5,4 | 2,9 | –5,5 | 3,5 | | |
| India | 5,3 | 0,6 | 7,8 | 1,9 | 7,4 | | |
| China | 6,1 | 1,1 | 9,0 | 1,2 | 9,2 | 1,0 | 7,8 |

**FUENTE:** Afi, FMI y Comisión Europea.

## 4. Desigual intensidad de las políticas nacionales

Frente a esas amenazas depresivas de la actividad económica, las reacciones de los Gobiernos europeos han sido, en esta ocasión, relativamente rápidas y contundentes. Desde luego, si las comparamos con las tardías e inadecuadas repuestas que se adoptaron en la gestión de la crisis de 2008. Las decisiones adoptadas han procurado sustancialmente los mismos propósitos: la atención prioritaria a la gestión sanitaria, el apoyo bajo distintas vías a la liquidez de las empresas y la compensación por las pérdidas de horas de trabajo o directamente de empleos. En algunos casos, de forma más o menos simbólica la mayoría de los Gobiernos han aumentado el gasto en investigación para la detección de vacunas o medicamentos con los que combatir la infección y, como el español, han reforzado los controles y autorizaciones previas sobre la inversión extranjera directa en las más relevantes empresas españolas.

La diferencia fundamental entre los países ha radicado en la cantidad de recursos públicos asignados a esos propósitos. En general, las economías con mejor posición fiscal de partida están estimulando mucho más, incluso utilizando de forma más amplia y discrecional no solo la flexibilización de los límites de déficit y deuda pública

definidos en el Plan de Estabilidad y Crecimiento, sino que también lo han hecho con la suspensión de las normas sobre las ayudas de Estado. Es, por tanto, probable que la recuperación de aquellas economías de la eurozona con mayor margen de maniobra fiscal se asiente en mucha mayor medida en ayudas públicas a las empresas que fortalezcan no siempre de forma limpia sus ventajas competitivas. Las economías periféricas, que han sufrido un mayor impacto del contagio y disponen de menor margen de maniobra fiscal, se han visto en cierta medida obligadas a priorizar en la administración de sus recursos públicos las ayudas a las personas que a las empresas. Y eso es probable que acabe traduciéndose en desiguales ventajas competitivas en el seno del propio mercado interior.

En el Cuadro 2, elaborado por Bruegel (2020), se reflejan los principales grupos de medidas adoptadas por los Gobiernos europeos. Los cálculos de las ratios de las medidas están expresados sobre los PIB de 2019. La categoría de liquidez y garantías incluye medidas iniciadas por los Gobiernos (excluyen las decisiones de los bancos centrales), ya dispuestas por el sector privado; no los compromisos para apoyar con liquidez o con garantías o avales.

En el primer grupo se reflejan medidas de gasto adicional, fundamentalmente destinado a la atención sanitaria, a mantener el empleo, a subsidiar pequeñas y medianas empresas (pymes) o a incrementar la inversión pública. También están ingresos dejados voluntariamente de percibir, ya sea por cancelación de determinados impuestos o de contribuciones a la seguridad social. En definitiva, decisiones que generan aumentos de los déficits públicos sin ninguna compensación posterior. Ello, en contraste con los ingresos diferidos, aplazados, también fundamentalmente de algunos impuestos y contribuciones a la seguridad social que procuran frenar el deterioro de la liquidez, fundamentalmente de las empresas, pero no cancelar las obligaciones de pago. Esos desembolsos deteriorarán los déficits en 2020, pero si la gestión es adecuada, se recuperará en el futuro mediante la generación de mayor recaudación. El último de los conceptos son apoyos a la financiación de las empresas, ya sea mediante liquidez inmediata o líneas de crédito y suministros de avales a través de instituciones crediticias públicas, como el Instituto de Crédito Oficial (ICO) español.

Aun con las reservas que supone la fecha del 16 de abril, como referencia de esas asignaciones, se deduce fácilmente el mayor esfuerzo fiscal de aquellas economías con mayor margen de maniobra en sus finanzas públicas antes de la crisis. España no está entre ellas. Alemania, en el extremo opuesto está facilitando liquidez a pequeñas y medianas empresas, a través del sistema bancario con trámites muy simplificados y con el 100 % de garantía. De hecho, más de la mitad de todos esos apoyos a las empresas en la UE han tenido lugar en Alemania.

El denominador común de las decisiones de los Gobiernos nacionales ha sido tratar de ayudar a las empresas. Como se destaca por los investigadores de Bruegel, Anderson, Tagliapietra y Wolff (2020), «la aproximación básica ha sido proporcionar apoyo generoso e indiscriminado de emergencia para ayudar a las empresas a satisfacer sus necesidades inmediatas de liquidez». Pero en la medida en que la recesión permanece y se profundiza, es cierto que es necesaria una estrategia de más alcance.

Con todo, esa utilización mucho más moderada de la política fiscal no va a impedir que el aumento del gasto público derivado de esos estímulos y el desplome de los ingresos tributarios sitúen el déficit y la deuda pública españoles muy por encima del promedio europeo. Con dicha estructura de las finanzas públicas, la vulnerabilidad a episodios adversos en los mercados financieros es significativa. Es por eso que la economía española, al igual que las otras periféricas, depende de las políticas comunitarias para asentar su recuperación y la consiguiente reconstrucción.

**CUADRO 2**
**MEDIDAS FISCALES DISCRECIONALES EN RESPUESTA AL CORONAVIRUS AL 16 DE ABRIL DE 2020**
**(En % del PIB)**

| | Impulso fiscal inmediato | Diferidos | Liquidez y garantías |
|---|---|---|---|
| Bélgica | 0,70 | 3,00 | 10,90 |
| Dinamarca | 2,10 | 7,20 | 2,90 |
| Francia | 2,40 | 9,40 | 14,00 |
| Alemania | 6,90 | 14,60 | 38,60 |
| Grecia | 1,10 | 2,00 | 0,50 |
| Hungría | 0,40 | 8,30 | 0,00 |
| Italia | 0,90 | 13,20 | 29,80 |
| Holanda | 1,60 | 3,20 | 0,60 |
| España | 1,10 | 1,50 | 9,10 |
| Reino Unido | 4,50 | 1,40 | 14,90 |
| Estados Unidos | 9,10 | 2,60 | 2,60 |

**FUENTE:** Bruegel (2020).

## 5. La necesidad de las medidas paneuropeas

Esa desigual capacidad de las políticas fiscales nacionales es una de las razones que apoya la adopción de medidas conjuntas por las instituciones comunitarias para garantizar una recuperación europea suficientemente sólida, susceptible de compensar el empleo perdido; capaz igualmente de señalizar a las empresas para asentar las decisiones de inversión privadas necesarias para la reconstrucción. Está en lo cierto la propia Comisión Europea (2020) cuando advierte que, de no asumirse esas decisiones, «la crisis puede conducir a severas distorsiones en el seno del mercado único y arraigar divergencias económicas, financieras, y sociales entre los Estados miembros de la eurozona que podrían amenazar la estabilidad de la Unión Monetaria», y, se podría añadir, la propia legitimación de su existencia ante sus ciudadanos. Las reacciones comunitarias a la crisis han sido, cuando se concluyen estas notas,

bien distintas entre las dos instituciones con capacidad decisional, el Banco Central Europeo y la Comisión.

### *5.1. Banco Central Europeo (BCE)*

De las instituciones europeas, la que sin duda ha definido un mayor compromiso con la neutralización de la recesión y con la evitación de la fragmentación financiera en el seno de la eurozona ha sido el BCE. A partir del 12 de marzo reforzó las actuaciones que adoptó en la gestión de la crisis de 2008 y decidió otras específicas tendentes a asegurar los flujos crediticios a las empresas y evitar que restricciones de liquidez precipitaran situaciones de insolvencia. Se comenta someramente a continuación los rasgos fundamentales de cada una de ellas.

- **Ampliación de las operaciones de refinanciación a largo plazo (LTROs),** con el fin de proporcionar un respaldo a la liquidez del Eurosistema. Serán de aplicación hasta la entrada en vigor de la tercera edición de las subastas de liquidez a los bancos, las denominadas operaciones de financiación a plazo más largo con objetivo específico (TLTRO III, por sus iniciales en inglés), a partir de junio de 2020 hasta junio de 2021. Todo ello con el fin de garantizar que la liquidez llegue a las pymes con tipos de interés subsidiados a los bancos.
- **Reducción temporal de los requerimientos de capital por riesgo de mercado.** Junto a esas decisiones favorecedoras de la liquidez, el BCE permitió reducir temporalmente los requerimientos de capital por riesgo de mercado. El banco central reaccionaba así a la elevada volatilidad en los mercados financieros que surgió tras la extensión de la pandemia. Permite que las entidades de crédito ajusten el componente supervisor de dichos requerimientos, con el fin de mantener las actividades de creación de mercado y la liquidez en los mercados.
- **Flexibilización de los requisitos exigibles a los activos de garantía** en las operaciones de crédito del Eurosistema, aceptando activos cuya calificación crediticia esté por debajo de los requisitos mínimos de calidad crediticia actuales.
- **Ampliación del Programa de compra de bonos en el mercado secundario (PSPP),** establecido en 2015**,** hasta 870.000 millones de euros, mediante la creación del Programa de Compras de Emergencias Pandémicas (PEPP). Italia y España han sido los más claros beneficiarios de esos programas, evitando la ampliación de los diferenciales de tipos de interés de sus bonos frente a las referencias homogéneas alemanas. Entre marzo y mayo, 37.000 millones de euros fueron las adquisiciones de bonos españoles, un 14,4 % de las totales. En su inequívoca disposición a evitar males peores, tras la reunión de su Consejo de Gobierno del 4 de junio, junto a unas previsiones de manifiesta debilidad del crecimiento económico y ausencia de inflación, anunciaba la ampliación

de ese último programa de compra de bonos en 600.000 millones adicionales. Las dudas acerca de la insuficiencia del programa ante el crecimiento de las necesidades de emisión de nueva deuda de los Estados quedaba así despejada. Semanas antes, había manifestado su disposición no solo a ampliarlo, sino a admitir otros instrumentos o activos de mayor riesgo, como podrán ser bonos privados de reducida calificación crediticia, los denominados «ángeles caídos». Son decisiones consecuentes con la intensa debilidad del crecimiento, la ausencia de inflación y la persistencia de riesgos de fragmentación financiera en el seno de la eurozona, de ampliación de los diferenciales de tipos de interés entre las economías periféricas y las centrales.

La actitud del banco central es tanto más encomiable cuanto que muchas de sus decisiones han sido objeto de controversia en el seno de su Consejo de Gobierno. Las más recientes, tendentes a neutralizar la severidad de las consecuencias de la pandemia, coincidieron con la difusión de una resolución del Tribunal Constitucional alemán cuestionando precisamente la pertinencia del programa de compra de bonos públicos, lanzado en 2015. Con argumentos que fueron contestados, además de por el propio BCE, por el Tribunal de Justicia Europeo, no es la primera vez que se condicionan decisiones favorables al conjunto por parte de instituciones alemanas. Con independencia de la validez jurídica de esa resolución, lo que sí parece evidente es que no es el momento de cuestionar el grado de ortodoxia de decisiones tales. La gravedad de la situación, las amenazas ciertamente severas que pesan sobre la economía europea, sobre la estabilidad de la unión monetaria, y sobre la propia dinámica de integración regional, aconsejan no poner demasiados obstáculos a la satisfacción de ese propósito por reducir los costes muy severos de la crisis abierta.

De esa situación extrema que sufren la totalidad de las economías avanzadas tomó buena nota el Banco de Inglaterra, al acceder a la solicitud del Gobierno británico de favorecer la financiación directa al Tesoro de su país, sin mediación alguna con el fin de minimizar costes y ganar en celeridad en la lucha contra las consecuencias económicas de la pandemia. Una suerte de seguro para la financiación a corto plazo, una forma de sortear las dificultades para financiar enormes demandas de efectivo, difíciles de cubrir en el mercado de letras del Tesoro. De esa forma, el Gobierno no apela a la emisión en el mercado primario de bonos, financiando directamente por el banco central todos los costes excepcionales derivados de la gestión de la pandemia. Ante una decisión tal, los mercados financieros, los de bonos o los de divisas, ni se inmutaron.

### *5.2. Comisión Europea*

Las actuaciones de la Comisión también contrastan con las llevadas a cabo durante las crisis de 2008. Ha actuado de forma rápida, aunque las decisiones adoptadas hasta el momento han sido relativamente menos agresivas que las asu-

midas por algunos Gobiernos nacionales, incluso en Europa. Han sido 540.000 millones de euros los comprometidos en los tres programas que se comentan más adelante. Al final de este epígrafe se destaca la posible actuación que ha generado mayores expectativas y que de hecho podría contribuir en mayor medida a la reconstrucción: la constitución de un fondo específico para esa finalidad, que sigue sin concretarse a la fecha de cierre de estas páginas.

- **Mecanismo Europeo de Estabilidad (MEDE):** 240.000 millones de euros. Se trata de un instrumento precautorio destinado a financiar los gastos directos e indirectos del Covid-19. La nueva línea de crédito específica, no tendría la condicionalidad propia que llevaron consigo los «rescates» que se realizaron durante la crisis de 2008. Se mantiene la posibilidad abierta a que su uso pueda ir más allá del estrictamente sanitario. Es verdad que, a pesar de las circunstancias distintas a las que presidieron las actuaciones del MEDE en la crisis anterior, el estigma político asociado a la supuesta «inferioridad» de quien lo solicita, «que nos vean llamar a la puerta», sigue siendo un freno significativo.
- **Banco Europeo de Inversiones (BEI):** 200.000 millones de euros. Destinado a movilizar financiación a empresas, especialmente pymes. Abierto a aportaciones del presupuesto UE y agencias financieras nacionales.
- **SURE (Support to mitigate Unemployment Risks in an Emergency):** 100.000 millones de euros. Destinado a financiar las ayudas al desempleo de los países miembros, como podría ser el caso de los ERTE nuestros. Para los más optimistas debería constituir el embrión de un seguro de desempleo europeo, que vienen reivindicando España, Francia e Italia.

### *5.3. El necesario fondo para la recuperación y la reconstrucción de Europa*

Siendo relevantes esas tres modalidades de apoyo, la magnitud de las amenazas que pesan sobre el bienestar de los europeos y sobre su identificación con el proyecto de integración requiere de medidas mucho más ambiciosas. Aquellas que facilitando la recuperación y la reconstrucción no amplíen la brecha ya existente entre las economías. Han de ser, por tanto, decisiones que sin incorporar más deuda pública inmediata a las economías nacionales garanticen más demanda en el corto plazo y más inversión en el medio. Ese era el propósito del plan de reconstrucción sobre el que discutió el Consejo Europeo, que permitiría acelerar esas dos transiciones pendientes de la UE: la energética invirtiendo en tecnologías e infraestructuras compatibles con el cumplimiento de los compromisos de París 2015, demostrando en definitiva la estrecha complementariedad entre recuperación y sostenibilidad, y esa otra transición tendente a recuperar el terreno perdido en la inserción en la economía digital.

En ausencia de un plan inversor suficientemente ambicioso, expresivo de la necesaria cohesión, solvencia y modernización de la eurozona, esta volvería a conformar

un escenario propicio a tensiones financieras como las que culminaron en el verano de 2012. Es por ello que diversos Gobiernos, el español entre ellos, sometieron al Consejo Europeo diversas iniciativas en la dirección de un fondo creado mediante el endeudamiento de las instituciones europeas. El propio Parlamento Europeo respaldó el 15 de mayo un plan de reconstrucción con el apoyo de cinco grandes grupos parlamentarios totalizando el 80 % de los votos de la Cámara. Apenas unos días después conocíamos el acuerdo entre los máximos mandatarios de Alemania y Francia por apoyar una iniciativa con propósitos similares.

El 27 de mayo la Comisión daba a conocer una propuesta propia, que incorporaba buena parte de los elementos de la propuesta franco alemana, pero también las intenciones generales de la española. La descripción del denominado «Next Generation EU», es un instrumento para la recuperación o reconstrucción, «garantizando que estas son sostenibles e inclusivas de todos los Estados miembros», en el contexto de un presupuesto más ambicioso. Se presenta, efectivamente, junto al proyecto del presupuesto de la UE, del denominado Marco Financiero Plurianual, para los próximos siete años (2021-2027). Entre uno y otro, los recursos financieros a disposición de la Comisión han quedado muy significativamente reforzados.

El primero de los rasgos destacados del «Next Generation EU» es su cuantía: los 750.000 millones de euros, superior a los 500.000 millones propuestos en el plan precedente franco alemán, junto a su carácter excepcional y, desde luego, temporal. El segundo elemento singular es su financiación, mediante la emisión de deuda con vencimiento a largo plazo (entre 2028 y 2058) de la propia Comisión Europea, que aprovechará su máxima calificación crediticia. En tercer lugar, quizás el elemento más relevante para las economías periféricas, la asignación a los Estados miembros es mayoritariamente mediante transferencias (500.000 millones de euros), no solo préstamos. Es decir, sin la condicionalidad dominante en la crisis de 2008 que ejerció una presión adicional sobre los presupuestos nacionales de las economías periféricas. Aun cuando es posible que algunos aspectos de la propuesta puedan ser alterados, es razonable asumir que en las asignaciones se prioricen a los países más dañados por la pandemia, que se haga lo propio con proyectos de inversión en sostenibilidad energética y economía digital; y, que sin incorporar la clásica condicionalidad, si se exija a los beneficiarios un plan concreto de reformas en la dirección de una mayor eficiencia y modernización de las economías beneficiarias.

Sin menoscabo de la valoración inequívocamente favorable de esa propuesta de la Comisión, la experiencia nos obliga a ser cautos. El Consejo puede todavía alterar términos que ahora consideramos esenciales de la misma, e incluso reflejar algunas de las resistencias de los «cuatro frugales». El otro foco de tensión ya emergente es el relacionado con la distribución de esos fondos entre los Estados miembros y las eventuales exigencias en términos, no tanto de la condicionalidad convencional, sino de reformas estructurales a llevar a cabo en plazos determinados por los Gobiernos, incluidas las asociadas a las recomendaciones del semestre europeo para cada país. Con la necesaria flexibilidad en su aplicación, estas últimas exigencias tendentes a la modernización de las economías deberían ser bien recibidas por los Gobiernos nacionales.

### *5.4. Completar la Unión Bancaria*

Sin la urgencia de las anteriores, perfeccionar la unión bancaria es una condición necesaria para no aumentar la exposición a episodios como los que confluyeron en la crisis de la deuda soberana y bancaria en el verano de 2012. Aun cuando la situación de solvencia de los bancos en la UE es mejor que la que presidió aquel bucle diabólico, es adecuado, como ha sugerido el gobernador del Banco de España, Pablo Hernández de Cos (2020), acelerar la entrada en vigor de su fase final. Su completa centralización actuaría como el definitivo cortafuegos, ya que la prolongación de la recesión derivaría en aumentos de las tasas de morosidad y, en general, en deterioros de la calidad de los activos.

## 6. Consideraciones finales

Europa, algunas de sus economías pertenecientes a la unión monetaria, sufren las consecuencias de una de las más severas crisis de la historia cuyo origen no es precisamente económico ni financiero. Las decisiones orientadas a la neutralización de sus efectos dejarán las finanzas públicas en una situación de manifiesta precariedad, de exposición a episodios de tensiones financieras. En mayor medida si tenemos en cuenta que las economías más dañadas son las que abordaban esta crisis con un mayor déficit y deuda públicos. Las actuaciones del BCE, en especial la ampliación de su programa de compras de activos en el mercado secundario está manteniendo costes de financiación reducidos y diferenciales en los tipos de interés igualmente distantes de cualquier conformación de crisis en esos mercados. Pero junto a esa actitud del banco central será necesario que los inversores en esos mercados y en los de activos de riesgo asuman que la propia unión monetaria es viable. Eso significa hoy, disponer de mecanismos que garanticen no solo decisiones suficientemente eficaces tendentes a la recuperación, sino igualmente facilitadoras de la adopción de las propias decisiones de inversión privada hoy bloqueadas.

Deberían ser decisiones en la dirección de ese fondo tendente a la reconstrucción económica que asuma las prioridades energética y digital como vectores dominantes. Que incorporen, además, el perfeccionamiento de la transición hacia una unión económica y monetaria más completa, con más intensos grados de coordinación y de solidaridad fiscal, en la dirección que la principal organización patronal alemana, la Bundesverband der Deutschen Industrie (BDI), acaba de formular junto a sus colegas francesa e italiana. Hoy esa conjunción no es un necesario ejercicio de cohesión política, sino el prerrequisito para que la UE, y desde luego la UEM, sigan siendo plataformas adecuadas para el fortalecimiento de un mercado interior cuestionado. Es la única vía para que se satisfaga esa pretensión de la canciller Merkel cuando hacía pública la iniciativa franco-alemana del fondo de recuperación: «El objetivo es que Europa salga de esta crisis fortalecida, cohesionada y solidaria». Que así sea, podría añadirse, porque de lo contrario la próxima crisis es posible que se gestione con un censo menor de Estados miembros.

## Referencias bibliográficas

Anderson, J., Bergamini, E., Brekelmans, S., Cameron, A., Darvas, Z., Domínguez, M., & Midões, C. (2020). The fiscal response to the economic fallout from the coronavirus. *Bruegel datasets***.** Bruegel. https://www.bruegel.org/publications/datasets/covid-national-dataset/

Anderson, J., Tagliapietra, S., & Wolff, G. B. (2020). Rebooting Europe: a framework for a post Covid-19 economic recovery. *Policy Brief* 2020/01. Bruegel. https://www.bruegel.org/2020/05/rebooting-europe-a-framework-for-a-post-covid-19-economic-recovery/

Chinn, D., Klier, J., Stern, S., & Tesfu, S. (2020). Safeguarding Europe's livelihoods: Mitigating the employment impact of Covid-19. *Mckinsey*. https://www. mckinsey.com/industries/public-sector/our-insights/safeguarding-europes-livelihoods-mitigating-the-employment-impact-of-covid-19

Comisión Europea (2020). *European Economy Forecast*. Institutional Paper 125. https://ec.europa.eu/info/publications/european-economic-forecast-spring-2020_en

Hernandez de Cos, P. (24 de abril de 2020). Common policies against common shocks. *The EUROFI Magazine*. https://www.bde.es/f/webbde/GAP/Secciones/SalaPrensa/Articulos%20Prensa/hdc240420en.pdf

ILO (2020). *ILO Monitor: COVID-19 and the world of work,* 4th Edition. https://www.ilo.org/global/topics/coronavirus/impacts-and-responses/WCMS_745963/lang--en/index.htm

Ontiveros, E. (2019). *Excesos. Amenazas a la prosperidad global*. Editorial Planeta.

Ontiveros, E. (3 de mayo de 2020). *La vulnerabilidad de España*. El País, Suplemento Negocios.

Rogoff, K. (2020). Mapping the Covid-19 Recession. *Project Syndicate* (versión española en El País, 26 de abril). https://www.project-syndicate.org/commentary/mapping-covid19-global-recession-worst-in-150-years-by-kenneth-rogoff-2020-04/spanish

Summers, L. (14 de mayo de 2020). *Covid-19 looks like a hinge in history*. Financial Times. https://www.ft.com/content/de643ae8-9527-11ea-899a-f62a20d54625

Tooze, A (2019). *Crashed. How a Decade of Financial Crises Changed the World*. Viking.

MINISTERIO DE INDUSTRIA, COMERCIO Y TURISMO
ICE
INFORMACIÓN COMERCIAL ESPAÑOLA

CUADERNOS
ECONÓMICOS

## NORMAS DE PUBLICACIÓN

Los manuscritos que se envíen para publicar en Cuadernos Económicos de ICE deberán ser originales, no publicados, ni en proceso de evaluación por otra publicación. Se aceptan trabajos en castellano o en inglés. Todos los trabajos recibidos serán sometidos para su aceptación a evaluación externa anónima.

1. Los originales se enviarán en formato Microsoft Word a la dirección de correo electrónico, revistasice.sscc@mincotur.es o directamente al editor, Javier Salinas (javier.salinasj@uam.es), señalando en el asunto «Cuadernos Económicos de ICE». También pueden enviarse a través de la página web de *Revistas ICE*.

2. La extensión del trabajo, incluyendo cuadros, gráficos, tablas, notas y referencias bibliográficas, no excederá de 30 páginas. La fuente será Times New Roman 12, márgenes de 3 cm, espaciado doble y paginado en la parte inferior derecha.

3. Cada original incluirá en una primera página independiente el título del artículo y la fecha de conclusión del mismo; el nombre y apellidos de los autores, afiliación institucional, dirección, teléfono y correo electrónico de cada uno de ellos.

4. En la primera página del texto se incluirá:
   - El título en castellano y en inglés.
   - Un resumen del trabajo, en castellano y en inglés, con una extensión que no exceda de 100 palabras.
   - Una lista de 2 a 6 palabras clave
   - Los códigos (entre 1 y 5) de la clasificación del Journal of Economic Literature (JEL) que puede consultarse en el siguiente enlace https://www.aeaweb.org/jel/guide/jel.php

5. En las siguientes páginas se incorporará el texto, la información gráfica y la bibliografía con la siguiente organización:
   - Se incluirá, por este orden, introducción, desarrollo, conclusiones y referencias bibliográficas. Si hubiera anexos, se insertarán tras las referencias bibliográficas y deberán llevar título.
   - En un archivo Excel independiente se incluirá la representación gráfica (cuadros, gráficos, diagramas, figuras, etcétera), que debe llevar título, estar numerada y referenciada en el texto. En la parte inferior se incluirán la fuente de información y, en su caso, las notas aclaratorias.
   - Los apartados y subapartados se numerarán en arábigos según el siguiente modelo:

**1. Título del apartado** (primer nivel)

*1.1. Subapartado (segundo nivel)*

1.1.1. Sección del subapartado (tercer nivel)

- Las notas a pie de página irán integradas en el texto y su contenido debe estar al final de su misma página en tamaño 10, Times New Roman.
- En las citas en el texto, de acuerdo con las normas de la American Psychological Association (Normas APA), se indicará entre paréntesis el apellido del autor y el año. Ej.: (Martínez, 1991)
- Las referencias a siglas deben ir acompañadas, en la primera ocasión en que se citen, de su significado completo.
- Al final del texto se recogerá la bibliografía utilizada, bajo el título de «Referencias bibliográficas», ordenada alfabéticamente según las Normas APA. A continuación se recogen los ejemplos de referencias más utilizados:

***Libro***

Apellido, A. A. (Año). *Título*. Editorial.
Pilling, D. (2019). *El delirio del crecimiento*. Taurus.

***Libro electrónico con DOI (o URL)***

Apellido, A. A. (Año). *Título*. https://doi.org/xxx
Nicolleti, B. (2017). The Future of Fintech: Integrating Finance and Technology in Financial Services. Palgrave Macmillan. https://doi.org/10.1007/978-3-319-51415-4

***Capítulo de libro***

Apellido, A. A. (Año). Título del capítulo o la entrada. En C. C. Apellido (Ed.), *Título del libro* (pp. xx-xx). Editorial.
Demas, M., & Agnew, N. (2013). Conservation and sustainable development of archaeological sites. En I. Rizzo y A. Mignosa, *Handbook on the Economics of Cultural Heritage* (pp. 326-343). Edward Elgar.

***Publicaciones periódicas formato impreso***

Apellido, A. A. (Año). Título del artículo. *Nombre de la revista*, *volumen* (número), pp-pp.
Fernández-Blanco, V., Prieto-Rodríguez, J., & Suárez-Pandiello, J. (2017). A quantitative analysis of reading habits in Spain. *International Journal of Arts Management, 19*(3), 19-32.

***Publicaciones periódicas online con DOI (o URL)***

Apellido, A. A. (Año). Título del artículo. *Nombre de la revista*, *volumen*(número), pp-pp. https://doi.org/xxx
Amador, J., & Cabral, S. (2017). Networks of Value-added Trade. *The World Economy, 40*(7), 1265-1512. https://doi.org/10.1111/twec.12469

***Informe oficial en web***

Organismo (Año). *Título del informe*. http://www…

Fondo Monetario Internacional (2019). *Global Financial Stability Report*. https://www.imf.org/en/Publications/GFSR/Issues/2019/10/01/global-financial-stability-report-october-2019

***Ley/Reglamento***

Título de la ley. *Publicación*, número, fecha de publicación, pp-pp. http://www.boe.es....

Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa. *Boletín Oficial del Estado*, nº. 295, de 10 de diciembre de 2013, pp. 97858 a 97921. https://www.boe.es/buscar/pdf/2013/BOE-A-2013-12886-consolidado.pdf

**Orden de la lista de referencias bibliográficas**

Las referencias se ordenan alfabéticamente y, en caso de varios trabajos realizados por el mismo autor/a, el criterio es el siguiente:

- Primero los trabajos en los que el/la autor/a figura solo/a. Correlativos de año más antiguo al año más actual de publicación.
- Segundo, aquellos trabajos colectivos en los que el/la autor/a es el/la primero/a. Correlativos por fecha.
- Tercero, en caso de coincidencia exacta de autores y fechas, debe citarse cada trabajo añadiendo una letra a la fecha. Ej.: 2014a, 2014b, etc.

En la página web del grupo de revistas de Información Comercial Española se publican los artículos en formato electrónico y se recogen los principios editoriales de la publicación: ***www.revistasICE.com***

**CUADERNOS ECONÓMICOS DE INFORMACIÓNCOMERCIAL ESPAÑOLA (CICE)**
**ISSN 0210-2633**

## SUSCRIPCIÓN ANUAL

| CUADERNOS ECONÓMICOS DE INFORMACIÓN COMERCIAL ESPAÑOLA (2 NÚMEROS/AÑO) | | | |
|---|---|---|---|
| | ESPAÑA 1 año | UNIÓN EUROPEA 1 año | RESTO DEL MUNDO 1 año |
| SUSCRIPCIÓN | 30,00 € | 40,00 € | 40,00 € |
| Gastos de envío España | 1,80 € | 6,60 € | 11,32 € |
| Más 4% de IVA. Excepto Canarias, Ceuta y Melilla | 1,27 € | | |
| TOTAL | 33,07 € | 46,60 € | 51,32 € |

## EJEMPLARES SUELTOS

| CUADERNOS ECONÓMICOS DE INFORMACIÓNCOMERCIAL ESPAÑOLA | | | |
|---|---|---|---|
| | ESPAÑA 1 ejemplar | UNIÓN EUROPEA 1 ejemplar | RESTO DEL MUNDO 1 ejemplar |
| NÚMERO SUELTO | 18,00 € | 23,00 € | 23,00 € |
| Gastos de envío España | 0,90 € | 3,30 € | 5,66 € |
| Más 4% de IVA. Excepto Canarias, Ceuta y Melilla | 0,76 € | | |
| TOTAL | 19,66 € | 26,30 € | 28,66 € |

**Suscripciones y venta de ejemplares sueltos**
Ministerio de Industria, Comercio y Turismo. Centro de Publicaciones.
C/ Panamá, 1. Planta 0, despacho 3.
Teléfonos: 91 3 49 51 29 – 91 3 49 51 33 (venta de publicaciones)
91 3 49 43 35 (suscripciones)
Correos electrónico: CentroPublicaciones@mincotur.es

GOBIERNO DE ESPAÑA
MINISTERIO DE INDUSTRIA, COMERCIO Y TURISMO
SECRETARÍA GENERAL TÉCNICA
SUBDIRECCIÓN GENERAL DE DESARROLLO NORMATIVO, INFORMES Y PUBLICACIONES
CENTRO DE PUBLICACIONES

MINISTERIO DE INDUSTRIA,
COMERCIO Y TURISMO

CUADERNOS ECONÓMICOS
ICE
INFORMACIÓN COMERCIAL ESPAÑOLA

# CUADERNOS ECONÓMICOS

**ÚLTIMOS NÚMEROS PUBLICADOS**

60. Juegos dinámicos
61. Distribución de la renta y crecimiento económicos
62. La Nueva Economía Política
63. Aspectos del mercado de trabajo español
64. Crisis y reformas de los sistemas de Seguridad Social: Temas a debate (I)
65. Crisis y reformas de los sistemas de Seguridad Social: Temas a debate (II)
66. Microeconometría y conducta: Aplicaciones para el caso español
67. La economía de la salud
68. Modelos de microsimulación
69. Instrumentos derivados
70. Nuevas tendencias en Historia Económica
71. Diseño y evaluación de políticas de protección ambiental
72. Decisiones de inversión y adopción de nuevas tecnologías
73. La gestión de la actividad innovadora en las empresas
74. Capital humano y empleo
75. Desigualdad, equidad y eficiencia en salud y servicios sanitarios
76. Contribuciones recientes a la Economía del Desarrollo
77. Contribuciones a la Economía Experimental
78. Aportaciones recientes a la Economía Industrial
79. Contribuciones a la economía de la energía
80. Evaluación económica de infraestructuras de transporte
81. Análisis económico de las telecomunicaciones en España
82. Comercio internacional, empresas y competitividad
83. Economía de las energías renovables
84. Mejorar la productividad: un reto pendiente en la economía española
85. Temas actuales de Economía Política
86. Gestión ambiental de las organizaciones
87. Inmigración y crisis económica en España
88. Defensa de la Competencia en España: un análisis sectorial
89. Contribuciones a la Economía de la Innovación: análisis de la I+D+i empresarial
90. La vivienda tras la crisis financiera global: nueva visión, viejos problemas
91. Política económica europea
92. Una panorámica del sector bancario después de la crisis
93. Turismo sostenible: un debate todavía abierto a la discusión
94. Sector exterior y economía internacional
95. Capital humano y mercado laboral
96. Sistema Nacional de Salud: retos y tareas pendientes tras la Gran Recesión
97. ¿Economía colaborativa?: antecedentes, situación y desafíos en su regulación
98. Economía de la Cultura

CUADERNOS ECONÓMICOS

ICE

INFORMACIÓN COMERCIAL ESPAÑOLA

GOBIERNO DE ESPAÑA

MINISTERIO DE INDUSTRIA, COMERCIO Y TURISMO

9 770210 263007 00099

P.V.P.: 18 € + IVA