



# **UNIVERSIDAD NACIONAL DEL SUR**

TESIS DE MAGISTER EN DERECHO CON ORIENTACIÓN EN  
ANÁLISIS ECONÓMICO

NORMAS, PREFERENCIAS SOCIALES E INTENCIONES EN  
DOS ESTUDIOS EXPERIMENTALES SOBRE CORRUPCIÓN

CARLOS MAXIMILIANO SENCI

BAHÍA BLANCA

ARGENTINA

2014



## **PREFACIO**

Esta Tesis se presenta como parte de los requisitos para optar al grado Académico de Magíster en Derecho con Orientación en Análisis Económico, de la Universidad Nacional del Sur y no ha sido presentada previamente para la obtención de otro título en esta Universidad u otra. La misma contiene los resultados obtenidos en investigaciones llevadas a cabo en el ámbito del Departamento de Derecho durante el período comprendido entre el 2011 y el 2014, bajo la dirección del Dr. Fernando Tohmé y la co-dirección del Dr. Esteban Freidin.

-----

UNIVERSIDAD NACIONAL DEL SUR  
Secretaría General de Posgrado y Educación Continua

La presente tesis ha sido aprobada el...../...../....., mereciendo calificación la de.....(.....)

---

---

*A mi madre*

## Agradecimientos

La sección dedicada a los agradecimientos suele ser la más pequeña de una tesis. Sin embargo, estoy convencido que la importancia que se le dedica es inversamente proporcional a la gratitud que inspira las palabras de quien la escribe.

En primer lugar agradezco a mi director de tesis, el Dr. Fernando Tohmé, por la excelente predisposición que mostró siempre y la motivación que me transmitió en todo momento. A mi co-director, el Dr. Esteban Freidin, lo hago responsable de haberme hecho creer que un filósofo puede hacer ciencia. A él, infinitas gracias. Espero que este sea el comienzo de una fructífera colaboración con ambos en el futuro.

Agradezco asimismo al Departamento de Derecho de la Universidad Nacional del Sur. La Maestría en Derecho creada recientemente aporta un espacio creativo e intelectualmente sólido, donde estoy seguro que se escribirán tesis de excelente calidad.

Agradezco también al Departamento de Sociología de la Universidad de Groningen, y en general a todos los miembros del ICS de Groningen.

En el camino que culminó en esta tesis, muchas personas tuvieron alguna que otra responsabilidad en el resultado final: le agradezco a Guille Zampar, porque cuando se aburría de los experimentos con levaduras, se interesaba por los económicos. A Manu Muñoz-Herrera, porque me ayudó en las complejidades del z-tree, siempre a ritmo de cumbia.

Por último, agradezco profundamente a quienes me acompañaron con su amistad y me brindaron su apoyo e hicieron de mi larga experiencia invernal en Groningen una experiencia cálida: a Soledad Ayala, James “Jaimito” Leigh, Pedrito Miguel, Andrés Tello, Lieuwe Zijlstra, Menno Rol, Fernanda Andrés, Marina Montelongo, Britta Rüschoff, Guille Zampar, Lia Nunes, Isabel Estrada, Michael Richardson, Katharina Fietz, Rik Smit, Angélica López Angarita, y a Silvana Della Penna por prestarme una notebook cuando a algún habitante groninguense se le ocurrió que bien valía la pena correr el riesgo de apropiarse ilegalmente de la mía. Gracias a todos.

Bahía Blanca, Septiembre de 2014

Maximiliano Senci

## Resumen

Los estudios experimentales sobre corrupción se consideran como un área particular de los estudios más generales sobre preferencias sociales. La tesis que se presenta está compuesta por dos artículos experimentales sobre corrupción, en los cuales se analiza el rol de las preferencias y las intenciones en juegos de corrupción diseñados a tal fin.

En el primer estudio examinamos el llamado “efecto de marco”, es decir, el efecto de que las decisiones que tomamos los individuos son sensibles al contexto en el que se presentan, y el rol de los incentivos egoístas en un juego de corrupción, concebido como una variante de un juego de coima de Abbink et al. (2002). Los efectos de marco varían entre estudios, y los estudios sobre corrupción han reportado resultados mixtos. La literatura sugiere que los efectos de marco pueden estar mediados por los costos morales que inducen en los participantes. Para alcanzar ambos objetivos comparamos un marco abstracto o “libre de contexto” frente a un marco “cargado” en un juego de corrupción. Al mismo tiempo, variamos el monto que los sujetos pueden ganar en los diferentes escenarios lo que nos permite estudiar el punto de quiebre en el que los sujetos pasan de elegir la opción “honesta” a elegir la opción corrupta. Adicionalmente implementamos un test socio-psicológico para investigar hasta qué punto la disposición manipulativa de los participantes puede explicar la variación en el comportamiento individual en los diferentes marcos. Encontramos que los sujetos consistentemente se comportan de manera menos corrupta en el marco cargado. Es más, las medidas comportamentales post-experimento utilizadas proveen soporte parcial a la hipótesis de que las respuestas individuales al marco podrían estar vinculadas a las reacciones suscitadas por el marco cargado, que interpretamos como un incremento del costo moral. El trabajo elabora en mayor detalle la última afirmación.

En el segundo trabajo argumentamos que en la literatura experimental sobre corrupción no se ha tenido en cuenta debidamente el contrato fiduciario que define a la relación entre Principal y Agente. Introducimos un juego experimental en el que el tercero perjudicado por las externalidades es al mismo tiempo el Principal que puede tomar un rol activo o pasivo en el juego. Ser activo significa que puede entrar en un juego de confianza con el Agente, mientras que si es pasivo esa interacción entre Principal y Agente está ausente,

y el primero se transforma en un tercero pasivo. Comparando las dos condiciones, podemos testear hipótesis de la literatura de preferencias sociales, específicamente si los actores exhiben preferencias sociales y si evitan perjudicar a terceros que pueden señalar intenciones amables. Contraviniendo el abordaje de la teoría de preferencias sociales en corrupción, no encontramos respaldo a la hipótesis de que las intenciones de los terceros que sufren la externalidad puedan disminuir la corrupción. Los análisis subsiguientes no brindan respaldo a la noción de que la reciprocidad entre los actores corruptos esté modulada por las percepciones de amabilidad.

## **Abstract**

Experimental studies on corruption can be considered as a sub-field of the broader area of social preferences. The present thesis is comprised of two experimental papers on corruption, in which we analyze the effect of social preferences and intentions in two games designed for that purpose.

The first paper examines the effects of framing and potential gains in a corruption game, conceived as a variant of Abbink et al's (2002) bribery game. The effect of framing refers to the fact that individuals' decisions are sensitive to the context in which they are presented. Studies on framing effects in corruption have reported mixed results. The literature suggests that framing effects might be moderated by psychological or moral costs associated to norm breaking or corruption, which might be differentially stressed by the different frames. To reap both objectives we compare an abstract vs. a loaded framing of the stage game. By incrementally varying the amount which subjects potentially earn throughout several scenarios, we study the tipping point at which subjects switch to the corrupt option. In addition we implement a social-psychological test to assess to what extent personality measures could explain variation in behavior across between subjects' conditions. We find that subjects consistently exhibit less corruption in the loaded framing. Moreover, post-experimental personality measures add support to the hypothesis that subjects' responses to frames may be entwined with the emotional reactions cued by them. This paper elaborates further on this topic.

In the second paper we study social preferences and intentions in an experimental game designed to explicitly take into account the fiduciary contract that binds the Principal and the Agent. We cogently argue that this is a structural element of corruption that should be taken into account in any experimental setting. A decisive difference between an ultimatum game or a trust game and any experimental game that realistically intends to mimic corruption is that in the latter case successful corruption imposes a negative externality to a third party. If actors value the welfare of thirds, they will give externalities due weight in their decisions. We developed an experimental game in which the Principal suffering from the negative externalities can take on an active or a passive role. By comparing these two

conditions, we can test behavioral predictions coming from the social preferences literature, specifically whether or not actors display social preferences and refrain to harm a third party. Contravening the prominent social preferences' approaches to corruption, we did not find support for the claim that intentions of thirds play a role in shaping actor's behavior. Further analyses did not support the notion that reciprocity between corrupt parties is driven by the perception of kindness.

## ÍNDICE

<b>Lista de Figuras</b> .....	12
<b>Lista de tablas</b> .....	13
<b>Introducción general</b> .....	14
<b>Bibliografía de la Introducción general</b> .....	19
<b>Capítulo 1</b> .....	21
<b>Efectos de marco e incentivos egoístas en un juego de coima*</b> .....	21
<b>1. Introducción</b> .....	22
<b>2. Literatura relacionada: versiones anteriores de JdC</b> .....	26
<b>3. El experimento</b> .....	30
3.1 Nuestra variante del juego de coima .....	30
3.2. <i>Principales diferencias metodológicas entre nuestro JdC y versiones anteriores</i> .....	33
<b>4. Predicciones</b> .....	35
<b>5. Procedimiento</b> .....	38
5.1. <i>La explicación pública del asistente</i> .....	39
5.2. <i>Contenido de los cuadernillos</i> .....	39
5.3. <i>Pago</i> .....	42
5.4. <i>Marco</i> .....	43
5.5. <i>Variables dependientes</i> .....	44
<b>6. Análisis estadísticos</b> .....	45
<b>7. Resultados y discusión</b> .....	46
7.1. <i>Participantes A</i> .....	46
7.2. <i>Participantes B</i> .....	53
<b>8. Conclusiones</b> .....	58
<b>Referencias bibliográficas del capítulo 1</b> .....	62
<b>Preferencias sociales, externalidades e intenciones en un juego de corrupción</b> .....	66
<b>1. Introducción</b> .....	67
<b>2. Bases teóricas</b> .....	74
2.1. <i>Teorías de preferencias sociales</i> .....	74
2.2 <i>Trabajos relacionados</i> .....	78
<b>3. Diseño experimental</b> .....	83
3.1. <i>Tratamiento Base</i> .....	87
3.2. <i>Tratamiento Activo</i> .....	88

<b>3.3. Procedimiento</b> .....	88
<b>4. Predicciones</b> .....	90
<b>5. Datos y variables</b> .....	95
<b>6. Resultados y discusión</b> .....	96
<b>6.1. Estadística descriptiva</b> .....	97
<b>6.2. Análisis estadístico</b> .....	98
<b>6.3. Resultados de las regresiones</b> .....	99
<b>7. Consideraciones finales</b> .....	113
<b>Referencias bibliográficas del Capítulo 2</b> .....	117
<b>Apéndice 1</b> .....	125
<b>Apéndice 2</b> .....	128
<b>Apéndice 3</b> .....	129
<b>Referencias Bibliográficas completas</b> .....	132

## Lista de Figuras

**Figura 1** (p. 31): Árbol de decisión del Juego de Coima en nuestro experimento.

**Figura 2** (p. 48): Porcentaje de participantes A que eligieron la opción L en el marco abstracto (n =26) o que eligieron “ofrece trato corrupto” en el marco cargado (n = 27).

**Figura 3** (p. 52): Correlación entre puntajes de maquiavelismo y frecuencia de coima en los participantes A en función de los tratamientos de marco.

**Figura 4** (p. 54): Porcentaje de participantes B que aceptaron el trato corrupto graficado en función de los escenarios y los tratamientos.

**Figura 5** (p. 86): Ilustración del juego en forma extensiva correspondiente al tratamiento base tal como fue utilizado en el experimento. Los pagos entre corchetes corresponden a los pagos de A, B y C, respectivamente.

**Figura 6** (p. 87): Ilustración del juego en forma extensiva correspondiente al tratamiento activo tal como fue utilizado en el experimento. Los pagos entre corchetes corresponden a los pagos de A, B y C, respectivamente.

**Figura 7** (p. 97): Porcentaje de participantes A que eligieron *in*.

**Figura 8** (p. 98): Elecciones de B en función de ronda y tratamiento.

**Figura 9** (p. 98): Elecciones de C en función de ronda y tratamiento.

## Lista de tablas

**Tabla 1** (p. 103): Estimaciones de MCMC de los coeficientes de una regresión logística multinivel. Interacciones de Tratamiento con una variable *dummy* para Rol C.

**Tabla 2** (p. 106): Estimación usando MCMC de una regresión logística multinivel con una variable de nivel 1 que mide la percepción de amabilidad atribuida a la acción de A.

**Tabla 3** (p. 109) Estimación de los coeficientes para AmabAB a partir de un modelo multinivel de regresión logística con intercepto aleatorio.

**Tabla 4** (p. 110): Regresión logística multinivel con intercepto aleatorio estimada con MCMC. La variable AmabAC representa la evaluación de C de la amabilidad de A de acuerdo a una escala de 7 puntos de Likert. La acción *in* representa nuestra variable dependiente.

**Tabla 5** (p. 111): Regresión logística multinivel con intercepto aleatorio. Estimación de Cadenas de Markov de Monte Carlo (MCMC) de los parámetros de los coeficientes. Modelos con amabilidad de los jugadores en rol B. Sólo jugadores en rol C.

**Tabla 6** (p. 129): Estimación de los coeficientes de regresión por medio de IGLS (*Iterative generalized least squares*) de MlwiN. Atribuciones de amabilidad de B hacia A como variable dependiente.

**Tabla 7** (p. 130): Estimaciones de los coeficientes con IGLS (*Iterative generalized least squares*) de MlwiN. La variable dependiente es la atribución de amabilidad de C hacia A.

**Tabla 8** (p. 131): Estimaciones de los coeficientes con IGLS (*Iterative generalized least squares*) de MlwiN.

## Introducción general

La presente tesis contiene dos artículos independientes en el área de los estudios experimentales sobre corrupción. El primer artículo examina el llamado efecto de marco (*framing effect*) en un juego de corrupción, que intenta capturar el costo moral por incurrir en ese tipo de actividades ilegales. El segundo trabajo extiende estudios previos de corrupción incorporando la relación entre Principal y Agente por medio de un juego de confianza.

Los estudios experimentales sobre corrupción pueden considerarse una sub-área del marco más amplio de estudios sobre preferencias sociales. La literatura de preferencias sociales ofrece amplia evidencia de que las personas toman en cuenta el bienestar de los demás (Kahneman et al., 1986; Camerer & Thaler, 1995; Fehr & Gächter 2000; Simpson & Willer, 2008). Teorías como las de Fehr & Schmidt (1999) y Bolton & Ockenfels (2000) suponen motivos no egoístas en las personas que se exhiben en ciertas preferencias por las consecuencias distributivas que surgen de una interacción (i. e., preferencias “consecuencialistas”). Si las personas tienen preferencias sociales, entonces deberían tomar en cuenta la presencia de externalidades negativas a la hora de tomar sus decisiones. La corrupción típicamente genera externalidades negativas, o más llanamente, perjuicios a terceros pasivos. Sin embargo la literatura experimental sobre corrupción no ha encontrado evidencia clara de que las personas tomaran en cuenta el daño causado a terceros (ver Abbink et al., 2002; Abbink, 2006; Barr & Serra, 2009). Es posible que la práctica habitual en economía de utilizar juegos abstractos que no puedan asociarse directamente a un contexto, limitara la expresión de preferencias sociales en los participantes. Por eso, en estudios posteriores Abbink y colaboradores, y Barr & Serra, utilizaron presentaciones “cargadas”, es

decir con palabras que sugerían un contexto de interacción determinado. Es claro que las preferencias de los individuos dependerán del contexto de interacción y de las normas sociales relevantes en dicho contexto (ver Bicchieri, 2006). Si la “presentación” puede hacer que las consecuencias negativas asociadas a los actos de corrupción adquieran mayor disponibilidad cognitiva (e. g., estén en el foco de atención), podrían disuadir a los participantes de corromperse (Abbink, 2006). Sin embargo, los resultados obtenidos al respecto fueron mixtos (cfr. Barr y Serra, 2009). Con el objetivo de contribuir a esta discusión, el primer capítulo de la tesis examina el llamado “efecto de marco” o de presentación, es decir, el efecto de que las decisiones que tomamos los individuos son sensibles al contexto en el que se presentan, y el rol de los incentivos egoístas en un juego de corrupción, concebido como una variante del juego de coima de Abbink et al. (2002). La literatura sugiere que los efectos de marco pueden estar mediados por los costos morales que inducen en los participantes. Para alcanzar ambos objetivos comparamos un marco abstracto o “libre de contexto” frente a un marco “cargado”. Al mismo tiempo, variamos el monto que los sujetos pueden ganar en los diferentes escenarios, lo que nos permite estudiar el punto de quiebre en el que los sujetos pasan de elegir la opción “honesta” a elegir la opción corrupta. Adicionalmente implementamos un test socio-psicológico para investigar hasta qué punto la disposición manipulativa de los participantes puede explicar la variación en el comportamiento individual en los diferentes marcos. Encontramos que los sujetos consistentemente se comportan de manera menos corrupta en el marco cargado. Es más, las medidas comportamentales post-experimento utilizadas proveen soporte parcial a la hipótesis de que las respuestas individuales al marco podrían estar vinculadas a las reacciones suscitadas por el marco cargado, que interpretamos como un incremento del costo moral. El trabajo elabora en mayor detalle la última afirmación.

Otro abordaje en la literatura de preferencias sociales sostiene que además de los resultados los individuos tienen en cuenta las intenciones con los que interactúan en sus decisiones (Rabin, 1993; Falk & Fischbacher, 2006). El trabajo seminal en esta dirección pertenece a Rabin (1993). La idea básica de Rabin consiste en que los individuos queremos ser amables con quienes son amables con nosotros, lo que da lugar al establecimiento de una relación de reciprocidad. Responder con amabilidad las intenciones amables de otros (reciprocidad positiva) o de manera no amable las motivaciones egoístas o competitivas (reciprocidad negativa) son ejemplos de preferencias sociales “procesuales” (Falk & Fischbacher, 2006).

Los experimentos sobre corrupción han demostrado que la reciprocidad es importante para el establecimiento de relaciones corruptas (Lambsdorff, 2012). El trabajo pionero de Abbink, Irlenbusch & Renner (2002) modela una situación de coima entre una Empresa y un Funcionario, que se caracteriza por tres elementos: en primer lugar, un componente de reciprocidad: la Empresa y el Funcionario pueden establecer una relación de beneficio mutuo. Las partes no pueden invocar la ley para hacer cumplir el acuerdo corrupto, por lo tanto recurren a mecanismos extralegales como la confianza y la reciprocidad. En segundo lugar, la corrupción genera externalidades negativas en terceros. Y en tercer lugar, la corrupción es riesgosa, lo que significa que quienes incurren en actos de corrupción son pasibles de ser sancionados. La conclusión más importante del trabajo de Abbink y colaboradores, consiste en que el nivel de corrupción (ofrecimiento y aceptación de coimas) es una función de la reciprocidad entre las partes.; cuanto mayor es la coima, mayor reciprocidad hay entre las partes. Por otra parte, el establecimiento de confianza y reciprocidad entre las partes parece excluir consideraciones de equidad distributiva hacia terceros (Abbink, 2005).

El juego de Abbink y colaboradores tiene muchos méritos, sin embargo, se reduce a un juego de confianza simple con externalidades. En este sentido, le falta un componente que está presente en la gran mayoría de las definiciones de corrupción, nos referimos a la relación entre Principal y Agente. En el segundo artículo incorporamos el contrato implícito entre Principal y Agente. Difícilmente pueda hablarse corrupción sin tener en cuenta este contrato. De hecho el modelo del Principal-Agente ha sido el modelo teórico más extensamente utilizado en la literatura sobre corrupción (e. g., Banfield, 1975; Rose-Ackermann, 1978, 2010; Klitgaard, 1988; Shleifer & Vishny, 1993; Lambsdorff et al., 2005; della Porta & Vanucci, 2012). En la corrupción burocrática o política el principal es el Estado, o los ciudadanos, mientras que el actor corrupto es un Funcionario. La corrupción es, entonces, básicamente un juego de tres jugadores en el que el intercambio entre el Funcionario y un tercero (el cliente) distorsiona los incentivos que tiene el Funcionario para cumplir el contrato con el Principal. Es claro que, el Funcionario se corromperá dependiendo de la combinación que exista entre los incentivos monetarios, la posibilidad de ser castigado y los propios escrúpulos morales.

En el segundo capítulo que compone esta tesis argumentamos que en la literatura experimental sobre corrupción no se ha tenido en cuenta debidamente el contrato implícito que define a la relación entre Principal y Agente. Introducimos un juego experimental en el que el tercero perjudicado por las externalidades es al mismo tiempo el Principal que puede tomar un rol activo o pasivo en el juego. Ser activo significa que puede entrar en un juego de confianza con el Agente, mientras que si es pasivo esa interacción entre Principal y Agente está ausente, y el primero se transforma en un tercero pasivo. Comparando las dos condiciones, podemos testear hipótesis de la literatura de preferencias sociales,

específicamente si los actores exhiben preferencias sociales y si evitan perjudicar a terceros que pueden señalar intenciones amables. Contraviniendo el abordaje de la teoría de preferencias sociales en corrupción, no encontramos respaldo a la hipótesis de que las intenciones de los terceros que sufren la externalidad puedan disminuir la corrupción. Los análisis subsiguientes no brindan respaldo a la noción de que la reciprocidad entre los actores corruptos esté modulada por las percepciones de amabilidad.

## Bibliografía de la Introducción general

- Abbink, K. (2005) Fair salaries and the moral costs of corruption. En Boicho Kokinov (Ed.), *Advances in Cognitive Economics*. Sofia: NBU Press.
- Banfield, E. (1975) Corruption as a Feature of Governmental Organization, *Journal of Law and Economics*, 18(3): 587–605.
- Bicchieri (2006) *The Grammar of Society. The Nature and Dynamics of social norms*, Cambridge, Cambridge University Press.
- Bolton, G. & Ockenfels, A. (2000) ERC: A Theory of Equity, Reciprocity, and Competition, *American Economic Review*, 90 (1), 166-193.
- Camerer, C. & Thaler, R. (1995) Ultimatums, Dictators and Manners, *Journal of Economic Perspectives*, 9(2), 209-219.
- Della Porta, D. & Vanucci, A. (2012) *The Hidden order of corruption. An Institutional Approach*, Ashgate.
- Falk, A. & Fischbacher, U. (2006) A Theory of reciprocity, *Games and Economic Behavior*, 54(2), 293-315.
- Fehr, E. & Gächter, S. (2000) Cooperation and punishment in public goods Experiments, *American Economic Review*, 90, 980-994.
- Fehr, E. & Schmidt, K. (1999) A Theory of Fairness, Competition, and Cooperation, *Quarterly Journal of Economics*, 114(3), 817-868.
- Kahneman, D., Knetsch, J. & Thaler, R. (1986) Fairness as a Constraint on Profit Seeking: Entitlements in the Market, *American Economic Review*, 76(4), 728 -741.
- Klitgaard, (1988) *Controlling Corruption*, California, University of California Press.
- Lambsdorff, J. G. Taube, M., & Schramm, M. (Ed.) (2005) *The New Institutional Economics of Corruption*, New York, Routledge.

- Lambsdorff, J. G. (2012) Behavioral and experimental economics as a guidance to anticorruption, en Serra, D. & Wantchekon, L. (Eds.) (2012) *New advances in the experimental research on corruption*
- Rabin, M. (1993) Incorporating Fairness into Game-Theory and Economics, *American Economic Review*, 83(5), 1281-1302.
- Rose-Ackerman, Susan (1975) The Economics of Corruption, *Journal of Public Economics*, IV, 187-203.
- Rose-Ackerman, S. (2010) The Law and Economics of Bribery and Extortion, *Annu. Rev. Law Soc. Sci.*, 6: 217–38.
- Shleifer, A. & Vishny, R. (1993) Corruption, *The Quarterly Journal of Economics*, 108 (3), pp. 599-617
- Simpson, B. & Willer, R. (2008) Altruism and Indirect Reciprocity: The Interaction of Person and Situation in Prosocial Behavior, *Social Psychology Quarterly*, 71:37-52.

## Capítulo 1

### Efectos de marco e incentivos egoístas en un juego de coima\*

---

\* Trabajo en conjunto con Esteban Freidin.

## 1. Introducción

Según una caracterización comúnmente aceptada en las ciencias sociales la corrupción involucra, por excelencia, la violación de normas (Graeff, 2005; Schweitzer, 2005). Una norma social representa un consenso al que ha llegado una sociedad o grupo de personas respecto de la legitimidad de un comportamiento (Coleman, 1990; Ostrom, 2000, Bicchieri, 2006), y supone expectativas en torno a su cumplimiento; dicho de otra manera, cuando los miembros de un grupo consideran aplicable una norma en un determinado contexto, esperan asimismo que los otros miembros del grupo respeten la norma y que castiguen su incumplimiento (ver Bicchieri, 2006).

Hay dos mecanismos principales que suscitan el seguimiento de las normas; por un lado, la transgresión de las expectativas normativas asociadas al seguimiento de una norma da lugar a que la transgresión sea castigada por pares (ver Fehr & Gächter, 2002); por otro lado, el castigo social puede ser auto-infligido a partir de emociones como la vergüenza y la culpa (Elster 1989; Benabou & Tirole, 2004). En este sentido puede también decirse que las normas sociales forman parte de la identidad social de un individuo (Akerlof & Kranton, 2000), por eso una desviación de la norma puede afectar la propia identidad del actor. Hay una serie de trabajos que muestran que el deseo de conservar la propia identidad guarda una relación estrecha con la percepción que se tiene de sí mismo en tanto persona moral, y que dicha relación define el castigo auto-infligido (Amir et al., 2009). Cuando un individuo ha internalizado una norma, desviarse del cumplimiento de ella representa un costo en términos psicológicos o morales. Dicho de otro modo, el costo moral representa la pérdida de utilidad

que sufre un individuo cuando se corrompe (Della Porta & Vanucci, 2012, p. 57-76). Es claro, entonces, que el costo moral entra en la ponderación subjetiva que realiza el individuo entre la recompensa potencial y el escrúpulo moral que pueda generarle desviarse de una norma o incurrir en un acto delictivo.

En la investigación experimental de la corrupción, se ha abordado el estudio de la relación entre el costo moral asociado a la transgresión de las normas sociales y el comportamiento corrupto, entre otras formas, a través del método del *marco* social, es decir, de “presentar” a los sujetos la tarea de decisión induciendo un contexto determinado. En un sentido pre-teórico, *marco* se refiere a la idea de que además de la estructura de pagos existen otras características del proceso de decisión que son sistemáticamente explotadas por los individuos.<sup>1</sup> El adjetivo “social” califica el hecho de que los comportamientos estudiados tienen una interpretación pretendidamente social, es decir, que se refiere a situaciones de interacción social entre dos o más individuos y no a situaciones de decisión individual.

La literatura sobre *framing* muestra que ciertas claves contextuales sutiles pueden tener influencia sobre la orientación pro-social de los sujetos. Por ejemplo, hay estudios que utilizan técnicas de *priming* y muestran que con exponer a los sujetos a adjetivos que resalten la competitividad o la cooperación de la situación puede modificarse su comportamiento (Hertel & Fiedler, 1994, 1998). Otro estudios muestran que sólo basta con simplemente *nombrar* un dilema social como “Juego de la Comunidad” o “Juego de Wall-Street”, para lograr una influencia decisiva sobre las preferencias sociales y la cooperación (Lieberman et

---

<sup>1</sup>Esto es evidente ya en el artículo de Tversky y Kahneman de 1981, en el que introducen el concepto de *marco* de la siguiente manera: “We use the term 'decision frame' to refer to the decision maker conception of acts, outcomes and contingencies associated with a particular choice” (1981, p. 451). A continuación especifican aún más su significado: “The frame that a decision maker adopts is controlled partly by the formulation of the problem and partly by the norms, habits, and personal characteristics of the decision maker” (1981, p.451).

al., 2004; Lindenberg & Steg, L., 2007). Esta literatura sugiere que no es posible, o al menos que podemos obtener resultados sesgados, si medimos las preferencias “fuera de contexto”, precisamente porque el contexto tiene un efecto sobre las preferencias sociales.

Es por eso que la razón intuitiva para favorecer el uso de instrucciones “cargadas” en un experimento económico diseñado para estudiar el fenómeno de la corrupción, consiste en que las connotaciones morales negativas que con seguridad afectan las actitudes que tenemos frente a la corrupción en la vida real, pueden resultar neutralizadas en un experimento que utilice un *marco* abstracto. Esto es, un fraseo cargado puede hacer que las consecuencias negativas asociadas a los actos de corrupción adquieran mayor disponibilidad cognitiva (e.g., estén en el foco de atención) y por ello mismo disuadan a los participantes de corromperse (Abbink, 2006).

No obstante los intentos de algunos investigadores por comprender el fenómeno, no hay evidencia concluyente respecto de los efectos de *marco* en juegos de corrupción. Comparaciones directas entre tratamientos sugieren que el nivel de corrupción puede ser insensible al *marco* (Abbink y Hennig-Schmidt, 2006). También se ha argumentado que la utilización de palabras moralmente cargadas puede resultar completamente artificial a los participantes y puede inducir un juego de roles (*role playing*), a la manera por ejemplo en que lo haría en un juego de naipes (Bardsley, 2005; Barr y Serra, 2009).

Con el objetivo principal de estudiar el efecto de *marco* en un contexto de corrupción y contribuir a la discusión en torno a su utilidad, introducimos una variante de un Juego de Coima (JdC a partir de ahora) en el cual el potencial oferente de coima puede elegir entre tres opciones: abstenerse de iniciar una transacción, elegir jugar un juego de confianza (*Trust Game*), o elegir un Juego de confianza *con externalidades negativas* (simulando una situación de corrupción). Incrementamos los incentivos egoístas de la opción corrupta a lo

largo de los escenarios, lo que nos permite estudiar el punto de quiebre en el cual los individuos pasan de la opción honesta a la opción corrupta.

Además, en el presente trabajo vamos más allá de los estudios anteriores sobre *marco* en corrupción, ya que además de la tarea experimental nuestros participantes responden al cuestionario de Maquiavelismo Mach IV (Christie & Geis, 1979). El cuestionario de Maquiavelismo mide el nivel de egoísmo y oportunismo de un individuo, características psicológicas que muchos considerarían típicamente inmorales. De ahí que de aquellos que puntúen bajo en la escala de maquiavelismo (*Mach scale*) puede decirse que exhiben tendencias pro-sociales. Esta combinación entre toma de decisiones y cuestionarios socio-psicológicos permite explorar las soluciones motivacionales a dilemas sociales (ver Kollock, 1998). Estas últimas asumen que los actores no son completamente egoístas y despliegan algún tipo de pro-socialidad que se exhibe en la ponderación que realizan entre la ganancia propia y ajena. Es útil recurrir a esta distinción, ya que comprender e identificar que hay individuos que debido a determinadas características de personalidad (e.g. Maquiavelismo) consistentemente otorgan un peso relativo diferente a los pagos propios y ajenos, es importante para comprender las tasas de cooperación en dilemas sociales. Medidas como la de maquiavelismo, pueden ayudarnos a dilucidar respuestas diferenciales a los efectos de *marco*. Es más, evaluar diferencias individuales de las respuestas de los sujetos a los tratamientos es importante, porque, en primer lugar, puede que uno encuentre efectos nulos de *marco* que en el fondo pueden ocultar diferencias individuales lo suficientemente sutiles como para que el *marco* no permita expresarlas; por otro lado, precisamente encontrar diferencias individuales al *marco* puede ayudar a desentrañar el mecanismo subyacente que opera en el *marco*.

Naturalmente, esto puede ser de gran utilidad para interpretar efectos de *marco*. Si

son los individuos pro-sociales quienes fundamentalmente responden a él, es posible argumentar que el *marco* trabaja incrementando los costos morales (por no adecuarse a la norma). El cuestionario nos provee de una medida independiente para evaluar esa afirmación.

El presente estudio está organizado de la siguiente manera: en la sección 2, presentamos los JdC de Abbink et al. (2002; Abbink & Hennig-Schmidt, 2006) y Barr & Serra (2009) y discutimos sus resultados tomando en consideración hasta qué punto los diseños capturan los costos morales que la corrupción involucra. En la sección 3, discutimos literatura relacionada. La sección 4 presenta nuestra variante del JdC y los métodos utilizados en nuestro experimento, incluyendo la implementación de los tratamientos de *marco* (abstracto versus “cargado”) y el uso del cuestionario de Maquiavelismo Mach IV. En la sección 5 presentamos nuestras predicciones. En la sección 6 describimos en detalle los resultados de nuestro experimento cuya conclusión principal consiste en que tanto potenciales oferentes de coimas como potenciales coimeados fueron sensibles al fraseo cargado. Además se encontró una correlación positiva entre la probabilidad de elegir la opción corrupta y la medida de Maquiavelismo (aunque sólo para potenciales oferentes de coimas en el *marco* cargado).

Para finalizar, en la última sección discutimos nuestros resultados, así como las limitaciones y potenciales aplicaciones de nuestro estudio.

## **2. Literatura relacionada: versiones anteriores de JdC**

Abbink y colaboradores (2002) propusieron el Juego de Coima (*Bribery Game*) original, con la pretensión de ser interpretado como una interacción entre una empresa y un funcionario público. En el juego pares de participantes interactuaron por 30 rondas (siempre

con los mismos compañeros) ya sea en un *Juego de Confianza* (tratamiento control), un *Juego de Confianza* con externalidades negativas (tratamiento NE), o un *Juego de Confianza* tanto con externalidades negativas como con posibilidad de castigo (tratamiento NE + C). En primer lugar, en el Juego de Confianza el Jugador 1 (la empresa a partir de ahora) debía optar por transferir puntos experimentales (convertibles a dinero al final del experimento) al Jugador 2 (el Funcionario Público a partir de ahora), quien a su vez debía elegir entre las opciones X e Y. Asumiendo que la Empresa transfirió dinero al Funcionario Público, la elección de la opción Y indicaba reciprocidad de parte del Funcionario Público, tanto porque Y significaba un pago más alto para la Empresa que X, como porque la opción X maximizaba el pago para el Funcionario Público. En segundo lugar, para modelar las externalidades negativas de la corrupción, los autores implementaron el tratamiento NE, en el que cada vez que la Empresa transfería dinero al Funcionario Público (simulando un ofrecimiento de coima) y el Funcionario Público elegía la opción Y (simulando la provisión del servicio corrupto), todos los otros pares de jugadores en la misma sesión experimental sufrían una pérdida importante de puntos. Por último, Abbink y colaboradores modelaron el costo potencial por ser descubiertos y castigados implementando una variación al tratamiento NE, en el que incluían una posibilidad baja de detección, pero que significaba un castigo monetario drástico (pérdida de todos los puntos y exclusión del juego), cada vez que se concretaba un trato corrupto (es decir, cada vez que la Empresa transfería dinero y el Funcionario Público elegía la opción Y).

Los resultados principales del experimento fueron los siguientes: a) Empresas y Funcionarios Públicos desarrollaron confianza y reciprocidad en todos los tratamientos, lo que significa que hubo transferencias significativas de dinero de las Empresas a los Funcionarios Públicos, y que transferencias mayores estuvieron asociadas con una mayor

probabilidad de que los Funcionarios Públicos elijan Y; b) tanto el promedio de transferencias de las Empresas como el promedio de elección de Y por parte de los Funcionarios Públicos fue más alto en los tratamientos control y NE que en el tratamiento NE + C; c) los participantes en los tratamientos control y NE no difirieron en cuanto al promedio de transferencias ni en cuanto a la propensión a elegir la opción Y. De estos resultados, Abbink y colaboradores concluyeron que los participantes fueron insensibles al daño que significaban las externalidades para los terceros, y que sólo vieron modificados su comportamiento cuando existió la posibilidad de que sean detectados y castigados (Abbink et al., 2002).

El patrón de resultados encontrados en Abbink y colaboradores (2002), esto es, que la reciprocidad por parte del Funcionario Público es una función del nivel de la coima recibida, ha encontrado apoyo y extensión en otros estudios. Primero, en juegos de coima similares (Abbink, 2004, 2005; Abbink & Hennig-Schmidt, 2006) y en otros trabajos que se han enfocado en otros tópicos (Fran & Schulze, 2000; Schulze & Frank, 2003; Armantier & Boly, 2013).

A pesar de los interesantes resultados de Abbink y colaboradores (2002), es necesario hacer una importante salvedad respecto del procedimiento utilizado por los autores que dificulta la interpretación de sus resultados, particularmente del referido a la supuesta insensibilidad de los participantes respecto del daño a terceros. Y es que los terceros no fueron representados por jugadores pasivos (como en nuestro experimento), sino que los mismos pares Empresa-Funcionario Público participando de la misma sesión hacían las veces de terceros perjudicados. Por lo tanto, el desinterés por el bienestar de los terceros pudo estar confundido con una expresión de reciprocidad negativa (esto es, a los participantes puede no haberle importado perjudicar a los terceros porque pueden haber anticipado que algunos de

esos terceros podían asimismo afectarlos negativamente en sus pagos; Barr & Serra, 2009). A pesar de que puede argumentarse que la reciprocidad negativa involucra un elemento moral, la preocupación que exhiben los sujetos experimentales por terceras partes que no pueden afectar su propio pago es generalmente considerada un indicador claro de sensibilidad a las normas sociales (Fehr & Fischbacher, 2004). Por lo tanto, interpretar la ausencia de un efecto de tratamiento en la condición NE del experimento de Abbink y colaboradores, como un indicio de que los sujetos no poseen preferencias pro-sociales frente al daño a terceros podría ser un error.

En un estudio posterior, Abbink & Hennig-Schmidt (2006) realizaron un *JdC* similar, pero esta vez implementaron un marco cargado en el que en las instrucciones utilizaron un fraseo relacionado con corrupción (e.g., “empresa” en lugar de “jugador 1”, “funcionario público” en lugar de “jugador 2”, “pago privado” en lugar de “transferencia”, “otorgar licencia” en lugar de “opción Y”). Compararon esta condición con una condición abstracta. La hipótesis que plantearon consistía en que dado que la corrupción es un fenómeno que generalmente consideramos inmoral, el marco cargado resaltaría los aspectos negativos normalmente asociados al ofrecimiento y aceptación de coimas, y por ende debería disuadir a los sujetos de corromperse si exhibían preferencias sociales. Sin embargo, los resultados no arrojaron diferencias significativas de tratamiento en las decisiones de los sujetos.

Teniendo en cuenta el resultado negativo del experimento de Abbink y colaboradores, es razonable preguntarse si el *JdC* original utilizado por los autores es realmente capaz de permitir a los sujetos expresar las preferencias pro-sociales que presumiblemente modulan la corruptibilidad de las personas. En efecto, Barr & Serra (2009) propusieron una variante del *JdC*, enmarcada como una transacción entre un ciudadano y un funcionario público, que resultó parcialmente exitosa en mostrar la implicación de las preferencias sociales de los

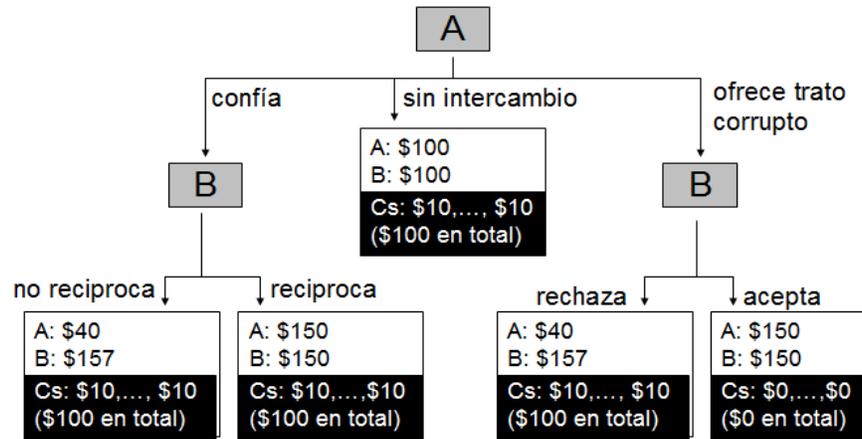
sujetos en el juego. Barr & Serra encontraron que el marco cargado disuadía a los participantes en mayor medida que el marco abstracto, aunque el efecto sólo estaba presente en los jugadores que iniciaban o que movían primero (es decir, los participantes en el rol de ciudadanos o potenciales oferentes de coimas). También encontraron que subir el nivel de las externalidades negativas causadas por la corrupción reducía el promedio de las transferencias de los ciudadanos en el marco cargado. Paradojalmente, los segundos jugadores resultaron ser menos propensos a aceptar la coima en el tratamiento con externalidades altas en el marco abstracto que en el cargado.

### 3. El experimento

#### 3.1 Nuestra variante del juego de coima

El árbol de decisión de nuestro juego consiste en tres ramas e involucra a dos jugadores activos, que a continuación llamaremos Participante A y Participante B. En el juego el Participante A mueve primero, mientras que B lo hace a continuación de A. También participan 10 jugadores pasivos como terceras partes (Participantes C), quienes pueden convertirse en las víctimas de la coima entre A y B (ver figura 1). Inicialmente, el primer jugador tiene tres opciones: 1) no realizar ninguna transacción (opción J); 2) jugar un Juego de Confianza (opción K); 3) jugar un juego de coima (opción L; i.e., un juego de confianza con externalidades negativas hacia las terceras partes). Por su parte, el jugador B puede elegir entre reciprocitar (opción N) o no reciprocitar (opción M) la confianza depositada por A en el Juego de Confianza (i.e. la rama K del árbol de decisión), o entre aceptar (opción P) o rechazar (opción O) el trato corrupto ofertado por A en la otra rama del árbol (i. e. opción L,

ver figura 1).



**Figura 1.** Árbol de decisión del Juego de Coima en nuestro experimento. Los participantes fueron asignados al azar a grupos independientes. Las instrucciones, incluyendo los nombres de las opciones del árbol de decisión, se referían a las opciones utilizando un lenguaje abstracto (J, K, L, M, N, O, y P) o un lenguaje cargado (sin transacción, confía, ofrece trato corrupto, no reciproca, reciproca, rechaza, y acepta). Los pagos que se muestran en la figura corresponden a los pagos reales utilizados en el primer escenario experimental de nuestro procedimiento. En los escenarios 2 a 5, los pagos asociados con las ramas J y K permanecieron fijos, mientras que los incentivos monetarios para los participantes A y B en la rama L se incrementaron a razón de 10% en cada escenario consecutivo. La elección de la opción P en la rama L (i.e. el participante A ofrece un trato corrupto y el participante B acepta) genera una externalidad negativa en todos los escenarios experimentales (i.e., los participantes C pierden su chance de ganar dinero).

Si A elige la opción J, el segundo jugador B no tiene la posibilidad de tomar ninguna decisión y el juego finaliza ahí. En este caso todos los jugadores A y B, y los participantes C reciben el mismo pago intermedio (para las terceras partes, el monto se reparte equitativamente entre los 10 participantes, es decir que reciben un 1/10 del pago

correspondiente a los participantes A y B). Si A elige la opción K (el Juego de Confianza) o la opción L (el Juego de Coima), el pago para B (independientemente de su elección) resulta mayor que en el caso de que A elija la opción J (se señaló este punto a los participantes para resaltar que el participante A le hacía un favor a B evitando la opción J). Nótese que las opciones K y L poseen potenciales riesgos para A, ya que en ambos casos, de las dos opciones disponibles para B (opciones M y N en la rama K, y opciones P y O), la opción dominante en términos de pagos para B es la que peor deja a A. Es claro que si A quiere evitar arriesgarse debe elegir la opción que le asegura un pago de 100. En efecto, en cada rama lateral, B puede optar por reciprocitar la confianza de A o aceptar el trato corrupto eligiendo N o P respectivamente, o puede maximizar de manera egoísta su propio pago eligiendo la opción M o la opción O (para B, las opciones M y O involucran siempre un pago 5% más alto que las opciones N y P respectivamente). Nótese que las opciones M y N dejan a los terceros con la misma ganancia que la que hubieran obtenido si no había transacción (opción J), lo que implica que en la rama K del árbol no hay externalidades negativas. En cambio, en la rama L, si el participante B acepta el trato propuesto por A (es decir, si elige la opción P), la ganancia de los participantes C pasa a ser nula (simulando la externalidad negativa asociada a la corrupción). Sin embargo, si el participante B rechaza el trato de corrupción (es decir elige la opción O), los participantes C conservan intactos sus pagos.

El lector podría preguntarse el motivo por el cual el Jugador A elegiría la opción L en lugar de la opción K, cuando ambas son idénticas en términos de pagos, mientras que la opción L presenta externalidades negativas hacia los terceros. En el experimento, sin embargo, a los participantes se les presentaba este mismo escenario en repetidas ocasiones pero aumentando los incentivos monetarios en la rama L. Dicho de otra manera, ofrecer un trato corrupto se volvía cada vez más tentador para los participantes a medida que pasaban

de escenario en escenario, mientras que la ganancia asociada a las opciones J y K permanecía igual. Por lo tanto, los participantes A se enfrentaban a una ponderación entre incentivos egoístas y preferencias pro-sociales (presumiblemente por no dañar a terceros). Por su parte, en la rama L los participantes B enfrentaban un compromiso entre reciprocidad al participante A y no dañar a los participantes C.

### *3. 2. Principales diferencias metodológicas entre nuestro JdC y versiones anteriores*

Nuestro experimento difiere en algunos puntos de los trabajos mencionados anteriormente, y por lo tanto es conveniente dedicar una sección para analizarlos.

En primer lugar, a diferencia de los JdC de Abbink y colaboradores y Barr & Serra, en nuestro procedimiento el Jugador A no transfiere dinero ni puntos experimentales al Jugador B. En su lugar, el Jugador A puede elegir entre jugar “seguro” (opción J) o tomar un riesgo eligiendo la opción que más favorece a B (y que más lo puede favorecer potencialmente a él también). En segundo lugar, y con el objetivo de simplificar el procedimiento, en nuestro experimento no implementamos ningún sistema de detección o castigo. Desde luego que esta simplificación acarrea el costo de no modelar una variable que en la realidad es ubicua; en este sentido, nuestra medida de corruptibilidad puede estar sobreestimando la corruptibilidad de los participantes, ya que aquellos que fueran propensos a entablar tratos corruptos en nuestro juego quizás se verían disuadidos de hacerlo si supieran que existe la posibilidad de ser detectados y castigados. De todas formas, en la vida real las personas muchas veces no conocen los detalles concretos de las sanciones por actos de corrupción, y nuestro juego puede interpretarse como simulando una interacción corrupta en un contexto de impunidad al castigo de terceros.

En tercer lugar, siguiendo a Barr & Serra, y a diferencia de Abbink y colaboradores, nuestro juego es un verdadero juego de una ronda (*one-shot*). Si bien los participantes jugaron en varios escenarios consecutivos, no recibieron ningún tipo de *feedback* entre escenarios, y se les dijo que el pago que recibirían al final de la sesión iba a estar determinado por las decisiones que hubieran tomado en un sólo escenario seleccionado a tal fin (ver más detalles de la forma en la que se prosiguió a pagarles a los participantes en la sección 5.3). La última diferencia que queremos destacar entre nuestro JdC y las versiones anteriores concierne al punto de si era estratégico para el Jugador 2 aceptar el trato corrupto y proveer el servicio. En el JdC original utilizado en Abbink y colaboradores (2002), la opción Y (i.e., la opción que favorece más a la Empresa) no era la opción maximizadora para el Funcionario Público en ninguna ronda; de todas formas, cuando el Funcionario Público elegía la opción Y reforzaba la confianza de la Empresa, lo que resultaba en transferencias más altas de la Empresa en rondas subsiguientes. Por lo tanto, la naturaleza repetida del juego hacía que los incentivos para el Funcionario Público estuvieran confundidos, ya que bien podía reciprocarse a la Empresa por satisfacer una preferencia de reciprocidad, o bien podía reciprocarse por razones meramente estratégicas (resultaba ser una buena estrategia a lo largo del juego reciprocarse a la Empresa si esta elevaba las transferencias). Por su parte, en el juego de Barr & Serra, si la transferencia del ciudadano superaba cierto umbral (el umbral correspondía a la transferencia que en caso de ser reciprocada culminaba en pagos equitativos para ambos jugadores), entonces reciprocarse resultaba la estrategia dominante para el Funcionario. En nuestro juego, por el contrario, el Jugador 2 siempre resulta mejor en términos monetarios si elige no reciprocarse y rechaza el trato.

En nuestra opinión este rasgo de nuestro juego modela adecuadamente la intención original de Abbink et al. (2002) de estudiar situaciones de coima donde pudiera desarrollarse

confianza y reciprocidad. En efecto, puede argumentarse que en muchas situaciones de corrupción menuda (*petty corruption*) el Funcionario a quien se le ofrece una coima puede maximizar su beneficio egoísta aceptando la coima, pero no proveyendo el servicio corrupto (esta posibilidad está vedada en el juego de Barr y Serra, pero puede tener expresión en el nuestro).

#### 4. Predicciones

El *marco* abstracto mide la valoración que hacen los individuos de las externalidades negativas, y en particular la solución de compromiso que encuentran, es decir, cómo ponderan entre los incentivos egoístas y las valoraciones pro-sociales. El *marco* cargado por su parte mide el impacto que tiene el fraseo cargado sobre el comportamiento de los participantes, en particular si dicho *fraseo* incide sobre la valoración que hacen los individuos de las externalidades hacia los terceros aumentando los costos morales por seguir la opción corrupta. Por lo tanto, comparando ambas condiciones podemos evaluar si el fraseo cargado modifica las propiedades motivacionales salientes del juego, promoviendo un cambio en las preferencias sociales de los individuos. Esto debería ser evidente en la proporción de participantes que eligen la opción corrupta en el *marco* abstracto versus el *marco* cargado. Esperamos que el *marco* cargado incremente el costo moral de seguir la opción corrupta, y por lo tanto que disminuya la proporción de participantes que elija la opción corrupta.

Consideremos en primer lugar el juego *one-shot*. En ambas condiciones, el juego es un verdadero juego *de una ronda* que se repite en todos los escenarios. En nuestra variante del JdC, suponiendo que ambos jugadores son racionales y egoístas y que además eso es conocimiento común, el equilibrio de Nash perfecto en sub-juegos es para el primer jugador

elegir la opción J (por sobre la opción K o L) y finalizar el juego allí. Esto es así, porque en caso de que elija la opción K o la opción L, el Jugador B tiene un incentivo egoísta para elegir la alternativa que menos favorece a A (la opción M en la rama K, o la opción O en la rama L), en cuyo caso A gana menos de lo que ganaría quedándose en J. Es más, el incentivo egoísta para el jugador B se incrementa a lo largo de los escenarios. Por lo tanto, como para cualquier juego de confianza standard (e.g., Berg, Dickhaut, & McCabe, 1995) la predicción juego-teórica anticipa que el Jugador B tiene incentivos egoístas para no reciprocarse la confianza del Jugador A (o aceptar el trato corrupto en nuestro procedimiento). Sabiendo esto, el Jugador A no debería confiar (u ofrecer el trato corrupto a B). Por lo tanto, el nivel de corrupción debería ser nulo, ya que A nunca debería elegir la opción que lo pone en riesgo.

Sin embargo, hay una extensa literatura que indica que una porción importante de los sujetos experimentales no es egoísta (ver Fehr & Fischbacher, 2003). De acuerdo con esta literatura, los sujetos exhiben una amplia gama de motivaciones sociales, entre ellas aversión a la inequidad (Fehr & Schmidt, 1999; Dawes et al., 2007), o una preferencia por reciprocidad basada en la intención (Rabin, 1993; Falk & Fischbacher, 2006). Los Jugadores A que sean aversos a la inequidad (la predicción vale también para los jugadores que sean aversos al riesgo) elegirán la opción J (no transacción), mientras que los Jugadores B que interpreten como “amable” la elección de la opción corrupta por parte de A (ya que aumenta sus pagos) elegirán con mayor probabilidad la opción P, esto es, reciprocarán. Pero es de esperar, como en otros juegos que presentan incentivos similares (e. g. Abbink et al., 2002), que los jugadores A anticipen que van a ser reciprocados por B y que por lo tanto confíen y elijan la opción L. En este caso, jugadores que no sean aversos al riesgo exhibirán una mayor propensión a elegir la opción corrupta. Por lo tanto, en consonancia con la literatura experimental sobre corrupción esperamos un nivel de corrupción no nulo en nuestro

experimento.

Una cuestión clave, entonces, reside en saber si el marco modifica las preferencias sociales de los participantes. Dado que hay trabajos previos que utilizando diferentes versiones del JdC han encontrado variaciones significativas en el comportamiento de los sujetos frente al *marco*, medir el nivel de Maquiavelismo de los sujetos nos permite examinar si los sujetos que obtienen puntajes diferentes en el cuestionario Mach IV responden de manera diferente asimismo al *marco*. Hay indicios que así lo sugieren. Quienes obtienen un puntaje alto en el cuestionario (llamados en la literatura *high-Mach*) tienden a exhibir la característica actitud del *Homo Economicus*, es decir, de búsqueda de la maximización de los pagos egoístas, independientemente del efecto de las acciones sobre el pago de los otros (Wilson, Near & Miller, 1996; Gunnthorsdottir, McCabe, & Smith, 2002). En línea con esto, aparentemente mayor maquiavelismo se asocia con una menor propensión a regirse por una moral deontológica (Bartels & Pizarro, 2011). Esto también es consistente con estudios que muestran que los *high-Mach* tienen una actitud flexible con respecto a las normas, y pueden seguirlas o violarlas conforme les resulte más conveniente (Meyer, 1998; Czibor & Bereczkei, 2012). Por ejemplo, Spitzer, Fischbacher, Herrnberger, Grön, & Fehr (2007) encontraron una asociación entre mayor Maquiavelismo y transferencias más bajas en un juego de Dictador. Sabido es que el monto que se transfiere en este tipo de juegos se toma como un indicador de una preferencia por la equidad y la justicia. En este sentido, esperamos encontrar una correlación positiva entre *Maquiavelismo* y la opción corrupta para los Jugadores A, tanto en el *marco* abstracto como en el *marco* en contexto.

El caso de los jugadores B es claro. Si los sujetos que puntúan alto en *maquiavelismo* se comportan como verdaderos *homini economici* entonces esperamos que los jugadores B no exhiban una preferencia social por reciprocitar la confianza de A, y den rienda suelta a su

egoísmo optando por la acción que maximiza sus pagos. Nótese que si A elige la opción corrupta, B siempre está mejor no reciprocando. Por lo tanto para los B no esperamos una correlación entre maquiavelismo y la opción corrupta, pero sí con la opción O (no reciprocación el trato corrupto), que es la estrategia dominante para B. Esta predicción vale tanto para el *marco* abstracto como para el *marco* cargado.

## 5. Procedimiento

El experimento fue realizado en la Universidad Nacional del Sur, Bahía Blanca, Argentina, en el mes de Noviembre de 2011. Los 127 participantes fueron reclutados asistiendo a un aula con previo consentimiento del profesor a cargo de la materia. Una vez en el aula, invitamos amablemente a los estudiantes presentes a participar en nuestro estudio de “toma de decisiones”, resaltando explícitamente que la participación era voluntaria, y que podían ganar dinero tomando parte del estudio. El experimento se realizó en dos sesiones diferentes, tomando parte de la primera 73 estudiantes de la carrera de Ingeniería Química, y en la segunda 54 estudiantes de Bioquímica. Todos los participantes realizaron la tarea experimental y completaron el test Mach IV. En el aula nos aseguramos de que hubiera espacio suficiente para que todas las decisiones y las respuestas al test se realizaran en privado. Cada sesión experimental siguió el orden siguiente: 1) un asistente introdujo el objetivo del experimento a los participantes, explicó un escenario a modo de ejemplo, y mencionó asimismo cómo iban a determinarse los pagos al final de la sesión; 2) los cuadernillos fueron distribuidos y se corrió el experimento; y 3) una vez que todos los participantes devolvieron los cuadernillos a los asistentes, realizamos los pagos correspondientes. A continuación explicamos en mayor detalle las tres partes del

experimento.

### *5.1. La explicación pública del asistente*

Al comenzar la sesión, uno de los asistentes explicó públicamente a los estudiantes en qué consistía el experimento. El asistente les explicaba que estaban participando de un estudio de toma de decisiones en el que algunos de ellos iban a poder ganar dinero, y que lo que pudieran ganar dependería de las decisiones que tomaran. Asimismo se hizo hincapié en que no había respuestas correctas o incorrectas, y que el objetivo del estudio consistía en medir preferencias individuales. A continuación se dibujó a modo de ejemplo un árbol de decisión del juego en el pizarrón, y se explicaron los diferentes roles del juego así como las consecuencias de cada una de las posibles acciones. A los participantes se les comunicó que cada uno participaría en uno de los tres posibles roles, A, B, o C, y que los roles quedarían determinados al azar de acuerdo al cuadernillo que recibieran.

Por último, antes de entregar los cuadernillos, el asistente explicó de qué manera se iban a realizar los pagos reales al final de la sesión (véase más abajo). Es importante destacar que se les aseguró a los participantes que sus decisiones y el pago potencial permanecerían anónimos. Una vez que se terminó de impartir las instrucciones, se entregaron los cuadernillos, y se les pidió a los participantes que permanecieran en silencio hasta que la sesión finalizara.

### *5.2. Contenido de los cuadernillos.*

Los cuadernillos para los participantes A y B consistían en ocho páginas escritas en español. En las dos primeras páginas, los participantes encontraban una introducción, un

ejemplo de un escenario de decisión, y la información sobre su propio rol. La página 3 consistía de cuatro preguntas destinadas a evaluar la comprensión de los participantes de las consecuencias de los pagos de las posibles decisiones en el escenario ejemplo. Las páginas 4 a 6 incluían los escenarios de decisión 1 a 7 (cinco escenarios experimentales seguidos de dos escenarios control). Por último, las páginas 7 y 8 contenían una versión en español del cuestionario de Maquiavelismo Mach IV (Christie & Geis, 1970, ver apéndice 1).

Escenarios experimentales: Antes de presentar los escenarios, los participantes podían leer la frase “Ahora es tiempo de que tomes tus decisiones” al comienzo de la página. Como se puede ver en la Figura 1, cada escenario experimental se compone de tres alternativas para los participantes A (opciones J, K, o L), y dos alternativas de decisión para los participantes B (opciones M y N en la rama K, y O y P en la rama L). En la Figura 1 también presentamos los pagos asociados a cada combinación de decisiones del Escenario 1 tal como fue presentado a los participantes. A diferencia del Escenario 1, de los escenarios 2 al 5, incrementamos sistemáticamente en 10% en cada escenario la ganancia que podían obtener los participantes en la rama L del árbol. Si el trato corrupto se consumaba, es decir, si A elegía ofrecer el trato y B aceptaba, entonces los participantes A y B obtenían la misma ganancia: en el escenario 2, \$165; en el escenario 3, \$181; en el escenario 4, \$199; en el escenario 5, \$219; la ganancia de los terceros en este caso era siempre \$0. En todos los escenarios experimentales, si el participante B no reciprocaba en la rama L (i. e., si elegía la opción O en lugar de P), el participante A ganaba \$40, los participantes C ganaban \$100 en total, y la ganancia de B era siempre 5% más alta que en la opción alternativa P (i. e., en el escenario 2, \$173; en el escenario 3, \$190; en el escenario 4, \$209; y en el escenario 5, \$230. Con esta opción modelamos el incentivo de B de aceptar la coima, pero no proveer el servicio corrupto, y el correspondiente riesgo que representa para A confiar en B). Nótese que los pagos

asociados a las ramas J y K no varían entre escenarios.

Escenarios de control: Teniendo escenarios repetidos, podemos medir la frecuencia con que los participantes eligen la opción corrupta (i. e., la cantidad de escenarios en los que eligen la opción con externalidades negativas) y el nivel mínimo de incentivos en el que los participantes eligen ofrecer el trato corrupto o aceptarlo. Sin, embargo, la interpretación de estas variables presenta dos ambigüedades que necesitan ser controladas. Primero, si un participante A eligió la opción J en todos los escenarios, del 1 al 5, podría parecer que lo hiciera porque es la decisión “moralmente correcta”. Sin embargo, su motivación puede no estar relacionada con la moralidad, sino más bien con la aversión al riesgo. Para desambiguar este patrón de decisiones, en el escenario control 6 modificamos los incentivos de tal manera que la opción que generaba externalidades (un daño a los participantes C) también dejaba de ser riesgosa para A (opción K en escenario 6; ver Apéndice 2 para más detalles). Por lo tanto, si un participante A eligió la opción J de los escenarios 1 al 5, y además evitó la opción que dañaba a los terceros en el escenario 6), computamos un 0 en su frecuencia de coima y establecimos su incentivo mínimo de coima por encima del nivel capturado en el experimento. En cambio, si el Participante A eligió la opción J de los escenarios 1 a 5, pero luego cambió a la opción K en el escenario 6, entonces lo consideramos excluible de los análisis de la correlación entre los puntajes de maquiavelismo y las decisiones dado que interpretamos que en su comportamiento el motivo predominante no es la preferencia social por no dañar a terceros, sino meramente la aversión al riesgo.

La segunda ambigüedad por la que queríamos controlar ocurre cuando el participante B eligió siempre la opción O por sobre la opción P. Para considerar que ese patrón de decisiones se debe a una preferencia por no dañar a terceros, tenemos que descartar la posibilidad de que en realidad el participante B estuviera simplemente maximizando sus

ganancias. Con este objetivo en mente, diseñamos el escenario control 7 (ver apéndice 2), en el que el participante B tenía la opción de maximizar su propia ganancia (opción N), o evitar dañar a las terceras partes a costa de un sacrificio de 5% en su ganancia (opción M). Entonces, si un participante B eligió la opción O de los escenarios 1 a 5, y a su vez evitó dañar a los terceros en el escenario 7, computamos 0 en su frecuencia de coima y establecimos su nivel de incentivo mínimo por encima del nivel máximo medido en el experimento. En cambio, si el participante B eligió la opción O de los escenarios 1 a 5, pero luego cambió a la opción con externalidades para los terceros en el escenario 7, entonces lo consideramos excluible de los análisis de correlación entre las decisiones y los puntajes de maquiavelismo porque interpretamos que el componente que motivaba sus decisiones era la maximización de la ganancia y no un interés por no dañar a terceros.

Mach test: A los participantes A y B se les presentó una traducción al español en dos páginas del *Mach Test* que tuvieron que completar a continuación del último escenario. El Mach Test consiste en 20 enunciados, como por ejemplo, “Uno debe actuar sólo cuando hacerlo es moralmente correcto” (“*One should take action only when sure it is morally right*”) o “Quien confía completamente en cualquiera, busca problemas” (“*Anyone who completely trusts anyone else is asking for trouble*”). Los participantes deben expresar su nivel de acuerdo con cada enunciado en una escala de Likert de 5 puntos que va de “totalmente de acuerdo” a “en total desacuerdo” (ver Apéndice 1).

### 5.3. Pago.

Una vez que todos los participantes completaron sus cuadernillos y los devolvieron al asistente, seleccionamos al azar a dos participantes, uno en el Rol A y otro en el Rol B. A

continuación seleccionamos uno de los 7 escenarios a partir del cual calculamos los pagos que correspondían a cada participante. Para asegurarnos completa anonimidad en las elecciones y en los pagos, realizamos el siguiente procedimiento: a) los 12 participantes que recibieron un pago monetario real (esto es, un participante A, un participante B y los 10 participantes C) fueron convocados individualmente a un aula diferente de la que se realizó la sesión; b) en el caso de los participantes A y B, le pedimos al participante A (B) seleccionado que eligiera al azar un cuadernillo de un participante B (A) (cuya identidad no era revelada) de entre todos los cuadernillos de los participantes B (A) y su pago se determinó de acuerdo a las elecciones que realizaron ambos en el escenario seleccionado; y c) en el caso de los 10 participantes C, cada uno debía seleccionar al azar dos cuadernillos, uno correspondiente a un participante A y otro a un participante B (cuyas identidades no eran reveladas), y sus pagos se determinaban a partir de las elecciones que habían realizado dichos participantes en el escenario seleccionado.

#### 5.4. Marco

A diferencia de otros estudios, en los cuales los juegos eran enmarcados como transacciones entre empresas o ciudadanos y funcionarios públicos (Abbink y Hennig-Schmidt, 2006; Barr y Serra, 2009), utilizamos un marco de etiquetas (*Label framing*) (ver Dufwenberg et al., 2011; Elliott et al. 1998, donde lo llaman un “*pure framing effect*”), y por lo tanto evitamos cualquier referencia a roles que puedan parecer artificiales a estudiantes de grado. Los participantes fueron asignados a uno u otro de los tratamientos de acuerdo al cuadernillo que recibieron: un tratamiento con marco abstracto, n=53; y un tratamiento con un marco cargado, n=54. En el marco abstracto, en el árbol de decisión los participantes

tenían opciones con nombres abstractos como J, K, L, M, N, O y P. En cambio, en el tratamiento cargado las opciones tenían nombres cargados: la opción J como “sin transacción”, K como “confía”, opción L como “ofrece trato corrupto”, opción M como “no reciproca”, opción N como “reciproca”, opción O como “rechaza”, y opción P como “acepta”. Excepto por los nombres asociados a cada opción de juego, el resto de las instrucciones de juego y los escenarios, así como los pagos asociados eran idénticos entre ambos tratamientos.

### *5.5. Variables dependientes*

#### *Participantes A*

En cada escenario experimental registramos la decisión tomada por los participantes A entre las opciones J, K y L. Definimos la variable “frecuencia de coima”<sup>2</sup> sumando la cantidad de veces que el participante eligió la opción L a lo largo de todos los escenarios. Ya que el experimento consta de 5 escenarios, “frecuencia de coima” varía de 0 a 5. Además, para capturar el efecto del incremento del incentivo monetario egoísta en la rama izquierda del árbol, definimos una variable que llamamos “mínimo incentivo de coima” (MIC) y que indica el primer escenario en el que el participante A escogió la opción L. MIC tiene un rango de 1 a 6 (donde 1 significa que el participante escogió la opción corrupta en el escenario 1 mientras que 6 indica, por el contrario, que el participante no eligió la opción corrupta en ninguno de los escenarios).

---

<sup>2</sup> A lo largo del texto usamos las expresiones “frecuencia de coima” y “frecuencia de trato corrupto” de manera indistinta. Lo mismo con las expresiones “aceptación de coima” y “aceptación de trato corrupto”.

### *Participantes B*

En el caso de los participantes B, registramos dos decisiones por cada escenario experimental (M y N en la rama K; O y P en la rama L). Calculamos la variable “frecuencia de coima” para cada participante B sumando las veces que B optó por P en los 5 escenarios. Asimismo calculamos la variable “mínimo incentivo de coima” (MIC), que indica el primer escenario en el que el participante B escogió aceptar el trato corrupto (opción P). MIC tiene un rango de 1 a 6 (donde 6 significa que el participante nunca aceptó la opción corrupta).

## **6. Análisis estadísticos**

Reportamos los análisis estadísticos de manera independiente para los jugadores A y los jugadores B. Las variables dependientes “frecuencia de coima” y “mínimo incentivo de coima” resultaron tener una distribución no-normal de acuerdo al *Test de Normalidad de Shapiro-Wilk*, y por lo tanto utilizamos estadísticas no-paramétricas para realizar los análisis. Utilizamos el *test U de Mann-Whitney* para comparar las respuestas de los participantes entre tratamientos (abstracto versus “en contexto”). También recurrimos al *test exacto de Fisher* para evaluar si el número de participantes que eligió la coima (opción L para los jugadores A y opción P para los jugadores B) difirió entre tratamientos en cada escenario. Para estimar si el número de participantes que eligió la coima se modificó a lo largo de los escenarios usamos el *test Q de Cochran* para los participantes A y B en ambos tratamientos. A su vez, utilizamos el test Chi-cuadrado para evaluar si el número de participantes que eligió la opción corrupta difirió del azar en cada escenario (para los participantes A, en lugar de computar 33% como el nivel de azar para elegir la opción corrupta, comparamos la frecuencia con la cual eligieron la opción L con la suma de las frecuencias de la opción L más la opción K en

un escenario dado, esto es, sólo consideramos las dos opciones que suponen un riesgo para A, pero que varían en términos de sus externalidades; para los participantes B simplemente comparamos la frecuencia con que eligieron la opción corrupta contra el 50% de azar de elegir la opción P u O). Los análisis para estimar la diferencia en la frecuencia con que los participantes B eligieron la opción de coima P y la opción N (es decir, que reciprocaron cuando A eligió K) se basaron en el *Wilcoxon Matched Pairs test*. Por último, utilizamos *Spearman Rank Order correlations* para evaluar la asociación entre los puntajes del cuestionario Mach y las variables dependientes definidas más arriba.

## 7. Resultados y discusión

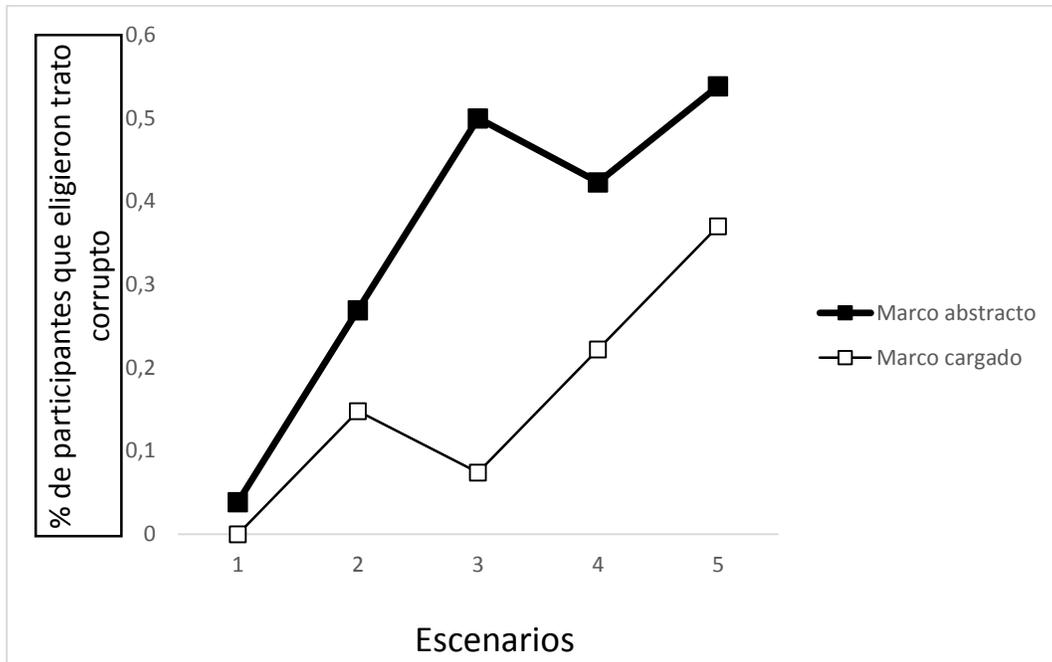
Ninguno de los participantes respondió de manera errónea a más de una respuesta; por lo tanto no excluimos a ningún participante por esta razón.

### 7. 1. *Participantes A*

Cuatro participantes A en el marco abstracto y uno en el marco cargado eligieron la opción J (i.e., sin transacción) del escenario 1 al 5, y luego eligieron la opción con externalidades negativas en el escenario 6, cuando no había ningún riesgo involucrado. Excluimos a estos participantes de los análisis, de acuerdo al criterio del escenario control 6. Por lo tanto reportamos los resultados incluyendo estos participantes (marco abstracto,  $n=26$ ; marco cargado,  $n=27$ ) y excluyéndolos (marco abstracto,  $n=22$ ; marco cargado,  $n=22$ ).

La figura 2 muestra el porcentaje de participantes A (sin excluir de acuerdo al criterio del escenario control 6) que eligieron la opción L (i.e., la opción de coima) en función del escenario experimental y el *marco*. Considerando los 5 escenarios en conjunto, el promedio de la frecuencia de coima fue casi el doble en el marco abstracto que en el marco cargado

(media $\pm$ 1 SE: 2.65  $\pm$ 0.33 en el marco abstracto, y 1.33 $\pm$ 0.23 en el marco cargado; *Mann-Whitney U test*,  $U_{26,27}=195$ ,  $Z=2.83$ ,  $p=0.005$ ). Complementariamente, la frecuencia de las elecciones de la opción K (esto es, el Juego de Confianza sin externalidades negativas) fue más alta en el marco cargado que en el abstracto (media $\pm$ 1 SE: 1.12  $\pm$ 0.22 en el marco abstracto, y 2.15 $\pm$ 0.32 en el marco cargado; *Mann-Whitney U test*,  $U_{26,27}=222$ ,  $Z=2.29$ ,  $p < 0.05$ ). En cambio, la frecuencia de la elección de la opción J no difirió entre tratamientos (media $\pm$ 1 SE: 2.12  $\pm$ 0.31 en el marco abstracto, y 1.96 $\pm$ 0.30 en el marco cargado; *Mann-Whitney U test*,  $U_{26,27}=338.50$ ,  $Z=0.22$ ,  $p=0.82$ ). Estas diferencias en la frecuencia con que los participantes eligieron las opciones L, K o J, se mantuvo cuando repetimos los análisis excluyendo a los participantes de acuerdo al criterio del escenario control 6 (frecuencia de la opción L: *Mann-Whitney U test*,  $U_{22,26}=131.50$ ,  $Z=3.20$ ,  $p < 0.05$ ; frecuencia de la opción K: *Mann-Whitney U test*,  $U_{22,26}=191.50$ ,  $Z=2.02$ ,  $p < 0.05$ ; frecuencia de la opción J: *Mann-Whitney U test*,  $U_{22,26}=191.50$ ,  $Z=2.02$ ,  $p=0.82$ ). Este resultado muestra claramente que el uso de un fraseo cargado que contiene palabras que hacen referencia a la corrupción, disuadió efectivamente a los participantes A de elegir la opción que incluía externalidades negativas hacia terceros.



**Figura 2.** Porcentaje de participantes A que eligieron la opción L en el marco abstracto ( $n = 26$ ) o que eligieron “ofrece trato corrupto” en el marco cargado ( $n = 27$ ). El eje horizontal representa los cinco escenarios de decisión. El incentivo monetario asociado a la opción L y a la opción “ofrece trato corrupto” se aumentó 10% en cada escenario, mientras que los pagos en las opciones alternativas permanecieron fijos.

También evaluamos la diferencia entre tratamientos comparando la frecuencia con que los participantes eligieron la opción L (la opción de coima) en cada escenario (ver figura 2). Encontramos que, sin excluir participantes, las diferencias de tratamiento son significativas en el escenario 3 (*Fischer's exact test*,  $p < 0.001$ ; escenario 1,  $p = 0.49$ ; escenario 2,  $p = 0.34$ ; escenario 4,  $p = 0.10$ ; escenario 5,  $p = 0.25$ ), mientras que excluyendo participantes de acuerdo al criterio fijado con el escenario control 6, también surgen diferencias significativas entre tratamientos en el escenario 4 (*Fischer's exact test*,  $p = 0.05$ ). Es más, el número de participantes que eligió la opción de coima resultó estar significativamente por debajo del nivel de azar en los escenarios 1 y 3 del marco cargado (*Chi-square tests*: en el

escenario 1, 0 participantes eligió la opción L de los 13 que eligieron o la opción K o L,  $\chi^2=7.80$ ,  $gl=1$ ,  $p < 0.01$ ; en el escenario 3, 2 participantes de 18 eligieron la opción L,  $\chi^2=6.41$ ,  $gl=1$ ,  $p < 0.05$ ; en el escenario 2, 5 de 15 eligieron la opción L,  $\chi^2=7.80$ ,  $gl=1$ ,  $p=0.46$ ; en el escenario 4, 6 de 14 eligieron la opción L,  $\chi^2=0.56$ ,  $gl=1$ ,  $p=0.46$ ; en el escenario 4, 6 de 14 eligieron la opción L,  $\chi^2=0.14$ ,  $gl=1$ ,  $p < 0.70$ ; y en el escenario 5, 11 de 19 escogió la opción L,  $\chi^2=0.09$ ,  $gl=1$ ,  $p < 0.76$ ). El número de participantes que eligió la opción de coima en el *marco* abstracto resultó marginalmente por debajo del nivel de azar en el escenario 1 (1 de 11,  $\chi^2=3.67$ ,  $gl=1$ ,  $p=0.055$ ), y marginalmente por encima del nivel de azar en el escenario 4 (11 de 13,  $\chi^2=2.89$ ,  $gl=1$ ,  $p=0.09$ ) y en el escenario 5 (14 de 17,  $\chi^2=3.36$ ,  $gl=1$ ,  $p=0.067$ ). Es importante destacar que los participantes evitaron la opción que generaba externalidades negativas para los participantes C cuando había una opción equivalente en términos de pagos, pero que no generaba externalidades (i. e., en el escenario 1). Este punto sugiere que en ambos tratamientos los sujetos no fueron indiferentes al bienestar de las terceras partes. Más aún, en el *marco cargado* los participantes A prefirieron la opción sin externalidades negativas (i. e., opción K) en el escenario 3, aun cuando la opción de coima en el mismo escenario presentaba un riesgo similar pero ofrecía mayores incentivos egoístas en términos monetarios. Esto sugiere que el efecto disuasivo del fraseo cargado puede haber funcionado incrementando los costos morales por elegir el “trato corrupto”.

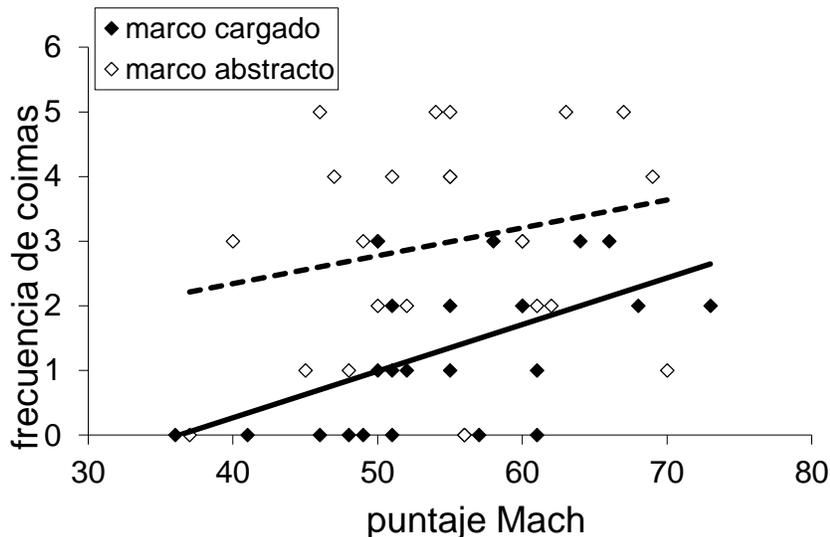
En términos de los incentivos mínimos de coima, cuando excluimos participantes en función del criterio establecido por el escenario control 6, encontramos que en general los participantes A eligieron *antes* la opción con externalidades negativas en el *marco* abstracto que en el cargado (media  $\pm$  1 SE: 3.45  $\pm$  0.36 en el marco abstracto, y 4.53  $\pm$  0.31 en el marco cargado; *Mann-Whitney U test*,  $U_{26,27}=185.50$ ,  $Z=2.08$ ,  $p < 0.05$ ). Este resultado también

indica que los participantes A ponderaban los incentivos monetarios egoístas y las preferencias pro-sociales a la hora de tomar una decisión, y que el peso relativo dado a las últimas se incrementó con el fraseo cargado de las opciones.

Los análisis usando el *Cochran Q test* para evaluar la variación en la propensión de los participantes a elegir la opción L a lo largo de los escenarios mostraron efectos significativos en ambos tratamientos (*marco* abstracto:  $\chi^2_Q=26.86$ ,  $gl=4$ ,  $n=26$ ,  $p < 0.001$ , *marco* cargado:  $\chi^2_Q=18.63$ ,  $gl=4$ ,  $n=27$ ,  $p < 0.001$ ). También comparamos los escenarios de a pares con el objetivo de identificar si los efectos de *marco* diferían entre escenarios. En el tratamiento abstracto, el número de participantes que eligió la opción L difirió significativamente entre los escenarios 1 y 2 ( $\chi^2_Q=4.50$ ,  $gl=1$ ,  $n=26$ ,  $p < 0.05$ ) y entre los escenarios 4 y 5 ( $\chi^2_Q=12.00$ ,  $gl=1$ ,  $n=26$ ,  $p < 0.001$ ); el cambio del escenario 2 al 3 en el número de participantes que eligió L resultó sólo marginalmente significativo ( $\chi^2_Q=3.60$ ,  $gl=1$ ,  $n=26$ ,  $p = 0.058$ ; además del escenario 3 al 4, ( $\chi^2_Q=1.80$ ,  $gl=1$ ,  $n=26$ ,  $p = 0.18$ ). En el *marco* cargado, la variación en el número de participantes que eligió la opción L también resultó significativa del escenario 1 al 2 ( $\chi^2_Q=5.00$ ,  $gl=1$ ,  $n=27$ ,  $p < 0.05$ ) y del escenario 4 al 5 ( $\chi^2_Q=6.00$ ,  $gl=1$ ,  $n=27$ ,  $p < 0.05$ ). La frecuencia de la elección de L no difirió del escenario 2 al 3, y tampoco del 4 al 5 ( $\chi^2_Q=1.28$ ,  $gl=1$ ,  $n=27$ ,  $p =0.26$ ; ( $\chi^2_Q=3.57$ ,  $gl=1$ ,  $n=27$ ,  $p =0.059$ , respectivamente). Los datos y los análisis en general confirman que los participantes tienden a comportarse más en consonancia con la predicción de la teoría de juegos a medida que los escenarios aumentan los incentivos monetarios egoístas. Como puede verse en la figura 2, el patrón indica que el número de participantes que elige la opción corrupta crece a medida que aumenta la ganancia que genera la corrupción.

Por último, evaluamos la asociación de los puntajes del Mach test con las variables *frecuencia de coima* e *incentivo mínimo de coima*. Realizamos los análisis excluyendo los participantes de acuerdo al criterio del escenario control 6; ver figura 3). Es importante destacar que los puntajes de maquiavelismo no difirieron entre tratamientos (*Mann-Whitney U test*,  $U_{22,26}=268$ ,  $Z=0.37$ ,  $p=0.72$ ). En el marco abstracto, no resultaron significativas ni la asociación de maquiavelismo con la *frecuencia de coima* ( $\rho=0.19$ ,  $n=22$ ,  $p=0.39$ ) ni la asociación de maquiavelismo con el *incentivo mínimo de coima* ( $\rho=-0.13$ ,  $n=22$ ,  $p=0.55$ ). En contraste, tanto la asociación del puntaje de *Mach* con la *frecuencia de coima* y la asociación del puntaje de *Mach* con el *incentivo mínimo de coima* resultaron significativas en el marco cargado ( $\rho=0.47$ ,  $n=26$ ,  $p < 0.05$ ;  $\rho=-0.41$ ,  $n=26$ ,  $p < 0.05$  respectivamente). Las asociaciones del puntaje de *Mach* con la frecuencia de las otras opciones no resultaron significativas (en el marco abstracto: asociación entre *Mach* y opción K,  $\rho=-0.08$ ,  $n=22$ ,  $p=0.72$ ; asociación entre *Mach* y opción J,  $\rho=-0.06$ ,  $n=22$ ,  $p=0.80$ ); en el marco cargado: asociación entre *Mach* y opción K,  $\rho=-0.01$ ,  $n=26$ ,  $p=0.94$ ; asociación entre *Mach* y opción J,  $\rho=-0.12$ ,  $n=22$ ,  $p=0.55$ ). Los análisis correlacionales precedentes sugieren que las respuestas a nuestro juego pueden capturar diferencias individuales en las actitudes morales de los participantes, aunque sólo en sus respuestas al marco *cargado*. Sucede lo contrario cuando las opciones son presentadas de manera abstracta, lo que indicaría que las motivaciones morales serían menos evidentes en las decisiones en el marco abstracto. Esta conclusión, sin embargo, colisiona con el hecho de que los jugadores A en ambos tratamientos se abstuvieron de elegir la opción de coima cuando había una alternativa equivalente en los pagos que no generaba externalidades. Esta inconsistencia aparente puede ser resultado de que el Mach test no mida la misma dimensión moral que nuestro JdC. Los enunciados del test son predominantemente sobre honestidad, y reflejan una tendencia al

cinismo en las relaciones interpersonales, y por lo tanto tiene sentido que los puntajes que refleja el cuestionario se correlacionen con las respuestas de los participantes cuando se presentaron las opciones del juego con términos tales como “trato corrupto”, que pueden estar semánticamente asociados al concepto de honestidad. Además puede ser el caso de que influir negativamente en los pagos de los otros jugadores no tenga relación con la honestidad de los individuos. De hecho, parece claro que dimensiones como el daño y la justicia son dimensiones distintas de la moralidad (Haidt, 2007). Por lo tanto, si relacionamos directamente Maquiavelismo y justicia (e. g., ver Gunnthorsdottir, 2002), podemos tener una premisa a partir de la cual entender la falta de asociación entre el puntaje en el test y las respuestas de los sujetos en el marco abstracto.



**Figura 3.** Correlación entre puntajes de maquiavelismo y frecuencia de coima en los participantes A en función de los tratamientos de marco (marco abstracto,  $n = 22$ ; marco cargado,  $n = 26$ ). La línea sólida representa una regresión lineal significativa en el marco cargado ( $\rho=0.47$ ,  $p < 0.05$ ), mientras que la línea punteada representa una regresión lineal no significativa en el marco abstracto ( $\rho=0.19$ ,  $p < 0.05$ ).

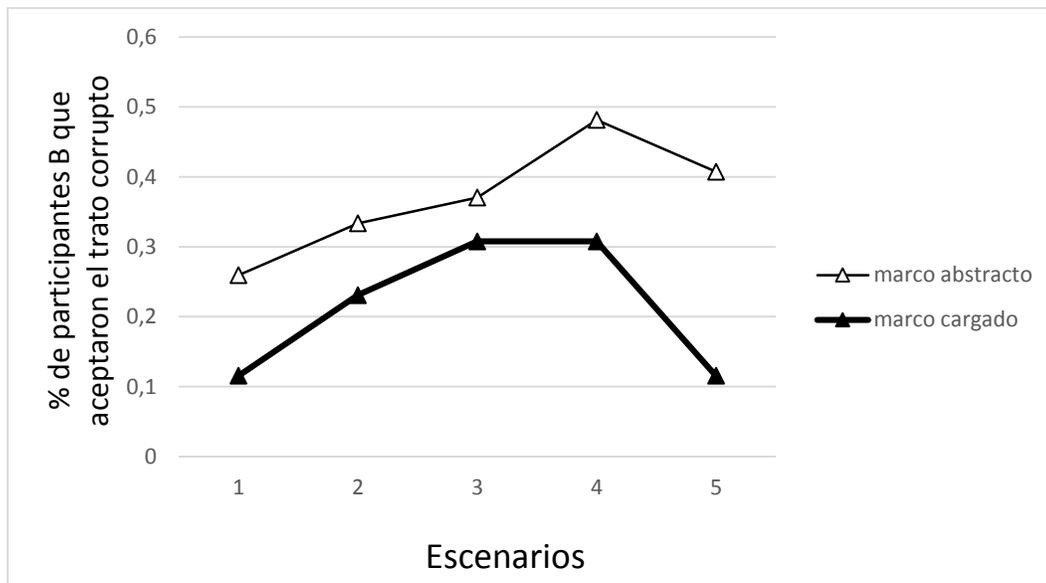
## 7.2. Participantes B

En el marco cargado un participante respondió de manera incompleta a la tarea experimental, por lo tanto fue descartado para los análisis. Asimismo, de acuerdo al criterio de control del escenario 7, cuatro participantes pueden ser excluidos en el marco abstracto y dos en el marco cargado. Por lo tanto reportamos resultados con (marco abstracto, n=27; marco cargado, n=26) y sin (marco abstracto, n=23; marco cargado, n=24) estos participantes.

La figura 4 muestra el porcentaje de participantes B (sin excluir participantes) que eligieron la opción P (o aceptar el trato corrupto) en función del escenario y el tratamiento. Al igual que sucedía con los participantes A, la frecuencia de aceptación del trato corrupto (elección de P en la rama L) en los cinco escenarios considerados conjuntamente, resultó ser casi el doble en el marco abstracto (media $\pm$ 1 ES: 2.48  $\pm$ 0.42 en el marco abstracto, y 1.26 $\pm$ 0.32 en el marco cargado; *Mann-Whitney U test*,  $U_{26,27}=235.50$ ,  $Z=2.05$ ,  $p < 0.05$ ). A su vez, la frecuencia con la que los B reciprocaron en la rama sin externalidades no resultó ser significativamente diferente entre tratamientos (media $\pm$ 1 ES: 2.89  $\pm$ 0.44 en el marco abstracto, y 3.19 $\pm$ 0.42 en el marco cargado; *Mann-Whitney U test*,  $U_{26,27}=312$ ,  $Z=-0.69$ ,  $p = 0.50$ ). Resulta interesante que entre las opciones P y N, es decir, entre reciprocación con y sin externalidades negativas, hayamos encontrado una correlación negativa en el marco cargado (*Spearman Rank Order Correlation*,  $\rho=-0.45$ ,  $p < 0.05$ ), pero no en el marco abstracto ( $\rho=0.19$ ,  $p=0.34$ ). Es más, en el marco cargado la frecuencia de elecciones de N (i. e., reciprocación en la rama sin externalidades negativas) resultó ser más del doble que la frecuencia de elecciones de P (i. e., reciprocación en la rama con externalidades negativas) en el marco cargado (*Wilcoxon Matched Pairs*,  $Z=2.97$ ,  $n=26$ ,  $p < 0.01$ ). Este patrón no se repitió, sin

embargo, en el marco abstracto ( $Z=2.97$ ,  $n=26$ ,  $p = 0.41$ ). Los resultados excluyendo los participantes de acuerdo al criterio del escenario control 7 no se modifican sustancialmente, por lo tanto decidimos no reportarlos.

Tal como sucedía en el caso de los participantes A, el marco cargado disuadió a los participantes B de aceptar el trato corrupto. La correlación negativa entre corrupción y reciprocidad (sin externalidades) en el marco cargado pero no en el abstracto sugiere que presentar las opciones con un fraseo cargado influyó decisivamente en las decisiones de los participantes.



**Figura 4.** Porcentaje de participantes B que aceptaron el trato corrupto graficado en función de los escenarios y los tratamientos. En el marco abstracto la aceptación del trato corrupto es la “opción P”, mientras que en el marco cargado, llamamos a esa opción “aceptar (el trato corrupto)”. En los escenarios de 1 a 5, incrementamos de a 10% por escenario el incentivo monetario para aceptar el trato, mientras que la ganancia en esa opción era siempre 5% más alta que la asociada a la opción alternativa (opción O u opción “rechazar (trato corrupto)”.

Asimismo evaluamos la diferencia entre tratamientos en el número de participantes

que aceptaron el trato en cada escenario (ver figura 4). Encontramos que, excluyendo o sin excluir participantes, las diferencias de tratamiento son significativas en el escenario 5 (sin excluir participantes, *Fischer's exact test*,  $p < 0.05$ ; escenario 1,  $p=0.16$ ; escenario 2,  $p=0.30$ ; escenario 3,  $p=0.42$ ; escenario 4,  $p=0.16$ ; mientras que excluyendo participantes de acuerdo al criterio fijado con el escenario control 7, las diferencias más importantes entre tratamientos surgen en el escenario 4,  $p=0.09$ , y en el escenario 5,  $p < 0.01$ ). Es más, el número de participantes que eligió la opción “aceptar” resultó estar significativamente por debajo del nivel de azar en los escenarios 1, 2 y 5 del marco cargado (*Chi-square tests*: en el escenario 1, 3 participantes de 26 eligieron la opción P,  $\chi^2=9.03$ ,  $gl=1$ ,  $p < 0.01$ ; en el escenario 2, 6 de 26 participantes eligieron la opción P,  $\chi^2=4.06$ ,  $gl=1$ ,  $p < 0.05$ ; en el escenario 5, 3 de 26 eligieron la opción P,  $\chi^2=9.03$ ,  $gl=1$ ,  $p < 0.01$ ; en los escenarios 3 y 4, 8 de 26 eligieron la opción P,  $\chi^2=2.00$ ,  $gl=1$ ,  $p=0.16$ ). Por su parte, en el marco abstracto la aceptación del trato corrupto no difirió del azar (en el escenario 1, 7 de 27 eligieron P,  $\chi^2=2.86$ ,  $gl=1$ ,  $p=0.09$ ; en el escenario 2, 9 de 27 eligieron P,  $\chi^2=1.23$ ,  $gl=1$ ,  $p=0.27$ ; en el escenario 3, 10 de 27 eligieron P,  $\chi^2=0.68$ ,  $gl=1$ ,  $p=0.41$ ; en el escenario 4, 13 de 27 eligieron P,  $\chi^2=0.00$ ,  $gl=1$ ,  $p=1.00$ ; y en el escenario 5, 11 de 27 eligieron P,  $\chi^2 = 0.30$ ,  $gl=1$ ,  $p=0.58$ ). Los resultados proveen respaldo otra vez a la idea de que las responsables de disuadir a los participantes no fueron las externalidades negativas *per se*, sino el uso de instrucciones cargadas con términos asociados a corrupción.

En términos de los incentivos mínimos por aceptar el trato corrupto, encontramos que los Participantes B eligieron aceptar el trato *antes* en el marco abstracto que en el cargado. Sin embargo, esta diferencia sólo es marginalmente significativa si excluimos participantes en función del criterio establecido por el escenario control 7 (excluyendo participantes,  $\text{media} \pm 1 \text{ ES}$ :  $3.17 \pm 0.43$  en el marco abstracto, y  $4.25 \pm 0.42$  en el marco cargado; *Mann-*

*Whitney U test*,  $U_{23,24}=191.50$ ,  $Z=1.80$ ,  $p=0.07$ ; sin excluir participantes, *Mann-Whitney U test*,  $U_{26,27}=271.50$ ,  $Z=1.41$ ,  $p=0.16$  ). El hecho de que los participantes B aplazaran la aceptación del trato a escenarios posteriores en el marco cargado indica claramente que el fraseo cargado de las opciones incrementó el costo moral por aceptar el trato.

Los análisis usando Cochran Q test para evaluar la variación en el número de participantes que aceptaron la coima a lo largo de los escenarios mostraron un efecto significativo sólo en el marco cargado (marco abstracto:  $\chi^2_Q=6.25$ ,  $gl=4$ ,  $p=0.18$ ; marco cargado:  $\chi^2_Q=10.08$ ,  $gl=4$ ,  $p<0.05$ ). También comparamos los escenarios de a pares con el objetivo de identificar cuál era el origen del efecto de marco general. Resultaron significativos los contrastes entre los escenarios 1 y 4 ( $\chi^2_Q=5.00$ ,  $gl=1$ ,  $p<0.05$ ) y entre los escenarios 3 y 5 ( $\chi^2_Q=5.00$ ,  $gl=1$ ,  $p<0.05$ ), mientras que resultaron marginalmente significativos los contrastes entre los escenarios 1 y 3 ( $\chi^2_Q=3.57$ ,  $gl=1$ ,  $p=0.06$ ) y los escenarios 4 y 5 ( $\chi^2_Q=3.57$ ,  $gl=1$ ,  $p=0.06$ ). Los demás contrastes resultaron ser no significativos. No haber encontrado cambios significativos en la frecuencia de aceptación de coimas entre tratamientos puede estar relacionado a un problema de falta de potencia experimental, especialmente teniendo en cuenta el patrón de decisiones que se observa en la figura 4. En cambio, la variación en la propensión de los participantes B a aceptar la coima en el marco cargado es más difícil de explicar. Por un lado, dado que el incentivo para rechazar la coima aumentaba a partir del escenario 2, y por lo tanto eso no puede explicar el patrón de elecciones de B. Por otro lado, a partir del escenario 2, aumenta la inequidad entre A y B a favor de este último cada vez que B rechaza la coima, y además la inequidad aumenta progresivamente a medida que pasan los escenarios. Es posible entonces que la sensibilidad frente a la pérdida en la que incurre A pueda explicar el incremento en el número de participantes B que eligieron aceptar la coima de los escenarios 1 a 3 y 4. Sin embargo, resulta

enigmático el motivo por el cual el número de participantes B que eligieron aceptar la coima declinó sensiblemente en el último escenario.

Por último, evaluamos la asociación de los puntajes del cuestionario de maquiavelismo con la *frecuencia de aceptación* y el *incentivo mínimo de aceptación* (del trato corrupto). Realizamos los análisis excluyendo los participantes de acuerdo al criterio del escenario control 7. Como habíamos anticipado en las predicciones, no encontramos evidencia de ninguna asociación significativa entre las decisiones de los participantes y los puntajes de maquiavelismo en ninguno de los tratamientos (*Spearman Rank Order Correlations*: en el marco abstracto, la asociación del puntaje de *Mach* con la *frecuencia de aceptación*,  $\rho=-0.09$ ,  $n=23$ ,  $p=0.69$ ; la asociación del puntaje de *Mach* con el *incentivo mínimo*,  $\rho=0.10$ ,  $n=23$ ,  $p=0.65$  ; en el marco cargado, la asociación del puntaje de *Mach* con la *frecuencia de aceptación*,  $\rho=-0.20$ ,  $n=24$ ,  $p=0.34$ ; la asociación del puntaje de *Mach* con el *incentivo mínimo de aceptación*,  $\rho=-0.07$ ,  $n=24$ ,  $p=0.73$ ). La falta de asociación entre las decisiones y el puntaje de Maquiavelismo refleja que quizás el juego no involucraba una dimensión moral para los participantes, independientemente del marco utilizado. Sin embargo, al igual que con los participantes A, esta conclusión colisiona con el hecho de que los participantes B evitaron elegir la opción P cuando presentamos esta opción usando términos cargados. En efecto, como mencionamos anteriormente, los resultados muestran que el incentivo mínimo de aceptación fue mayor en el marco cargado que en el abstracto, que nosotros explicamos a partir del aumento de los costos morales (o su prominencia) asociados con la aceptación del trato corrupto. A su vez, puede ser que la falta de correlación entre el puntaje de maquiavelismo y las decisiones se deba a que los participantes B hayan interpretado que la rama L del juego involucraba dos opciones similares en términos de las consecuencias para los otros participantes, ya que ambas opciones perjudicaban a otros

participantes, ya sea al participante A (opción O) o a los participantes C (opción P).

## 8. Conclusiones

En este trabajo nos propusimos como objetivo principal investigar de qué manera influye un *marco cargado* en el comportamiento de los sujetos en un juego de corrupción. Extendiendo investigaciones previas, intentamos identificar los mecanismos que subyacen al efecto de marco, intentando abrir así la caja negra de las investigaciones sobre marco en corrupción. En particular, pusimos a prueba la hipótesis de que los individuos serían menos sensibles al marco cargado cuanto mayor sea su puntaje en el Mach Test. Nuestros resultados proveen apoyo a esta hipótesis.

A diferencia de Abbink y Hennig-Schmidt (2006) y Barr y Serra (2009), que no encontraron efectos robustos de marco, nuestro experimento logró identificar un claro efecto de marco *tanto para los jugadores A como los B*. Creemos que la razón del éxito de nuestro experimento en este punto particular consiste en que en nuestro diseño el marco evita cualquier referencia a roles hipotéticos tales como “Firma” o “Funcionario Público”. Aparentemente cuando los roles resultan extraños, pueden resultar muy artificiales a los participantes e inducirlos a adoptar un comportamiento similar al que adoptarían en un juego de roles en lugar de expresar sus preferencias verdaderas (Bardsley, 2005; Barr & Serra, 2009). En cambio, sólo “etiquetando” las opciones logramos que el marco tuviera un efecto potente sobre las decisiones de los participantes, llevándolos a evitar las alternativas connotadas de manera negativa.

Los datos sugieren asimismo que nuestro JdC logró capturar el hecho de que los

participantes exhiben preferencias pro-sociales. Esto se hizo patente en el hecho de que evitaron la opción que generaba externalidades negativas en los terceros, al menos cuando existía otra opción equivalente en términos de ganancia. Encontramos este efecto tanto para los participantes A, es decir, los potenciales oferentes e coimas, en ambos marcos. Para los participantes B, el efecto sólo estuvo presente en el marco cargado. Es posible que los participantes B concibieran la rama “corrupta” del árbol en el marco abstracto como una decisión inevitable entre perjudicar a los participantes A o a los C. En este sentido, puede ser que hayan percibido las opciones como equivalentes en términos morales, lo que explicaría el motivo por el cual no exhibieron una preferencia en contra de la opción con externalidades. Por otro lado, como habíamos anticipado en las predicciones, para los participantes B la opción dominante siempre era rechazar el trato corrupto, con lo cual también evitaban perjudicar a los terceros. Una limitación que queremos mencionar de nuestro diseño en este punto es que los participantes A pueden haber estado motivados por razones de eficiencia para elegir la rama “corrupta” del juego. Nótese que la ganancia total es más alta en la rama L que en las demás alternativas del juego, por lo tanto los participantes A pueden haber elegido la rama L porque es la más eficiente, y no porque tengan una preferencia egoísta o una moral débil (es decir, porque sean corruptos). Queremos resaltar también que esa debilidad del diseño, también lo vuelve más robusto para testear el efecto que puede tener un fraseo cargado. En efecto, el efecto de marco encontrado disuadió a los participantes de elegir la rama “corrupta” del juego, contrarrestando incluso los motivos de eficiencia mencionados anteriormente.

Con el objetivo de capturar diferencias individuales en las respuestas al marco, y para tener una medida independiente de la pro-socialidad de los participantes, implementamos el cuestionario de maquiavelismo. En otros contextos se ha mostrado que las personas con

puntajes altos en maquiavelismo tienden a exhibir la característica actitud del *Homo Economicus*, es decir, de búsqueda de la maximización de los pagos egoístas, independientemente del efecto de las acciones sobre el pago de los otros (Christie & Geis, 1970; Gunnthorsdottir et al., 2002). Por lo tanto esperamos que los participantes con puntajes más altos en el cuestionario estuvieran más inclinados a elegir la opción con externalidades negativas, ya que presumiblemente tendrían menor sensibilidad a los costos psicológicos o morales implicados en la corrupción. En efecto, en el marco cargado encontramos correlaciones significativas (con los signos esperados) entre el puntaje de Maquiavelismo de los potenciales oferentes de coimas y las variables que construimos para medir la corruptibilidad de los sujetos. Sin embargo, no encontramos ninguna asociación entre los puntajes de maquiavelismo y las decisiones de los potenciales oferentes de coimas en el marco abstracto ni con las decisiones de los participantes B en ninguno de los dos marcos. Estos resultados “negativos” parecen sugerir que la dimensión moral capturada por el Mach test puede no coincidir con la medida por nuestro experimento, como mencionamos anteriormente cuando discutimos los resultados. También puede ser el caso que nuestro experimento no resultó ser lo suficientemente poderoso para detectar una asociación entre maquiavelismo y corruptibilidad en todos los casos. Esto puede deberse al menos a tres razones. En primer lugar, la dispersión de puntajes Mach en nuestro experimento resultó ser relativamente baja, especialmente si la relacionamos con otros estudios en los que los autores categorizan a los sujetos en *low-Mach* (los que puntúan dos desvíos estándar por debajo de la media) y *high-Mach* (los que puntúan dos desvíos estándar por encima de la media), es decir que sólo toman en cuenta los casos extremos (e. g., Gunnthorsdottir et al., 2002). Desde luego ello puede llevarse a cabo disponiendo de una muestra lo suficientemente grande. En segundo lugar, la externalidad negativa permanecía fija a lo largo de los escenarios, lo que

puede haber limitado la expresión de una preferencia por no dañar a terceros. Este problema podría ser subsanado implementando una variante del juego en la que las externalidades también se modifiquen. Y en tercer lugar, las alternativas de decisión para B no incluían la posibilidad de que los B pudieran rechazar la coima sin al mismo tiempo maximizar sus ganancias. Esto puede resultar problemático, en primer lugar, porque rechazar el trato corrupto puede ser moralmente ambiguo, en particular en el marco abstracto (el escenario 7 controlaba parcialmente esa opción); en segundo lugar, porque la existencia de incentivos externos puede limitar o incluso ir en contra de la expresión de una preferencia social intrínseca (e. g., ver Schulze y Frank, 2003).

Para finalizar, quisiéramos llamar la atención sobre una implicación potencial de nuestro estudio. Nuestros resultados sugieren que nombrar aquellas acciones socialmente no deseadas con etiquetas que señalen matices moralmente negativos podría ser un medio eficaz para disuadir a la gente de seguir dichas acciones. Creemos que no sólo las campañas anti-corrupción podrían verse beneficiadas, sino que, en principio, se podría aplicar en cualquier circunstancia en la cual se desee lograr una mayor conformidad de los ciudadanos hacia una norma particular.

## Referencias bibliográficas del capítulo 1

- Abbink, K. (2005) Fair salaries and the moral costs of corruption. En Boicho Kokinov (Ed.), *Advances in Cognitive Economics*. Sofia: NBU Press.
- Abbink, K. (2006) Laboratory experiments on corruption. En Susan Rose-Ackerman (Ed.) *The Handbook of Corruption* (pp. 418-437). UK and Northampton, USA: Edward Elgar Publishers, Cheltenham.
- Abbink, K., & Hennig Schmidt, H. (2006) Neutral versus Loaded Instructions in a Bribery Experiment, *Experimental Economics*, 9 (2), 103-121.
- Abbink, K., Irlenbusch, B., & Renner, E. (2002) An Experimental Bribery Game, *Journal of Law, Economics, and Organization*, 18 (2), 428-454.
- Akerlof, G. & Kranton, R. (2000) Economics and Identity, *The Quarterly Journal of Economics*, 115 (3), p. 715-753.
- Armantier, O. & Boly, A. (2013) Comparing Corruption in the Lab and in the Field in Burkina Faso and in Canada, *Economic Journal*, 123(573), p. 1168-1187.
- Amir, O., Ariely, D. & Mazar, N. (2008) The Dishonesty of Honest People: A Theory of Self-Concept Maintenance, *Journal of Marketing Research*, Vol. 45: 633-634.
- Bartels, D., & Pizarro, D. (2011) The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, 121, 154-161.
- Bardsley N. (2005) Experimental Economics and the Artificiality of Alteration, *Journal of Economic Methodology*, 12, 239-251.
- Barr, A., & Serra, D. (2009) The effects of externalities and framing on bribery in a petty corruption experiment, *Experimental Economics*, 12, 488-503.
- Benabou, R. & Tirole, J. (2004) Willpower and Personal Rules, *Journal of Political Economy*, 112 (4), 848-887.

- Berg, J., Dickhaut, J., & McCabe, K. (1995) Trust, reciprocity and social history, *Games and Economic Behavior*, 10 (1), 122–142.
- Bicchieri (2006) *The Grammar of Society. The Nature and Dynamics of social norms*, Cambridge, Cambridge University Press.
- Christie, R., & Geis, F. L. (1970) *Studies in Machiavellianism*. Academic Press: New York.
- Coleman, J. (1990) *Foundations of Social Theory*, Cambridge: Harvard University Press.
- Czibor, A. & Bereczkei, T (2012) Machiavellian people's success results from monitoring their partners, *Personality and Individual Differences*, 53 (3), p. 202–206.
- Dawes, C., Fowler, J., Johnson T., McElreath, R. & Smirnov, O. (2007) Egalitarian motives in humans, *Nature* 446 (7137), 794-796.
- Della Porta, D. & Vanucci, A. (2012) *The Hidden order of corruption. An institutional Approach*, Ashgate.
- Dufwenberg, M., Gächter, S., & Hennig-Schmidt, H. (2011) The framing of games and the psychology of play, *Games and Economic Behavior*, 73, 459-478.
- Elliott, C. S., Hayward, D. M., & Canon, S. (1998) Institutional framing: Some experimental Evidence, *Journal of Economic Behavior & Organization*, 35, 455–464.
- Elster, J. (1989) *The Cement of Society*, Cambridge, Cambridge University Press.
- Falk, A. & Fischbacher, U. (2006) A Theory of reciprocity, *Games and Economic Behavior*, 54(2), 293-315.
- Fehr, E. & Gächter (2002) Altruistic punishment in humans, *Nature*, 415: 137-140.
- Fehr, E. & Fischbacher, U. (2004) Third-party punishment and social norms, *Evolution and Human Behavior*, 25, 63-87.
- Fehr, E. & Fischbacher, U. (2003) The Nature of Human Altruism, *Nature* 425, 785-791.
- Fehr, E. & Schmidt, K. (1999) A theory of fairness, competition and cooperation, *The Quarterly Journal of Economics*, Vol. 114, No. 3, pp. 817-868.

- Graeff, P. (2005) Why should one trust in corruption? The linkage between corruption, norms and social capital. En Graf Lambsdorff, J., Taube, M. y Schramm, M. (Ed.) *The New Institutional Economics of Corruption*, p. 40-58.
- Gunnthorsdottir, A., McCabe, K., & Smith, V. (2002) Using the Machiavellianism instrument to predict trustworthiness in a bargaining game, *Journal of Economic Psychology*, 23, 49-66.
- Haidt, J. (2007) The new synthesis in moral psychology, *Science*, 316, 998-1002.
- Hertel, G., & Fiedler, K. (1994) Affective and cognitive influences in a social dilemma game, *European Journal of Social Psychology*, 24, 131–145.
- Hertel, G., & Fiedler, K. (1998) Fair and dependent versus egoistic and free: Effects of semantic and evaluative priming on the “ring measure of social values”, *European Journal of Social Psychology*, 28, 49 –70.
- Hodgson, G. M., & Jiang, S. (2007) The economics of corruption and the corruption of economics: An institutionalist perspective, *Journal of Economic Issues*, 41 (4), 1043-1061.
- Lambsdorff, J. G. Taube, M., & Schramm, M. (Ed.) (2005) *The New Institutional Economics of Corruption*, New York, Roulledge.
- Lieberman, V., Samuels, S. M., & Ross, L. (2004). The name of the game: predictive power of reputations versus situational labels in determining prisoners dilemma game moves, *Personality and Social Psychology Bulletin*, 30 (9), 1175–1185.
- Lindenberg, S., & Steg, L. (2007). Normative, gain and hedonic goal frames guiding environmental behavior, *Journal of Social Issues*, 63 (1), 117–137.
- Meyer, H. D. (1992) Norms and self-interest in ultimatum bargaining: The Prince’s Prudence, *Journal of Economic Psychology*, 13: 215-232.
- Miller, S. (2010) Institutional Corruption. In Seumas Miller *The Moral Foundations of Social Institutions: A Philosophical Study* (pp. 154-175). New York: Cambridge University Press.

- Ostrom, E. (2000) Collective action and the evolution of social norms, *Journal of Economic Perspectives*, 14 (3), 137-158.
- Rose-Ackerman, S. (2006) Introduction and overview. En Susan Rose-Ackerman (Ed.) *International handbook on the economics of corruption*, (pp. xvi-xxxviii). Cornwall: MPG Books Ltd, Bodmin.
- Schulze, G. G., & Frank, B. (2003) Deterrence versus intrinsic motivation: experimental evidence on the determinants of corruptibility, *Economics of Governance*, 4 (2), 143–160.
- Schweitzer, H. (2005) Corruption – its spread and decline. En Graf Lambsdorff, Taube y Schramm (Ed.) *The New Institutional Economics of Corruption*, p. 40-58.
- Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G., & Fehr, E. (2007) The neural signature of social norm compliance, *Neuron*, 56, 185-196.
- Wilson, D., Near, D. & Miller, R. (1996) Machiavellianism: a synthesis of the evolutionary and psychological literatures, *Psychological Bulletin*, 119(2):285-99.

**Preferencias sociales, externalidades e intenciones en un juego de corrupción**

## 1. Introducción

La corrupción es hoy uno de los problemas más acuciantes que enfrentan nuestras democracias. Sus efectos perniciosos sobre la riqueza social tienen un amplio alcance (ver Huntington, 1968; Shleifer & Vishny, 1993; Mauro, 1995; Goldsmith, 1999; Aidt, 2003; Jain, 2001; Seldadyo & Hayo, 2006; Pelligrini & Gerlagh, 2004), así como sus efectos perjudiciales sobre las instituciones políticas. El Banco Mundial ha identificado en la corrupción “the single greatest obstacle to economic and social development. It undermines development by distorting the rule of law and weakening the institutional foundation on which economic growth depends.”<sup>3</sup>

Además de la búsqueda de rentas (*rent-seeking*), que es general para muchas definiciones de corrupción (Nye, 1967; Huntington, 1968; Transparency International, 2001, 1; Jain, 2001; Lambsdorff, 2002a), la corrupción presupone un elemento normativo en la medida en que implica la violación de una regla asociada a un rol institucional (ver Banfield, 1975; Miller, 2005; Hodgson and Jiang, 2007; Graeff, 2010; Banerjee et al., 2012; Della Porta & Vanucci, 2012). El rol de los funcionarios públicos da contenido a esta regla, en la medida en que le son confiados los intereses y objetivos de aquellos a quienes representan. La traición de dichos intereses está en el corazón de la corrupción.

El abordaje experimental de la corrupción ha ganado recientemente una amplia atención. Los experimentos sobre sobornos permiten recolectar datos duros sobre un fenómeno que es, por su naturaleza intrínseca, difícil de observar y medir directamente. Además resulta relativamente fácil identificar relaciones de causalidad en experimentos

---

<sup>3</sup> [www1.worldbank.org/publicsector/anticorrupt/index.cfm](http://www1.worldbank.org/publicsector/anticorrupt/index.cfm).

económicos, ya que los factores causales de un efecto observado se pueden atribuir a la variación de tratamientos (Falk & Heckman, 2009). Es más, una mejor comprensión de los factores motivacionales del comportamiento corrupto puede resultar relevante en la formulación de políticas públicas destinadas a combatir la corrupción.

Los investigadores debaten sobre las motivaciones individuales que subyacen a la corrupción. Mientras que los abordajes más tradicionales se han enfocado en el egoísmo y el cálculo racional de costos y beneficios (Becker, 1968)<sup>4</sup>, la investigación sobre corrupción en economía experimental se basa en el reconocimiento de la confianza y la reciprocidad como principales impulsores de los actos de corrupción (ver por ejemplo Abbink et al. 2000; Barr & Serra 2009). Dado que la corrupción es una actividad ilegal y no es posible recurrir a instancias legales para hacer cumplir los acuerdos corruptos, las partes involucradas en una transacción ilegal de este tipo tienen que recurrir a mecanismos extra-legales para hacer cumplir sus “contratos corruptos” (Lambsdorff, 2002a). Los trabajos pioneros tienen el mérito de haber extendido el poder explicativo de las teorías de las preferencias sociales del dominio del intercambio en mercados laborales, a los que fueron inicialmente aplicadas (ver por ejemplo Akerlof, 1982; Rabin, 1993; Fehr and Gächter, 2000; Fehr et al. 2009), a un *explanandum* nuevo, la corrupción, i. e., a un dominio en el que las relaciones entre las partes no son *a priori* relaciones de mercado. Sobre los efectos de la reciprocidad hay evidencia abundante y variada en estudios experimentales y de campo (ver Malmendier, te Velde & Weber, 2013). Naturalmente, el movimiento hacia las preferencias sociales en la elaboración

---

<sup>4</sup> El abordaje económico tradicional suele enfatizar el rol de los incentivos económicos y de las oportunidades que tienen los individuos para corromperse. Según Becker: “A person commits an offense if the expected utility for him exceeds the utility he could get by using his time and other resources for other activities. Some persons becomes 'criminals', therefore, not because their basic motivations differ from that of other persons, but because their benefits and costs differ” (Becker, 1968, p. 172).

de modelos teóricos de la corrupción resulta necesario, ya que los modelos tradicionales basados en motivaciones egoístas pueden acomodar fácilmente el comportamiento corrupto cuando el "pago" depende de manera creíble del éxito de la operación o de que el servicio corrupto sea provisto (esto es, es en el interés del individuo pagar el soborno, si la relación corrupta es una interacción repetida), sin embargo su rendimiento teórico es pobre cuando se trata de casos únicos (*one-shot*). ¿Por qué los individuos devolverían un favor y optarían por proveer el "servicio corrupto"? En lugar de ello, los egoístas racionales embolsarían el dinero y traicionarían a la otra parte. En estos casos se puede ver cómo las preferencias sociales podrían explicar el motivo por el cual la corrupción es un resultado probable planteando una preferencia por honrar la confianza y reciprocitar acciones amables.

En los juegos canónicos de corrupción, el ofrecimiento de coimas del primer jugador (la Empresa) genera un beneficio al segundo jugador (el Funcionario Público). Es decir que motivos de reciprocidad positiva están en juego cuando este último provee el servicio corrupto. En este último caso, la corrupción aumenta sin importar si hay externalidades negativas.<sup>5</sup> Reciprocidad positiva, en estos contextos, es de hecho una función del monto de la coima transferido por la Empresa al Funcionario Público (Abbink, Renner & Irlenbusch, 2002). Las transferencias más altas son presumiblemente juzgadas como más amables que las transferencias más bajas (aunque esta característica se infiere, pero no se mide de forma explícita). Téngase en cuenta que el JdC de Abbink et al. (2002) se reduce a un juego de confianza simple, en el que el funcionario público después de recibir la coima tiene una elección binaria, esto es, corresponder la coima (concediendo el favor corrupto) o

---

<sup>5</sup> Una externalidad surge cuando las acciones de un actor focal generan costos o beneficios a otro actor, que no son tenidos en cuenta por el actor focal en sus decisiones (e.g. Coleman, 1990; Mas-Colell, Whinston y Green, 1995). Frecuentemente las externalidades son negativas, como en el caso de la corrupción.

defeccionar. De hecho, estos autores encontraron que la propensión a la coima no fue diferente en relación al comportamiento de los participantes en un juego de confianza (es decir, en el tratamiento sin externalidades negativas). Esto significa que los participantes en los roles de Empresas y Funcionarios públicos aparentemente no tuvieron en cuenta el daño causado a los demás.

Las transacciones de mercado consideradas corruptas se han estudiado principalmente en juegos de intercambio de regalos bilaterales (*gift-exchange games*) en los que la transacción "exitosa" entre los actores incumbentes puede generar externalidades negativas para terceros pasivos (ver Abbink et al. 2000; Barr & Serra, 2009; Malmendier & Schmidt, 2012). Sin embargo, mientras estos estudios explican el hecho de que los intercambios económicos se convierten en actividades corruptas, en el sentido de actividades ilegales o criminales, cuando afectan a terceros pasivos no involucrados, no incorporan, sin embargo el elemento normativo representado por el contrato fiduciario entre el Principal y el Agente. Esto apunta también al reconocimiento de los límites de una evaluación económica de la corrupción, ya que el incumplimiento de un contrato no es un concepto de carácter económico, sino normativo y/o legal.

La idea de que las partes en una transacción corrupta se benefician a costa de los demás miembros de la sociedad transgrediendo una norma es una idea aplicable a una amplia gama de comportamientos que la corrupción comprende. Piénsese, por ejemplo, en la captura del regulador (Spiller, 1991; Dal Bó, 2006), el tráfico de influencias (Harstad & Svensson, 2011), e incluso la malversación de fondos (ver Shleifer & Vishny, 1993). El modelo de Principal-Agente, que ha sido extensamente utilizado en la literatura sobre corrupción, permite capturar la naturaleza estratégica de las diversas instancias de corrupción

mencionadas anteriormente (e.g., Banfield, 1975; Rose-Ackerman, 1975, 1978; Klitgaard, 1988, 1991; Schleifer & Vishny, 1993, Gambetta, 2000; Della Porta & Vanucci, 2012). Una característica central de la relación entre el principal y el agente es la obligación fiduciaria que surge de la relación entre ambos. Las acciones del agente se encuentran subordinadas a ciertas reglas, lo que significa que al confiarle recursos al agente, el Principal espera que el agente respete y avance sus intereses.

Siguiendo a Gambetta (2000), creemos que la mejor manera de describir dicha relación fiduciaria es en el contexto de un Juego de Confianza entre principales y agentes. El tráfico de influencias ofrece un buen ejemplo que sirve a modo de ilustración. Muchos políticos asumen posiciones con altas remuneraciones como lobistas una vez que abandonan o finalizan su función pública, un fenómeno que se ha denominado "puerta giratoria" (*revolving door*) en la política de los Estados Unidos (véase Vidal i Blanes, J., Draka, M. & Fons-Rosen, C., 2012; Vetter, 2013). Aunque no es una práctica corrupta *per se*, guarda un parecido de familia, debido a la estructura de incentivos que los políticos enfrentan, lo que hace inferir que muchas veces aunque aparentemente trabajando en nombre de la ciudadanía, el político puede estar avanzando sus propios intereses privados, ayudando de forma proactiva a la parte de la que tiene más probabilidades de recibir beneficios en el futuro. Un político puede deber gratitud hacia sus votantes (los ciudadanos), y así inclinarse a honrar su confianza. Sin embargo, la asunción de un cargo público incluye la expectativa de mayores ganancias, lo que puede desalinearse los incentivos públicos y los privados, y los políticos pueden tener la tentación de ir en contra de los intereses de la ciudadanía en general. Las preferencias pro-sociales (por equidad o justicia) en este caso, entran en conflicto con las preferencias egoístas, que en el caso de corrupción van de la mano de la preferencia por

corresponder al socio corrupto. Así, a partir de las características de la corrupción explicadas anteriormente, es evidente que los políticos pueden hacer frente a un conflicto entre responder con amabilidad la confianza depositada en ellos o seguir su propio interés.

Este trabajo estudia el caso en el que reciprocarse la intención amable del Principal entra en conflicto con reciprocarse a la pareja corrupta. Tratamos de captar el significado normativo del contrato que vincula al Principal con el Agente mediante el diseño de un juego en el que el tercero que sufre del incumplimiento del contrato, y por lo tanto de las externalidades negativas que se derivan de ello, es el mismo Principal que puede entrar en una relación de confianza con el agente, o abstenerse de ella. La elección del Juego de Confianza para modelar dicha relación contractual responde a diversas razones: el juego de confianza se rige por normas de confianza (emisor) y reciprocidad (receptor), y se acepta generalmente que los participantes en los juegos experimentales juzgan el comportamiento de los demás de acuerdo a expectativas empíricas y normativas (Bicchieri, 2006). Las motivaciones morales que unen a los jugadores en un juego de confianza pueden ser fácilmente transferidas a nuestra historia motivacional. Mediante nuestra manipulación, es decir al permitir que el tercero sea un jugador pasivo, o que asuma un papel activo en el tratamiento alternativo, podemos evaluar si el Agente (el Funcionario o el político) toma en consideración el elemento normativo que surge de la interacción con el Principal (el ciudadano) al tomar sus decisiones. El Principal interactúa en un juego de confianza con el Agente, que después de asumir el cargo podría optar por honrar la confianza del Principal, o pasar el juego a un cliente. Pero aquí está el problema: un principal pasivo no puede señalar ningún tipo de intenciones, y el hecho de que pudiera señalar activamente intenciones amables debería tener un impacto en el comportamiento de los políticos (y tal vez del cliente).

La comparación entre ambas situaciones experimentales nos permite testear el efecto del componente fiduciario representado por el juego de confianza.

Identificar empíricamente los motivos que modelan el comportamiento corrupto es un problema complejo. Además de nuestros tratamientos, también hemos tratado de aumentar la comprensión de las prácticas de corrupción mediante la obtención directa de atribuciones de amabilidad. El hecho de que el Principal confíe en el Agente debe ser percibido como un gesto amable por parte de este, estimulando al agente a honrar la confianza del Principal y evitar generarle daños.

Naturalmente, nosotros no somos los primeros en aplicar los estudios de la literatura de preferencias sociales en el estudio experimental de la corrupción mediante experimentos: estudios previos examinan diferentes aspectos de la práctica del soborno, ya sea la influencia de los salarios (Abbink, 2005; Armantier y Boly, 2008; Van Veldhuizen, 2013), la cultura (Alatas et al, 2009; Cameron et al., 2009; Barr y Serra, 2010;), la influencia de los intermediarios (Drugov et al., 2014), los distintos sistemas de castigo (Engel, Goerg y Yu , 2011) y varias estrategias de lucha contra la corrupción (Banuri y Eckel, 2012; Serra, 2012; Schickora, 2013). Ver Andvig (2005), Abbink y Serra (2012) y Bobkova y Egbert (2012) para reseñas extensas sobre los estudios de la literatura experimental sobre corrupción.

En línea con otros estudios experimentales encontramos que los sujetos tienden a desestimar las externalidades negativas (ver Abbink et al., 2002), y que parece haber una disociación entre la percepción de la amabilidad y la reciprocidad, que exige una comprensión más profunda del papel de las intenciones en presencia de externalidades negativas.

El resto del artículo está organizado de la siguiente manera: en la próxima sección discutimos las bases teóricas y la literatura relacionada. Nuestro diseño experimental es presentado en la sección 3, seguido de las predicciones en la sección 4. En la sección 5 presentamos los datos y los resultados. Finalmente la sección 6 concluye.

## **2. Bases teóricas**

### *2.1. Teorías de preferencias sociales*

En las últimas décadas, la economía del comportamiento ha reformulado profundamente nuestra visión del ser humano en el campo de la economía (Camerer, 2003; Kahneman et al., 1986). Uno de los movimientos más fuertes de la economía del comportamiento se refiere a la modelización de las preferencias sociales. Los modelos de preferencias estándar para el comportamiento que se basan en el supuesto de maximización de la ganancia han ido dando paso gradualmente a modelos más complejos que permiten que las funciones de utilidad de los individuos incorporen preferencias pro-sociales.

En juegos de ultimátum, por ejemplo, los sujetos juegan de a pares y a cada uno de ellos se les asigna uno de dos roles, el de proponente o el de respondiente. El proponente recibe una dotación consistente en un monto positivo de dinero, y tiene que hacerle una oferta al segundo jugador, sobre la manera de compartirlo. Si el respondiente rechaza la oferta, el importe total se pierde y ambos se van con las manos vacías, mientras que si el respondiente acepta, el dinero se divide de acuerdo a la oferta inicial. Un patrón común de respuesta que se observa en este juego consiste en que el respondiente generalmente rechaza las ofertas que percibe como injustas o egoístas, con el resultado de que ambos jugadores se quedan sin nada

(Güth et al., 1982). En juegos de confianza *de una ronda* (Berg et al., 1995) los fiduciarios (*trustees*) suelen regresar al menos lo que se les ha transferido, lo que contradice el supuesto de que los individuos sólo tienen motivaciones egoístas y sólo persiguen la maximización de su propia ganancia. En los llamados juegos de intercambio de regalos (*gift-exchange games*) también se da este patrón de comportamiento, y se atribuye a motivos de reciprocidad que se genera entre los jugadores (Fehr, Kirchsteiger y Riedl, 1993). En general estas “anomalías”, desde el punto de vista de la teoría de juegos, son incorporadas por los economistas experimentales con el objetivo de refinar sus modelos (Camerer and Thaler, 1995).

Una gran cantidad de investigación, sobre todo en la economía del comportamiento y la psicología, muestra que las preferencias sociales son un componente importante de las decisiones individuales (Camerer, Loewenstein y Rabin, 2004). La idea de “preferencias sociales” consiste en que los individuos evalúan un resultado no sólo sobre la base de las consecuencias que tiene para ellos mismos, sino en función también de las consecuencias hacia otros (por ejemplo, Fehr y Schmidt, 1999; Bolton y Ockenfels, 2000). Además de tener preferencias en relación a los resultados (i. e., preferencias “consecuencialistas”), los individuos poseen preferencias en relación a los procesos que generan dichos resultados. Un tipo muy importante de preferencia social es la preferencia por la reciprocidad, según la cual los individuos recompensan las intenciones amables de otros jugadores o su naturaleza benevolente y castigan las intenciones o comportamientos que perciben como crueles o menos amables (por ejemplo, Rabin, 1993; Fehr y Gächter, 2002; Falk y Fischbacher, 2006). En el laboratorio, esto se refleja en la “tendency to respond to perceived kindness with kindness and perceived meanness with meanness” (Sobel, 2005, p. 392). Este tipo de preferencias de reciprocidad son distintas de los llamados “incentivos de juegos repetidos”,

según los cuales devolver amabilidad con amabilidad equivaldría a un movimiento estratégico con el fin de mantener una relación a largo plazo o para construir una reputación favorable.

Para dar cuenta de los resultados anómalos comentados más arriba, se han propuesto diferentes modelos a la par de los resultados experimentales. Intentaremos anclar la discusión de dichos modelos en referencia a dos categorías amplias, que se refieren a modelos consecuencialistas por un lado y a modelos de reciprocidad basados en la intención por otro. El consecuencialismo pone el énfasis en las consecuencias y los resultados de los procesos. Para una teoría ética consecuencialista el valor económico o moral de una acción deriva, no tanto de características intrínsecas de la acción como puede ser la intención del actor, sino de las consecuencias que se derivan de ella, y de las consecuencias de las posibles alternativas a dicha acción. Por lo tanto, el consecuencialismo es un concepto “paraguas” que comprende modelos de altruismo, eficiencia y teorías de la justicia basadas en los resultados, incluyendo los modelos de aversión a la inequidad (aversión a la inequidad consiste en la preferencia de las personas por resultados equitativos por sobre resultados inequitativos) (Fehr & Schmidt 1999; Bolton & Ockenfelds 2000; Eek y Gärling, 2006; Dawes et al., 2007), el criterio maximin rawlsiano (Charness & Rabin 2002) y la aversión a la injusticia (Konow, 2000). Un rasgo importante de estos modelos es que asumen que las características de los individuos son fijas y que no se modifican de acuerdo a las circunstancias. Consideremos por ejemplo de qué manera predice el comportamiento un modelo altruista en un juego de confianza. La función de utilidad de un altruista está de alguna manera cerrada a las intenciones de los otros jugadores con los que interactúa. Un altruista en el rol de fiduciario siempre debería devolver al primer jugador el máximo monto posible. Bajo el supuesto de que el fiduciario exhibe

preferencias altruistas, el primer jugador debería transferirle la mayor cantidad de dinero posible. La solución no apela en ningún momento a las intenciones de los actores, sino a características de los actores que son fijas y determinadas de manera exógena, y que no dependen realmente de la interacción con otros jugadores. Mientras que el consecuencialismo se enfoca exclusivamente en los resultados, en los últimos años se ha propuesto una clase diferente de modelos que toma en consideración las intenciones de los individuos. Las teorías de la reciprocidad basadas en la intención (e.g Rabin, 1993; y Duwenberg & Kirchsteiger, 2004) tienen su origen formal en la teoría de juegos psicológicos (Battigalli & Dufwenberg, 2009) y definen la justicia o la equidad en términos de las intenciones. Los individuos recompensan las intenciones amables dirigidas hacia ellos o eventualmente hacia terceros (reciprocidad indirecta) y también castigan las intenciones que perciben como no amables. Particularmente relevante para nuestro trabajo es el estudio de Falk y Fischbacher (2006). En su modelo, las preferencias dependen de las intenciones que los individuos perciben en una acción, y por lo tanto de las creencias en relación a las intenciones de los otros individuos. Según los autores, la reciprocidad es una respuesta comportamental a la amabilidad de otro individuo en una interacción, y por lo tanto es crucial entender de qué manera las personas evaluamos y atribuimos amabilidad a una acción. Para ello implementaron un cuestionario en el que los participantes respondían a una situación hipotética en el que un sujeto  $s$  interactuaba con otro sujeto  $i$ . Este último debía evaluar la amabilidad de diferentes divisiones monetarias propuestas por  $s$ . Variando el conjunto de alternativas que  $s$  tenía para elegir, los investigadores podían investigar qué era lo que tenían en cuenta los sujetos  $i$  para evaluar la amabilidad de  $s$ . Por ejemplo, si  $s$  tenía en su conjunto de estrategias la posibilidad de un pago equitativo, entonces cualquier otra decisión que tomara podía ser evaluada en función de la comparación con la acción equitativa. Nótese que en nuestro escenario, cada

individuo tiene en su conjunto de estrategias una acción equitativa. Sin embargo, señalar intenciones amables puede resultar difícil. Por ejemplo, podemos interpretar un regalo como un intento de ser influenciados, mientras que no recibir un regalo cuando era esperado puede dar lugar a un reproche (Malmendier & Schmidt, 2012). Dependiendo de las circunstancias, no reciprocitar una acción, cuando se espera que lo hagamos puede generar reciprocidad negativa (ver Mullainathan, 2004). En el modelo de Falk y Fischbacher, las acciones con idénticas consecuencias pueden generar respuestas diferentes dependiendo de cómo son interpretadas. En la siguiente sub-sección discutimos brevemente literatura experimental relacionada con corrupción.

## 2.2 Trabajos relacionados

Además de los trabajos mencionados más arriba, nuestro estudio se relaciona al menos con tres líneas de investigación de la literatura experimental sobre corrupción. En primer lugar, hay un número creciente de trabajos que estudian tópicos como confianza y reciprocidad en juegos de confianza (Abbink et al., 2000, 2002; Barr & Serra 2009, Lamsdorff & Frank, 2007; Jacquemet, 2012). Abbink y colaboradores (2002) propusieron el Juego de Coima (*Bribery Game*) original, con la pretensión de ser interpretado como una interacción entre una empresa y un funcionario público. En el juego pares de participantes interactuaron por 30 rondas (siempre con los mismos compañeros) ya sea en un Juego de Confianza (tratamiento control), un Juego de Confianza con externalidades negativas (tratamiento NE), o un Juego de Confianza tanto con externalidades negativas como con posibilidad de castigo (tratamiento NE + C). En primer lugar, en el Juego de Confianza el Jugador 1 (la empresa a partir de ahora) debía optar por transferir puntos experimentales (convertibles a dinero al final del experimento) al Jugador 2 (el Funcionario Público a partir

de ahora), quien a su vez debía elegir entre las opciones X e Y. Asumiendo que la Empresa transfirió dinero al Funcionario Público, la elección de la opción Y era un indicador de reciprocidad de parte del Funcionario Público, tanto porque Y significaba un pago más alto para la Empresa que X, como porque la opción X maximizaba el pago para el Funcionario Público. En segundo lugar, para modelar las externalidades negativas de la corrupción, los autores implementaron el tratamiento NE, en el que cada vez que la Empresa transfería dinero al Funcionario Público (simulando un ofrecimiento de coima) y el Funcionario Público elegía la opción Y (simulando la provisión del servicio corrupto), todos los otros pares de jugadores en la misma sesión experimental sufrían una pérdida importante de puntos. Por último, Abbink y colaboradores modelaron el costo potencial por ser descubiertos y castigados implementando una variación al tratamiento NE, en el que incluían una posibilidad baja de detección, pero que significaba un castigo monetario drástico (pérdida de todos los puntos y exclusión del juego), cada vez que se concretaba un trato corrupto (es decir, cada vez que la Empresa transfería dinero y el Funcionario Público elegía la opción Y).

Los resultados principales del experimento fueron los siguientes: a) Empresas y Funcionarios Públicos desarrollaron confianza y reciprocidad en todos los tratamientos, lo que significa que hubo transferencias significativas de dinero de las Empresas a los Funcionarios Públicos, y que transferencias mayores estuvieron asociadas con una mayor probabilidad de que los Funcionarios Públicos elijan Y; b) tanto el promedio de transferencias de las Empresas como el promedio de elección de Y por parte de los Funcionarios Públicos fue más alto en los tratamientos control y NE que en el tratamiento NE + C; c) los participantes en los tratamientos control y NE no difirieron en cuanto al promedio de transferencias ni en cuanto a la propensión a elegir la opción Y. De estos resultados, Abbink

y colaboradores concluyeron que los participantes fueron insensibles al daño que significaban las externalidades para los terceros, y que sólo vieron modificado su comportamiento cuando existió la posibilidad de que sean detectados y castigados (Abbink et al., 2002).

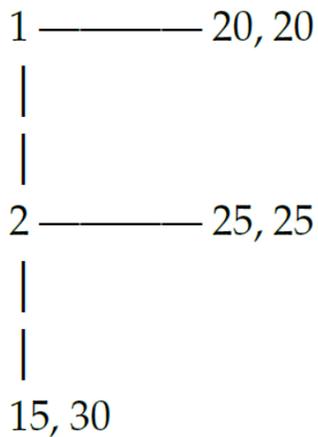
El trabajo que más se acerca al nuestro es el de Jacquemet (2012). Basándose en el juego de Abbink y otros, el estudio de Jacquemet se centra en el conflicto de reciprocidades que el agente, potencialmente corrupto, enfrenta entre el Principal y la pareja corrupta. El estudio compara un tratamiento denominado *Corrupción*, que es idéntico al tratamiento control utilizado en Abbink y otros, con un tratamiento llamado *Delegación Explícita*. En este tratamiento los mismos jugadores que jugaron durante 10 rondas en el tratamiento *Corrupción* son emparejados durante otras 10 rondas con otro jugador que hace las veces de Principal. Las acciones X e Y del juego de Abbink y otros, pasaron a representar bajo esfuerzo y alto esfuerzo, respectivamente. En este punto del juego, el Principal podía fijar un salario bajo o un salario alto para remunerar al agente. Dado que un incentivo salarial alto es un mecanismo poderoso para mantener relaciones de reciprocidad (ver Brown et al., 2004), Jacquemet argumenta que dicho incentivo le generaría un conflicto de reciprocidad al agente entre reciprocitar al Principal o a la pareja corrupta, lo que daría por resultado una disminución de la corrupción. El trabajo aporta como conclusión que incrementar los salarios puede disminuir la corrupción. A pesar de que el trabajo de Jacquemet (2012) es el que más se acerca al nuestro, difiere sin embargo en varios aspectos importantes. En primer lugar, mientras que Jacquemet está interesado en el efecto de los salarios sobre la corrupción, nuestro interés se centra en el elemento normativo que surge de la relación fiduciaria entre Principal y Agente. En segundo lugar, a diferencia de Jacquemet, en cuyo experimento los

participantes interactúan en todas las rondas siempre con las mismas parejas, en nuestro experimento los participantes interactúan cambiando de grupos continuamente. De este modo podemos descartar efectos reputacionales que pudieran surgir debido a la repetición. Y en tercer lugar, nuestro diseño es mucho más simple, lo que permite testear por efectos de tratamiento de una manera más limpia.

Un segundo conjunto de trabajos relacionados se centra en el efecto de las externalidades, ya sean positivas o negativas, sobre terceros pasivos. Una diferencia decisiva entre los juegos de ultimátum o confianza y cualquier juego que pretenda simular una situación de corrupción radica en que en el último caso la corrupción genera externalidades negativas en un tercero. A pesar de ser un tema relativamente poco estudiado, su importancia es central para entender el comportamiento corrupto. Tener en el foco de atención el daño causado a terceros debería aumentar los costos morales y por lo tanto disminuir la incidencia de corrupción. Sin embargo, la evidencia experimental sobre el efecto del daño o la externalidad a terceros sobre la incidencia de los niveles de corrupción en el laboratorio es mixta (ver Abbink et al., 2002; Abbink, 2005; Barr & Serra, 2009). Aún no se comprende cabalmente cómo es la relación entre el daño a terceros y la corrupción. También se ha argumentado que la corrupción es una acción intencional y por lo tanto a quienes incurren en ella simplemente no les interesa el bienestar de los posibles perjudicados (Engel et al., 2012).

En tercer lugar, también hay una serie de trabajos que muestran que los individuos tienden a diferenciar entre jugadores pasivos y activos. La mayoría de los modelos de reciprocidad brevemente comentados anteriormente (e. g., Rabin, 1993; Dufwenberg & Kirchsteiger, 2004; Falk & Fischbacher 2006) predicen un fenómeno que llaman reciprocidad

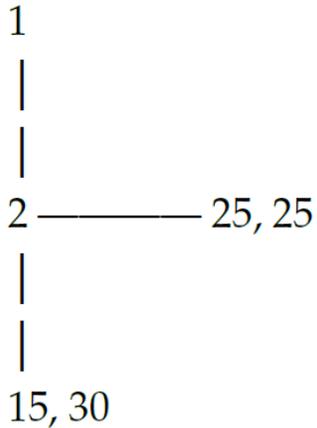
positiva.<sup>6</sup> Esto quiere decir simplemente que los individuos son, en promedio, más amables con aquellos que fueron amables con ellos, que con jugadores pasivos que no realizaron ninguna acción (o cuya acción no puede ser interpretada como intencional). Relevante en este sentido es el estudio de MacCabe et al. (2003), en el que presentaron a sus sujetos el siguiente juego de confianza:



El Jugador 1 puede jugar “derecha” (que es el equilibrio perfecto en el sub-juego), asegurándose un pago de 20, o puede jugar “abajo”. En este último caso, el Jugador 2 puede elegir entre el máximo beneficio para ambos (25, 25), o puede elegir la opción egoísta que perjudica al primer jugador. Ahora comparemos este tratamiento que los autores llaman “voluntario”, con el tratamiento alternativo que llaman “involuntario”, en el que el primer jugador no tiene más alternativa que jugar “abajo”:

---

<sup>6</sup> MacCabe et al. (2003) definen “reciprocidad positiva” de la siguiente manera: “Positive reciprocity can be described as the costly behavior of a second mover that rewards a first mover based on both the gains from exchange to the second mover as well as the second mover’s beliefs about the intentions motivating the action of the first mover” (p. 269)



El importante resultado que encuentran los autores es que más del doble de los jugadores eligieron “derecha” en el nodo 2 del tratamiento “voluntario” en comparación con el “involuntario”. Es claro que desde un punto de vista consecuencialista, no debería haber diferencias de tratamiento. Desde luego, estos resultados respaldan la idea de que la cooperación entre los individuos no depende exclusivamente de la distribución final de los resultados (en términos de los pagos finales), sino que los individuos toman en consideración lo que los otros individuos tienen en mente. ¿Por qué razón el jugador 2 elegiría “derecha” en lugar de “abajo”, en respuesta a la elección “abajo” del jugador 1? MacCabe et al. sostienen que si el jugador 1 elige “abajo” señala una intención de entrar en una relación recíproca de confianza, y el jugador 2 quiere reciprocitar esta confianza. A continuación introducimos nuestro diseño experimental del juego de corrupción.

### 3. Diseño experimental

Diseñamos un juego secuencial de tres jugadores. La figura 6 es una representación gráfica del juego en forma extensiva. Cada uno de los jugadores enfrenta una decisión

dicotómica entre elegir *in* o elegir *out*. Como es evidente a partir de la estructura de pagos, las distintas alternativas tienen consecuencias monetarias diferentes para cada uno de los jugadores. El jugador en el Rol A actúa como un Principal Benevolente, y tiene la opción de terminar el juego eligiendo *out* y asegurarse así un pago de 15 (los pagos están expresados en puntos experimentales), o pasarle el juego al jugador B (que simula el rol de Agente) eligiendo la opción *in*. La decisión de A de jugar *in* genera un excedente, que representa la renta que puede apropiarse B. El Agente, por su parte, puede elegir *out* y compartir de manera equitativa el excedente generado por el Principal, o puede elegir *in* y pasar el juego al jugador en el rol C (que simula el rol de un cliente y que representa la oportunidad para el Agente de traicionar la confianza del Principal). En el último caso, los pagos finales dependen de la decisión de C. Si C elige *out*, rechaza el trato con el Agente y la acción genera el resultado más equitativo (20, 20, 20). Nótese que además de ser el resultado más equitativo, también es una solución Pareto-superior a la prescripta por la teoría de juegos no cooperativa. Si por el contrario, el cliente acepta el trato corrupto del Agente eligiendo *in*, entonces el Principal sale perjudicado y B y C se reparten la porción más grande de la riqueza generada (10, 25, 25).

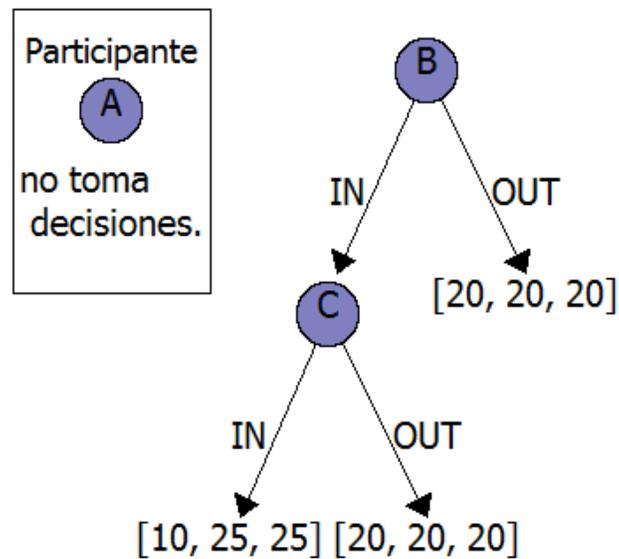
La estructura del juego es conocimiento común de todos los jugadores, así como el número de rondas, y el pago asociado a cada alternativa de juego. De la estructura de pagos se infiere claramente que modelamos un problema de confianza entre los jugadores A y B, en el cual pasando el juego o confiando, A puede inducir en B una motivación por reciprocidad positivamente (ver Camerer; 1988, Dasgupta, 1988; Kreps, 1990). Téngase en cuenta que en nuestro escenario cada jugador tiene a su disposición una opción equitativa para elegir. Por lo tanto hay una manera evidente de medir la amabilidad de cualquiera de los jugadores,

comparando la acción elegida con la alternativa equitativa disponible. Téngase en cuenta también que se dan las características mencionadas por MacCabe et al. para que exista reciprocidad positiva: primero, hay una ganancia mutua para A y B; en segundo lugar, A toma un riesgo en confiar en B; y en tercer lugar, B renuncia a algo para reciprocitar la decisión de A. La ganancia conjunta que pueden obtener A y B se mide en comparación con el equilibrio perfecto en subjuegos (que para A consiste en elegir *out*).

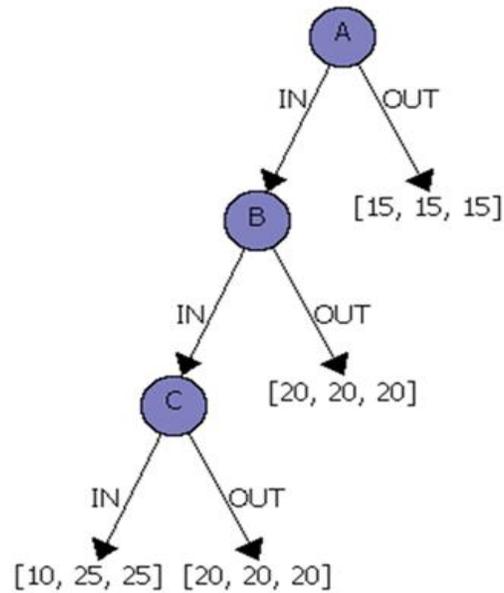
En lo que sigue argumentamos en favor de algunas características del presente modelo en comparación con modelos previos estudiados en la literatura experimental. En los juegos canónicos de corrupción (e.g., los de Abbink y otros, Barr y Serra, Alatas y otros, etc.) los socios corruptos se reparten un excedente en detrimento de un tercero. Sin embargo, en estos modelos, el agente corrupto tiene un incentivo para traicionar a la pareja corrupta, esto es, la maximización de ganancia y la reciprocidad no van de la mano. Por lo tanto estos modelos son útiles para demostrar la importancia de la reciprocidad en tanto mecanismo de protección frente al oportunismo (Lambsdorff, 2010). Desde luego, la razón para favorecer esta característica en un diseño depende del objetivo de investigación que uno tenga. A pesar de que este aspecto del diseño de los juegos de corrupción es muy útil para medir motivaciones de reciprocidad entre oferente de coima y coimeado, no es una característica necesaria *per se* de la corrupción. En nuestro escenario experimental, el político o funcionario perjudica a terceros (a los ciudadanos que depositaron su confianza) solamente dejándose llevar por su interés egoísta (otros estudios experimentales comparten esta misma característica, por ejemplo, Gneezy et al., 2014). Elegimos utilizar un protocolo en el que los sujetos interactúan en cada ronda con diferentes sujetos (llamado en la literatura “*stranger design*”). Esta característica del diseño fue elegida para evitar que las decisiones de los

sujetos estuvieran motivadas por los llamados incentivos de juegos repetidos. No permitir efectos reputacionales nos permite descartar explicaciones alternativas, y poner a prueba nuestras hipótesis de manera más clara. Siguiendo la práctica estándar en economía experimental, utilizamos un lenguaje neutral en las instrucciones para evitar referencias a la corrupción.

**Fig. 5.** Ilustración del juego en forma extensiva correspondiente al tratamiento base tal como fue utilizado en el experimento. Los pagos entre corchetes corresponden a los pagos de A, B y C, respectivamente.



**Fig. 6.** Ilustración del juego en forma extensiva correspondiente al tratamiento activo tal como fue utilizado en el experimento. Los pagos entre corchetes corresponden a los pagos de A, B y C, respectivamente.



### 3.1. Tratamiento Base

Nuestro experimento consiste en dos tratamientos: tratamiento base (TB) y tratamiento activo (TA). En el TB, el jugador en el Rol A es un jugador pasivo, es decir, que no toma decisiones en el juego, pero los pagos que recibe pueden ser potencialmente afectados por las decisiones de los otros jugadores. El jugador en el Rol B puede elegir entre las acciones *in* o *out*. Si elige *out*, el juego termina y los respectivos pagos son (20, 20, 20). Si elige *in*, le pasa el juego a C, quien a su vez tiene que elegir entre *in* o *out* (ver figura 5). Si el jugador en el rol C escoge *out*, el juego termina con pagos equitativos para cada jugador (20, 20, 20), mientras que si C elige *in*, B y C obtienen una ganancia mayor que A.

### 3. 2. Tratamiento Activo

Nuestro TA difiere del TB en que el jugador A es activo, lo que significa que participa del juego activamente tomando una decisión. De hecho, el jugador A es el primero en mover en la secuencia (véase la figura 6), y tiene que elegir entre salir del juego eligiendo *out*, lo que asegura una ganancia de 15 para cada jugador, o pasarle el juego a B, quien a su vez podría elegir la opción justa, que finaliza en pagos equitativos para cada jugador (20, 20, 20), o *in*, lo que implica pasar el juego a C. Obviamente, B recibe un pago superior cada vez que A renuncia a la ganancia segura de 15, y decide pasarle el juego. Por esta razón, el jugador A podría esperar que B ajuste su comportamiento a una norma de beneficencia recíproca. En ese caso, que ambos se beneficien significa para A que B elija *in*. Si B y C eligen *in*, A está siempre en una situación peor de lo que hubiera estado si no hubiera delegado su decisión a B.

### 3.3. Procedimiento

Llevamos a cabo el experimento en el Laboratorio de Sociología del Departamento de Sociología de la Universidad de Groningen en los Países Bajos. Los sujetos fueron invitados a participar a partir de una base de datos de aproximadamente 1.200 estudiantes voluntarios de diferentes carreras. Para ello utilizamos el software de reclutamiento ORSEE (Greiner, 2004). Tanto el reclutamiento como el experimento se ajustaron a las normas éticas establecidas por el Laboratorio de Sociología (<http://www.gmw.rug.nl/~orsee/public/privacy.php>). En total, 9 grupos participaron en el TB, y 10 grupos en el tratamiento TA. Los sujetos participaron en sólo uno de los tratamientos, por lo que nuestros tratamientos fueron entre sujetos. El experimento fue programado

utilizando el software z-Tree (Fischbacher, 2007). Los participantes ganaron aproximadamente 12 € (€1≈AR\$10 al momento del experimento) en promedio por persona y cada sesión duró alrededor de 45 minutos.

Una vez en el laboratorio, cada uno de los participantes se ubicó en un escritorio, cada uno de los cuales contaba con una computadora. A los participantes se les dijo que iban a participar en un experimento de toma de decisiones y que la comunicación entre ellos no iba a estar permitida durante el estudio. También se les comunicó que lo que ganaran en el experimento dependería de sus propias decisiones y de las decisiones de los demás. Se dejó en claro que todas las decisiones que tomaran permanecerían anónimas y que se les pagaría en efectivo al final del estudio dependiendo de la cantidad de puntos experimentales que hubieran ganado durante el experimento. Antes de iniciar el experimento se enfatizó en el hecho de que el experimento no incluía engaño en ninguna de sus formas.

Al principio de cada ronda los participantes fueron divididos aleatoriamente en grupos de tres (3) personas. En cada ronda, después de que se formaran los grupos, el programa asignaba aleatoriamente una identidad, es decir, un rol a cada participante. Los roles eran tres (3): rol A, rol B y rol C. Los roles también eran reasignados cada ronda. Asimismo la composición del grupo se modificaba después de cada ronda, lo que significa que los participantes eran reasignados a grupos diferentes en cada ronda. Por lo tanto cada sujeto asumía alternadamente los diferentes roles (A, B o C) durante 20 períodos. Esto significa que en promedio, los participantes interactuaban en cada rol durante alrededor de seis (6) rondas. En ambos tratamientos, los participantes recibían información sobre todas las decisiones tomadas en su grupo y sobre la cantidad de puntos ganados por cada miembro al final de cada período. Después de completar el experimento se les brindó información general

sobre los objetivos del experimento y se procedió a pagarles a los participantes. Para ello los ingresos de los sujetos fueron puestos en sobres, cada uno de los cuales fue recogido de forma privada por el participante correspondiente, asegurando así que no pudieran enterarse lo que los otros participantes habían ganado.

#### 4. Predicciones

El propósito de la manipulación experimental consistía en testear si el contrato fiduciario entre principal y agente podía inducir diferencias en la probabilidad de elección de la opción *in*. El escenario experimental nos permite investigar asimismo el rol de la amabilidad de A comparando la frecuencia con que los jugadores en los roles B y C eligieron la acción *in* en el TP *vis-a-vis* el TA. Para ser más precisos, mediante la comparación de los tratamientos podemos evaluar si las diferencias en las opciones podían ser resultado de las preferencias por reciprocidad basadas en la percepción de la amabilidad de A. Por ejemplo, en el TA un mayor nivel de *out* podría ser el resultado de reciprocidad positiva hacia A, ya que B puede percibir la delegación de A como amable. En este caso, deberían observarse más elecciones de *out* (en sentido inverso menor proporción de elecciones de *in*) en el tratamiento activo, ya que de esa forma los sujetos pueden beneficiar a A o al menos evitar generar externalidades negativas hacia A. Sin embargo, si comenzamos por considerar motivaciones económicas estándar del comportamiento, es evidente que B y C en ambos tratamientos elegirán las acciones que correspondan a su mayor nivel de utilidad. En un modelo de elección racional de maximización de la utilidad, tanto B como C elegirán jugar *in* en detrimento del tercero que sufre de las externalidades negativas. Nótese que desde el punto de vista de B y C el juego se transforma en un juego de aseguramiento (*Assurance Game*)

después de que A decide pasar el juego a B. Ambos, B y C, están en mejor situación si cooperan entre sí. Pero la cooperación no es claramente una mejora de Pareto, ya que A va a estar siempre peor, a pesar de que B y C mejoren sus pagos.

Esta característica de nuestro diseño nos proporciona un argumento de inducción retroactiva (*backward induction*) para obtener el único equilibrio perfecto en sub-juegos para el TA. Por paradójico que pueda parecer, si presuponemos que los individuos son egoístas racionales, sumado al supuesto de conocimiento común, entonces el equilibrio perfecto en el sub-juego predice un nivel de corrupción cero. Si el jugador B es maximizador de ganancias, entonces siempre tratará de llegar a un acuerdo corrupto pasando el juego a C, mientras que C siempre elegirá la opción que más favorece a B y C. Sabiendo esto, el jugador A no debería confiar nunca, y por lo tanto elegir siempre la opción *out*, lo que impide que se genere el excedente y por lo tanto que surja la corrupción. Esta solución juego-teórica se basa en la suposición de que los sujetos están motivados únicamente por sus propios beneficios. Por esta razón, no importa si el jugador A es un jugador activo o si tiene un rol pasivo, o si el jugador A puede señalar intenciones justas. Re-describir el problema de decisión, y manipular si el jugador A es activo o es un jugador pasivo, a la vez que mantener constantes los pagos, no debería tener ningún impacto en la conducta de los sujetos siempre y cuando la ponderación de la situación experimental dependa exclusivamente de los resultados. En términos más generales, podemos afirmar que el consecuencialismo no predice ninguna diferencia de comportamiento entre los tratamientos. De hecho, un consecuencialista no se ve afectado por las intenciones de otros actores. En los dos extremos del espectro consecuencialista, tanto el egoísmo como el altruismo, no deberían predecir diferencias de comportamiento entre los tratamientos. Sin embargo, en experimentos en los que los

jugadores se enfrentan a escenarios experimentales de corrupción que presentan incentivos similares se producen en el laboratorio altos niveles de ofrecimiento y aceptación de coimas (Abbink et al, 2002; Barr y Serra, 2009). En estos estudios, la reciprocidad ayuda a explicar el comportamiento por fuera del equilibrio. Por lo tanto, esperamos un nivel no despreciable de elecciones de *in* por parte de A. En base a estos resultados previos hipotetizamos que los jugadores en los roles B y C llegarán a un acuerdo corrupto. Esto nos permite formular nuestra primera hipótesis:

*Hipótesis 1:* Se observa un nivel no despreciable de corrupción, *independientemente del tratamiento.*

Por el contrario, los modelos de reciprocidad basados en la intención nos proporcionan un argumento diferente para evaluar la conducta de los individuos. Si los sujetos tienen preferencias sociales, entonces esperamos que atribuyan mayor amabilidad a los individuos que “pasan el juego” que a aquellos que no lo hacen. La percepción de la amabilidad de A debería aumentar cuando A elige la acción *in*, ya que incrementa las ganancias de los otros jugadores a la vez que asume un riesgo. La presencia de un jugador amable como jugador activo en el Rol A debería afectar, en alguna medida, la proporción de tratos corruptos entre B y C. Por lo tanto, esperamos una disminución de la corrupción en el TA, donde la presencia de un jugador amable que puede señalar intenciones debería generar un mayor costo moral en B y C por elegir la acción *in*. Esperamos asimismo que este efecto se dé tanto en B como en C. De lo dicho anteriormente derivamos las siguientes hipótesis:

*Hipótesis 2:* La probabilidad de que los jugadores en el rol B elijan *in* es más baja en el TA que en el TB.

Para la siguiente hipótesis, incorporamos el efecto de las atribuciones de amabilidad en el comportamiento.

*Hipótesis 2b:* Los jugadores en el Rol B son menos proclives a elegir *in* a medida que perciben mayor amabilidad por parte de A.

Argumentos similares a los que respaldan la hipótesis 2 se aplican también a los jugadores en el rol C. Sin embargo, como los pagos finales (en términos de la ganancia que reciben los jugadores en cada ronda) dependen en última instancia de la acción de C, es posible que esta característica del proceso de decisión pueda aumentar la sensación de responsabilidad para C. La reciente literatura experimental sobre el efecto psicológico de la delegación viene en apoyo de esta conjetura. En una serie de trabajos, se ha identificado que en ciertos juegos, por ejemplo en juegos de Dictador, la delegación de la decisión ayuda en la reducción de la responsabilidad del delegante, cada vez que la decisión implica un comportamiento egoísta o injusto de parte del delegante (ver Coffman, 2010; Hamman et al. 2010; Bartling y Fischbacher, 2012). Precisamente Bartling y Fischbacher estudiaron la atribución de responsabilidad empleando un juego de dictador de una ronda (*one-shot*) con cuatro jugadores, un dictador, un delegado y dos receptores. El dictador podía elegir entre un resultado justo, que implicaba que cada uno de los cuatro jugadores recibía la misma ganancia, o un resultado injusto, lo que daba mayores ganancias al dictador y al delegado y una ganancia sensiblemente menor a los otros jugadores. El dictador podía tomar la decisión por sí mismo o delegar la decisión al jugador en el rol de delegado. Los autores medían la atribución de responsabilidad observando la disposición de los receptores para castigar al dictador cuando delegaba su decisión y comparándola con un tratamiento en el que el dictador no podía delegar. Bartling y Fischbacher encontraron que cuando los dictadores

decidieron la asignación injusta por delegación, fueron menos castigados por los receptores. Como conclusión del trabajo, para los autores, la delegación conduce a un desplazamiento del *locus* de la responsabilidad y permite que se le “eche la culpa” al delegado. Esta idea también fue aplicada al estudio de los intermediarios en un experimento de soborno (Drugov et al., 2014). Podría decirse que, al trasladar la culpa, delegar las decisiones a un intermediario puede reducir los costos morales asociados a la decisión de sobornar. Aunque en nuestro juego no incluimos una medida de atribución de responsabilidad, siguiendo las ideas de estos estudios, podemos esperar que los jugadores en el Rol B elijan más a menudo *in* que los jugadores en el rol C, ya que la responsabilidad de los pagos depende en última instancia de la elección de C y B puede interpretar su decisión como una delegación hacia C. De lo anterior derivamos la hipótesis siguiente:

*Hipótesis 3:* La probabilidad de que C elija *in* es mayor en el TB que en el TA, pero en menor medida que los jugadores en el rol B.

Teniendo en cuenta las atribuciones de amabilidad declaradas por los sujetos, podemos afirmar la siguiente hipótesis asociada:

*Hipótesis 3b:* La probabilidad de que C elija *in* disminuye a medida que percibe mayor amabilidad en A.

Nótese que los jugadores en el Rol B, así como los jugadores en el rol C, se encuentran constantemente en un dilema, ya que tienen que decidir entre honrar la confianza de A o favorecer al otro jugador al tiempo que maximizan sus propios beneficios. En consecuencia, especulamos que tanto los jugadores en el rol B, como los jugadores en el rol C tenderán a percibir progresivamente menor amabilidad en los jugadores en el rol A, porque es la acción de A la que los pone en el dilema mencionado anteriormente.

*Hipótesis 4:* La percepción de los niveles de amabilidad de los jugadores en el rol A tiende a disminuir a medida que avanzan las rondas.

## 5. Datos y variables

El experimento se compuso de 4 sesiones, con 5 grupos de tres en tres sesiones, y 4 grupos de tres en la restante. Por lo tanto, se obtuvieron observaciones de 57 sujetos: 30 sujetos en el TB y 27 sujetos en el TA. Debido a que cada sujeto jugó durante 20 rondas alternando en diferentes roles, tenemos un total de  $57 \times 20 = 1140$  observaciones. Dado que utilizamos un protocolo de respuesta directa en vez del método de las estrategias (*strategy method*) tenemos observaciones de las decisiones de B sólo si A elige *in*. Lo mismo se aplica a las decisiones de C. Ya que los jugadores en el rol A son pasivos en el TB, no podemos comparar el comportamiento de A a través de los tratamientos. Naturalmente, no todos los jugadores recibieron el turno en cada ronda, por lo que el número de decisiones reales asciende a 760. Con el objetivo de testear nuestras hipótesis no podemos limitarnos a la comparación de los dos tratamientos. Más bien, tenemos que estimar la probabilidad de elegir *in* en función del tratamiento experimental, del efecto potencial de la ronda, y de las atribuciones amabilidad. La variable dicotómica Acción toma valor 1 si el actor elige *in*, y 0 si el actor elige *out*. Naturalmente, podemos comparar la probabilidad de elegir la acción *in* entre tratamientos para los jugadores B y C solamente, pero no para los jugadores A, porque en el tratamiento de referencia A no toma ninguna decisión. La variable Ronda-centrada se definió tomando el promedio de las rondas es decir  $\text{Ronda-centrada} = \text{Ronda} - 10.5$ . Por lo tanto los valores de Ronda-centrada están en un rango que va de -9,5 a 9,5. Por último

medimos las atribuciones de amabilidad pidiendo a los sujetos, luego de finalizada cada ronda, que valoren la amabilidad de los demás participantes en su grupo (¿Cuán amable era el Participante X para vos?) utilizando una escala de siete puntos de Likert que va de 1=muy poco amable a 7=muy amable. Para realizar las “atribuciones de amabilidad” se les pidió que tengan en cuenta las decisiones que habían tomado los otros participantes. Este método sigue de cerca el método utilizado por Falk et al. (2008).

## 6. Resultados y discusión

Dado que tenemos un diseño con medidas repetidas y con una variable dependiente dicotómica (i. e., la decisión de jugar *in* o *out*), para analizar los datos utilizamos modelos de regresión logística multinivel. Los modelos multinivel toman en consideración y permiten dar cuenta de la estructura jerárquica de los datos (Snijders y Bosker, 2012), así como manejar datos en los que las observaciones no son independientes entre sí. El experimento incluyó 57 sujetos que jugaron cada uno durante 20 rondas en interacciones *one-shot*, es decir, en juegos de una sola ronda por vez. Por lo tanto tenemos 1140 observaciones (que corresponden al Nivel 1) anidadas en 57 sujetos (Nivel 2). Dada la estructura jerárquica, no puede asumirse que las observaciones que pertenecen a un mismo sujeto sean independientes entre sí. Sujetos diferentes pueden tener disposiciones comportamentales idiosincráticas. Las condiciones o tratamientos experimentales son variables de Nivel 2. Como método de estimación usamos Cadenas de Markov de Monte Carlo (MCMC), ya que este procedimiento arroja estimaciones menos sesgadas que los métodos de estimación aproximada estándar. Además, MCMC calcula el desvío estadístico, lo que nos permite utilizar un test formal de Chi cuadrado para comparar modelos estadísticos anidados. La diferencia en el desvío estadístico de dos modelos anidados sigue aproximadamente una distribución de chi-cuadrado con grados de

libertad iguales a la diferencia en el número de parámetros de los modelos comparados (Hox, 1995). Valores significativos indican que el modelo con el desvío menor es una mejora con respecto al modelo con mayor desvío. Como variable dependiente tomamos la probabilidad de elección de *in*.

### 6.1. Estadística descriptiva

En el tratamiento base los jugadores en el rol B eligieron *in* un 83.5% de las veces, mientras que los C lo hicieron en un 75.6% de las veces. En el TA, el porcentaje de los jugadores A que eligieron *in* fue de 53.3% (ver figura 7 abajo). El porcentaje de jugadores en el Rol B que eligió *in* descendió, en relación al TB, a 80.9% (ver figura 8), mientras que el porcentaje de los jugadores en el rol C que eligió *in* descendió a 68.6% (ver figura 9).

La estadística descriptiva va en la dirección predicha por las hipótesis 2 y 3, dado que los jugadores en los roles B y C eligieron *in* menos en el TA que en el TB. Por lo tanto, podemos enunciar que los datos confirman la hipótesis 1, ya que se dan tratos corruptos a altos niveles, independientemente de los tratamientos.

Figura 7. Porcentaje de participantes A que eligieron *in*.

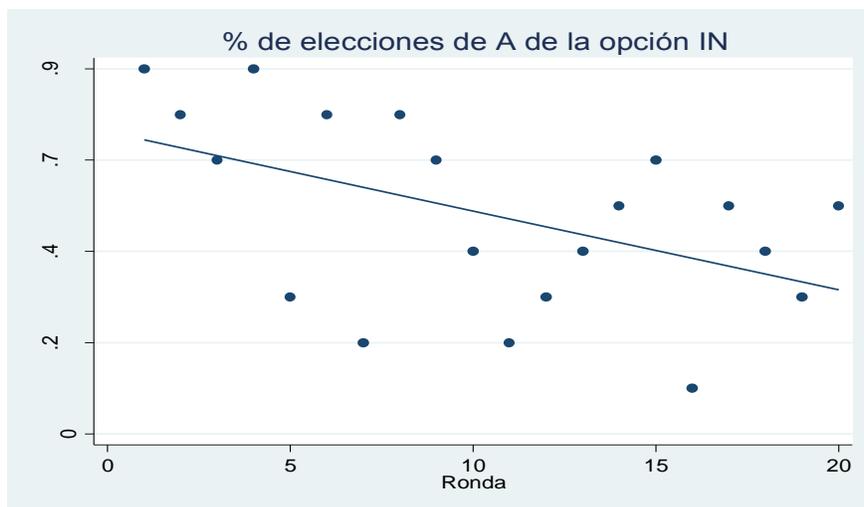


Figura 8. Elecciones de B en función de ronda y tratamiento.

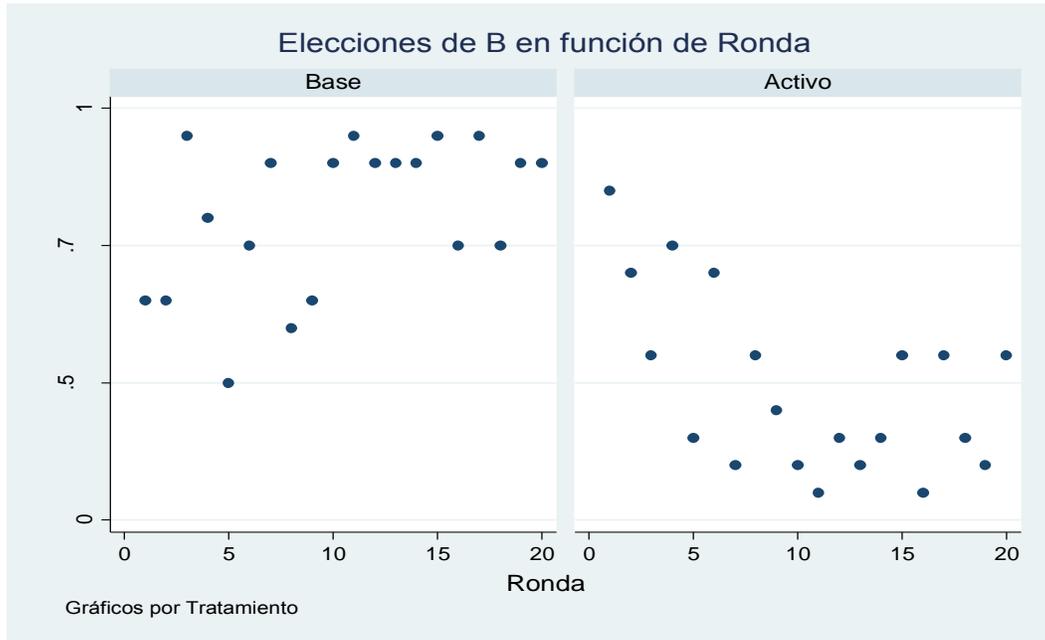
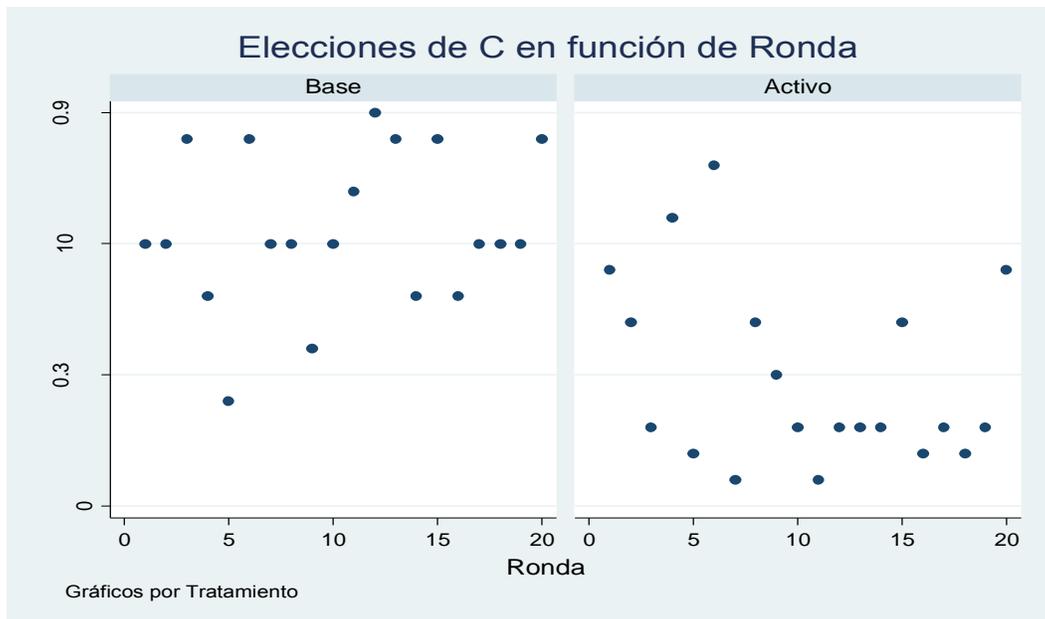


Figura 9. Elecciones de C en función de ronda y tratamiento.



6.2. Análisis estadístico

Los modelos multinivel se evaluaron utilizando el software *MLwiN* (Rasbash et al., 2009) a través del comando *runmlwin* en Stata (Leckie y Charlton, 2011), a menos que se indique lo contrario. Primero evaluamos modelos nulos o vacíos (i.e., modelos sin predictores). Siguiendo la práctica estándar, a partir de los modelos vacíos evaluamos medidas de correlación o agrupamiento de los datos. A continuación ajustamos modelos anidados con intercepto aleatorio (*random intercept*) y variables predictoras. Todos los predictores son efectos fijos, excepto el intercepto. También incluimos un efecto fijo para la variable Ronda en todos los modelos. Dado que la variable dependiente era binaria usamos regresión logística. Para el cálculo del coeficiente de partición de la varianza (CPV) para efectos aleatorios del nivel superior (en nuestro caso del nivel del individuo) es una práctica común en los modelos multinivel de respuesta binaria recurrir a la formulación de la variable latente del modelo logístico multinivel (Goldstein, 2010), en cuyo caso una CPV se calcula en términos de la propensión de un individuo a elegir *in*. En este caso la variación del nivel 1, es decir del nivel de la decisión se estandariza a la varianza logística  $\pi^2/3 \approx 3.29$  (sobre este punto ver Leckie & Charlton, 2011). Para evaluar la significatividad de los coeficientes individuales podemos recurrir a las estimaciones de los parámetros usando el criterio de Wald. Según el criterio de Wald si la razón entre el coeficiente y su error estándar excede 1.96 entonces podemos rechazar la hipótesis nula para ese parámetro en particular.

### **6.3. Resultados de las regresiones**

Para testear las hipótesis medimos el nivel de corrupción tomando como variable dependiente la proporción de elecciones de *in*. Usando modelos con intercepto aleatorio permitimos que la probabilidad de elegir *in* varíe entre sujetos, pero suponemos que los efectos de las variables explicativas son los mismos para cada sujetos, es decir, son efectos fijos.

#### *Variación a nivel individual*

El grado de agrupamiento en los datos se puede resumir en el coeficiente de correlación intra-clase (CIC), y el coeficiente de partición de la varianza (CPV). Normalmente, la fórmula para calcular la CIC es  $CIC = \text{Varianza entre grupos} / (\text{Varianza entre grupos} + \text{Varianza dentro del grupo})$ . En nuestros modelos la CIC y la CPV coinciden y miden la proporción en la variación de la probabilidad de elegir *in* que depende del nivel del individuo.

#### *6.3.1 Efectos de tratamiento*

Resumimos los resultados más importantes de nuestro experimento en las figuras 8 y 9. En la tabla 1 reportamos los modelos con intercepto aleatorio que utilizamos para testear diferencias de tratamiento. El primer resultado importante proviene de considerar un modelo que incluye una variable *dummy* para Rol, y puede ser expresado de la siguiente manera:

**Resultado 1:** La probabilidad de que los jugadores en el rol C elijan *in* es significativamente menor que para los jugadores en el rol B, *independientemente del tratamiento*.

**Respaldo:** La tabla 1 ofrece respaldo al resultado 1. Para testear si era más probable que los jugadores C eligieran *in*, comparamos el Modelo 2, que incluye una variable controlando por

el efecto de ronda, con un modelo en el que incluimos una variable *dummy* para el Rol C (tomando al Rol B como categoría de referencia). El modelo con Rol C resultó mejorar el ajuste del modelo (Diferencia en desvío = 10.62,  $gl=1$ ,  $p=0.0018$ ). Parece evidente que, tal como habíamos hipotetizado, ser el último jugador tiene un efecto sobre las decisiones. Presumiblemente, el hecho de que los pagos finales recaigan en última instancia sobre los hombros de C pueda incrementar su sentido de responsabilidad y hacer que la externalidad hacia A esté en su foco de atención. También puede ser el caso que los jugadores B hayan percibido la opción *in* como una delegación hacia C. Varios estudios (ver por ejemplo Bartling & Fischbacher, 2012) muestran que la responsabilidad por una decisión injusta puede ser efectivamente desplazada hacia otro jugador en el rol de delegado. Los jugadores B pueden haber anticipado que al pasar el juego a C (i. e., delegando la decisión a C), podrían evitar la responsabilidad por una acción injusta. Nótese que hay una importante diferencia entre los estudios de delegación y el nuestro, ya que en los primeros los dictadores pueden delegar, pero también pueden elegir ellos mismos la opción injusta, mientras que en nuestro juego los jugadores B tienen que delegar necesariamente a C para obtener la ganancia más alta.

Para testear nuestra hipótesis principal, en el siguiente modelo (Modelo 3) incorporamos una variable de nivel 2 para Tratamiento (TB como categoría de referencia). En línea con nuestra hipótesis, la estadística descriptiva mostraba que la cantidad de decisiones *in* era ligeramente superior en el TB. A pesar de eso, la inclusión de la variable para tratamiento no mejoró el ajuste del modelo (Diferencia en desvío = 0.29,  $gl = 1$ ,  $p = 0.59$ ). Este resultado va en contra de las hipótesis 2 y 3: para los roles B y C no hay diferencias estadísticamente significativas de tratamiento en la propensión a elegir *in*. Por lo tanto

debemos rechazar nuestra hipótesis principal, ya que no hay una diferencia principal de tratamiento, lo que sugiere que los jugadores en los roles B y C no se vieron afectados por la presencia de un jugador A que pudiera tomar un rol activo. Este resultado coincide con la literatura experimental sobre corrupción (ver Abbink et al., 2002; Barr & Serra, 2009) en que los jugadores tienden a enfocarse en la ganancia que pueden obtener de la transacción en detrimento del daño causado a terceros.

La CIC es 0.56, lo que significa que la correlación esperada entre dos decisiones del mismo individuo es 0.56. La CIC es una medida de la correlación de los datos en el mismo individuo y puede tomar valores en el intervalo  $[0, 1]$ . El valor 0.56 indica que la propensión a elegir *in* varía fuertemente entre individuo, lo que sugiere que existe cierta heterogeneidad individual que no es capturada por nuestros modelos.

**Tabla 1.** Estimaciones de MCMC de los coeficientes de una regresión logística multinivel. Interacciones de Tratamiento con una variable dummy para Rol C que toma valor 1 si el jugador fue asignado al Rol C y 0 en caso contrario. Modelamos la probabilidad de los jugadores de elegir *in* ( $in = 1$ ;  $out = 0$ ).

	Modelo Nulo	Modelo 1	Modelo 2	Modelo 3	Modelo Final
	Coef. (ES)	Coef. (ES)	Coef. (ES)	Coef. (ES)	Coef. (ES)
Intercepto	2.000 (0.337)*	1.694 (0.401)*	2.149 (0.467)*	2.331 (0.597)*	2.107 (0.531)*
Ronda		0.037 (0.022)	0.044 (0.023)	0.043 (0.024)	0.046 (0.024)
Rol C			-0.918(0.281)*	-0.920 (0.288)*	-0.709 (0.339)*
Tratamiento				-0.373 (0.678)	0.035 (0.747)
Trat * Rol C					-0.625 (0.619)
Parte Aleatoria					
Varianza Nivel 2	4.200 (1.443)				4.783 (1.749)
CIC	0.56				0.60
Desvío	381.47	379.13	<b>368.51</b>	368.22	368.16
N° de obs.	545	545	545	545	545

\*Valor absoluto de (Coeficiente/ES) mayor a 1.96.

Nota: La diferencia en desvíos en negrita indica que el modelo representa una mejora con respecto al modelo anterior.

En una segunda ola de análisis extendimos los análisis incorporando las atribuciones de amabilidad tal como surgían de las respuestas de los sujetos a la escala de amabilidad.

6.3.2. *Análisis de atribuciones de amabilidad.* ¿En qué medida las decisiones de los jugadores fueron sensibles a la amabilidad que percibieron en las acciones de los otros jugadores?

Para investigar el comportamiento de los jugadores B y C corrimos modelos similares a los anteriores, incorporando las evaluaciones que realizaron B y C de la amabilidad de A. A manera de control, primero investigamos si las atribuciones de amabilidad eran realmente explicadas por las acciones de los otros jugadores.

**Resultado 2:** Las acciones de los sujetos explican en gran medida las atribuciones de amabilidad.

**Respaldo:** Las tablas 5, 6 y 7 en el Apéndice 3 informan sobre las atribuciones de amabilidad realizadas por los sujetos. En el caso de los jugadores en los roles B y C, el modelo 2 de la Tabla 5 y el modelo 2 de la Tabla 6 muestran que la acción de A mejoró sustancialmente el ajuste del modelo ( $p < 0.001$  en ambos casos).

Lo mismo vale para la evaluación de C de la amabilidad de B. Un test formal de Chi-cuadrado comparando los modelos 1 y 2 muestra que incluir la variable que mide la amabilidad de B incrementó significativamente el ajuste del modelo (Dif. en desvío=82.93,  $gl=1$ ,  $p < 0.001$ ).

Estos últimos resultados nos sirven de control para ver si nuestro método fue exitoso en capturar la interacción entre las acciones y los juicios de los sujetos. Naturalmente, nos interesaba principalmente estudiar cómo las percepciones de amabilidad podían afectar el comportamiento en nuestro juego. Encontramos que la forma en la que los sujetos percibieron la amabilidad de A (i. e., qué tan amable lo encontraron), se vio correlacionada con las decisiones que tomaron. Nuestro diseño no nos permitía comparar entre condiciones qué tan amable era percibido el jugador A. Sin embargo, sí pudimos comparar si la percepción de la amabilidad de A tuvo una influencia en los jugadores en los roles B y C. La Tabla 2 informa sobre la influencia que tuvo la amabilidad percibida de A en la elección de la acción *in* por parte de B y C.

Los datos confirman el resultado siguiente:

**Resultado 3:** La probabilidad de que B y C elijan *in* aumenta a medida que perciben mayor amabilidad por parte de A.

**Respaldo:** Este resultado proviene de evaluar un modelo al que se le incluyó una variable de nivel 1, que mide la amabilidad de A (ver tabla 2 abajo). El coeficiente asociado a la variable Amabilidad-A es positivo y significativo (coef=0.758, ES = 0.234). La comparación entre los modelos 1 y 2 muestra que la inclusión de la variable que mide la amabilidad de A aumentó significativamente el ajuste del modelo (Diferencia en desvío=4.3,  $g^2=1$ ,  $p<0.05$ ). Este resultado es desconcertante desde el punto de vista de las teorías de la reciprocidad basadas en la intención. De hecho, parece contradecir sus premisas más básicas. El Jugador B tiene en su espacio de estrategias una forma fácil de reciprocitar la confianza de A, que consiste en elegir *out*. En efecto, en relación a lo que los jugadores pueden ganar, cuando B elige *out* se produce un resultado equitativo, que además es el resultado más eficiente en términos de pagos conjuntos para A y B. Lo que resulta desconcertante es que la amabilidad de A resultó ser significativa, pero en la dirección opuesta a lo que habíamos predicho en nuestra hipótesis. Nótese que nuestra hipótesis se desprende directamente de las teorías de la reciprocidad basadas en la intención. ¿Cómo podemos darle sentido a este resultado? En principio, parece coincidir con resultados previos en la literatura sobre corrupción y externalidades: en presencia de una externalidad débil, los jugadores tienden a reciprocitar a los compañeros corruptos y son indiferentes al daño que se pueda producir a terceros (ver Abbink et al., 2002; Abbink, 2005; Barr & Serra 2009; Engel & Zhurakhovska 2012).

Estos últimos resultados indican que pareciera haber un desacople entre las atribuciones de amabilidad y el comportamiento de los sujetos, lo que exige una comprensión más profunda del papel de las intenciones en nuestro juego.

**Tabla 2.** Estimación usando MCMC de una regresión logística multinivel con una variable de nivel 1 que mide la percepción de amabilidad atribuida a la acción de A.

	Modelo Nulo	Modelo 1	Modelo 2	Modelo 3
	Coef. (ES)	Coef. (ES)	Coef. (ES)	Coef. (ES)
Intercepto	1.88 (0.54)*	1.48 (0.75)*	2.39 (0.78)*	-2.09 (1.47)
Ronda		0.042 (0.044)	0.046 (0.045)	0.029 (0.045)
Rol C			-1.43(0.527)*	-1.678 (0.580)*
Amabilidad A				0.758 (0.234)*
Parte Aleatoria				
Varianza Nivel 2	5.01 (3.19)			6.15 (3.84)
CIC	0.60			0.65
Desvío	131.32	131.56	<b>123.46</b>	<b>119.16</b>
Nº de obs.	178	178	178	178

\*Valor absoluto de (Coeficiente/E. S.) mayor a 1.96.

Nota: La diferencia en desvíos en negrita indica que el modelo representa una mejora con respecto al modelo anterior.

Sin embargo, dado que el resultado anterior se refiere tanto a los jugadores en los roles B y C, tenemos que ver si el resultado difiere entre roles. En la tabla 3 reportamos modelos con intercepto aleatorio para analizar el efecto de la percepción de la amabilidad de A sobre la

acción de B (sobre la manera de estimar estos modelos usando cuadratura adaptativa (*adaptive quadrature*) ver Rabe-Hesketh et al., 2002). El algoritmo IGLS (*Iterative generalized least squares*) no consiguió proveer estimaciones para la varianza del nivel del sujeto. Esto puede ser problemático para estimar modelos con MCMC, ya que MCMC utiliza la estimación de IGLS como valores iniciales. Por esta razón decidimos estimar modelos basados en el algoritmo de Stata que utiliza el método de cuadratura adaptativa (*adaptive quadrature*). También estimamos modelos de efectos aleatorios y de efectos mixtos usando los comandos `xtlogit` y `xtmelogit` de Stata. Todos los abordajes nos arrojaron resultados muy similares. Los análisis, en efecto, confirman que la probabilidad de que el jugador en el rol B elija *in* aumenta a medida que percibe mayor amabilidad de A (ver tabla 3 más abajo;  $z = 1.97, p = 0.049$ ). Lo que este resultado demuestra es que los jugadores en el rol B tuvieron en cuenta la amabilidad de A, pero actuaron de manera exactamente opuesta a lo esperado. En este punto es necesario realizar una advertencia con respecto a este resultado, ya que hay evidencia bien establecida en el campo de la psicología moral y social, de que la conducta y el juicio no siempre se correlacionan (Nisbett & Wilson, 1977; Haidt, 2001). Esto quiere decir que el razonamiento explícito y el juicio no siempre van de la mano con la conducta manifiesta.

Nuestros resultados coinciden con una serie de trabajos (ver Dufwenberg & Gneezy, 2000; Charness & Rabin, 2002; Offerman, 2002), que utilizan juegos simples de dos jugadores para evaluar reciprocidad. En estos juegos, el jugador A tiene una opción de salida a la que puede renunciar o pasarle el juego a B. Los autores le presentan al jugador A diferentes opciones de salida, o ninguna. Es claro que la reciprocidad de B debería aumentar a medida que aumenta el monto al que renuncia A en la opción de salida para pasarle el juego

a B. Sin embargo, encuentran que el motivo de “reciprocidad positiva” es muy débil, y que incluso va en la dirección opuesta (véase particularmente Charness & Rabin, 2002, p. 843). Por otra parte, Charness & Rabin llaman la atención sobre el hecho de que el comportamiento del respondedor en los Juegos de Confianza no puede distinguirse del comportamiento del Dictador en un juego de Dictador, lo que parece indicar que no es la acción inicial del primer jugador la que induce motivos de reciprocidad en el segundo (para un resultado en este dirección ver también Dufwenberg & Gneezy, 2000).

Nótese también que en el tratamiento base podría existir un sentimiento de merecimiento, ya que B y C reciben un pago por hacer realmente algo, mientras que A no toma ninguna decisión. En varios trabajos se ha mostrado que los participantes son sensibles a este tipo de razones. Por ejemplo, Cherry y otros (2002) encuentran que en juegos de dictador cuando los dictadores tienen que ganar previamente el dinero que luego deben distribuir, el 95 se niega a compartir con el receptor. Aunque este motivo posiblemente no tenga ninguna influencia en el TA.

Sin duda, sin embargo, parece que los incentivos económicos que presenta la acción *in* para B y C fueran lo suficientemente elevados como para contrarrestar un efecto débil que pudiera tener la amabilidad de A.

**Resultado 4:** La probabilidad de que los jugadores en el rol B elijan *in* aumenta a medida que perciben mayor amabilidad en A.

**Tabla 3.** Estimación de los coeficientes para AmabAB a partir de un modelo multinivel de regresión logística con intercepto aleatorio usando el algoritmo de cuadratura adaptativa de Stata. La variable AmabAB representa la evaluación que hizo B de la amabilidad de A de acuerdo a una escala de siete puntos tipo Likert. La acción IN representa nuestra variable dependiente.

	Coef. (E.S.)	valor-Z	p-valor
Intercepto	-3.163 (3.182)	-0.99	0.320
Ronda	0.061 (0.078)	0.79	0.403
AmabAB	1.026 (0.522)	1.97	<b>0.049</b>
Varianza entre sujetos 8.40 (6.64)			
N° de sujetos 27			
N° de observaciones 97			

Basados en los resultados de los análisis previos tenemos que rechazar la hipótesis 2, ya que no encontramos un efecto significativo de tratamiento. Es más, el resultado 2 implica lo opuesto a la hipótesis 2b.

Para investigar si los jugadores en el rol C se vieron influenciados por la percepción de la amabilidad de A corrimos modelos similares.

**Resultado 5:** No hay efecto de la amabilidad de A sobre la acción de C en el TA.

**Respaldo:** En el modelo 2 (ver Tabla 4 abajo) la introducción de la variable AmabAC (amabilidad de A para C) no representó un incremento en el ajuste del modelo. El coeficiente individual 0.67 (ES = 0.51) no es significativo.

**Tabla 4.** Regresión logística multinivel con intercepto aleatorio estimada con MCMC. La variable AmabAC representa la evaluación de C de la amabilidad de A de acuerdo a una escala de 7 puntos de Likert. La acción *in* representa nuestra variable dependiente.

	Modelo Nulo	Modelo 1	Modelo 2
	Coef. (ES)	Coef. (ES)	Coef. (ES)
Intercepto	1.16 (0.59)*	0.951 (0.791)	2.11 (0.58)*
Ronda		0.0024 (0.059)	0.008 (0.059)
AmabAC			0.673 (0.515)
Parte aleatoria			
Varianza Nivel 2	3.87 (4.62)	4.11 (3.76)	4.29 (5.56)
Desvío	78.81	77.86	79.10
Nº de Obs.	81	81	81

\*Valor absoluto mayor a 1.96

**Resultado 6:** No hay efecto de la amabilidad de B sobre la acción de C en el TA.

**Respaldo:** Para evaluar el efecto global de la amabilidad de B, evaluamos las respuestas de los jugadores en el rol C a la escala de amabilidad. La tabla 5 muestra los modelos que incluyen la variable *Amabilidad B* como variable predictora. El coeficiente va en la dirección esperada ya que es positivo, pero no es significativamente diferente de 0 al 5 por ciento (Coeficiente=0.38, ES =0.23). Nótese que ajustar por la variable *Amabilidad B* no mejoró el modelo.

**Tabla 5.** Regresión logística multinivel con intercepto aleatorio. Estimación de Cadenas de Markov de Monte Carlo (MCMC) de los parámetros de los coeficientes. Modelos con amabilidad de los jugadores en rol B. Sólo jugadores en rol C.

	Modelo nulo	Modelo 1	Modelo 2	Modelo 3
	Coef. (e. s.)	Coef. (e. s.)	Coef. (e. s.)	Coef. (e. s.)
Intercepto	1.67 (0.38) *	1.71 (0.41) *	2.11 (0.58) *	0.42 (1.43)
Ronda		0.01 (0.03)	0.007 (0.03)	0.00 (0.03)
Tratamiento			-0.81 (0.71)	-0.96 (0.77)
Amabilidad B				0.38 (0.23)
Desvío	196.24	196.28	194.94	195.41
N° de Obs.	248	248	248	248

\*Valor absolute mayor a 1.96.

Los resultados 5 y 6 demuestran que las decisiones de C no se vieron influenciadas por la amabilidad que atribuyeron a los otros. Habíamos hipotetizado que la amabilidad de A, al elegir *in*, se vería reflejada en una menor probabilidad de que los roles B y C elijan *in*. A pesar de que la estadística descriptiva así lo sugería, ya que los jugadores en el rol C eligieron *in* menos veces en el TA (ver figura 9 arriba), el análisis estadístico que realizamos probó que la diferencia no es estadísticamente significativa.

Es posible que el motivo por el cual no encontramos un efecto se deba a que en muchas ocasiones, cuando los jugadores A eligieron *in*, lo mismo hicieron los jugadores B. De esta

manera las atribuciones de amabilidad de C hacia A y hacia B pueden haberse contrarrestado mutuamente, haciendo difícil que C pudiera ponderar la amabilidad de A y B de manera adecuada.

**Resultado 7:** B y C perciben menor amabilidad de parte de A a medida que avanzan las rondas.

**Respaldo:** En la hipótesis 4, habíamos predicho que B y C percibirían menor amabilidad de parte de A a medida que pasara el tiempo porque durante todo el juego están en el dilema de tener que honrar la confianza de A o reciprocarse mutuamente. Para analizar la hipótesis, en una nueva serie de análisis tomamos las atribuciones de amabilidad que cada sujeto realizó sobre los otros miembros de su grupo luego de enterarse de sus respectivas acciones (ver Tablas 5, 6 y 7 en el Apéndice 3). Estimamos un modelo de respuesta normal usando el algoritmo IGLS (*Iterative generalized least squares*) de MlwiN. En primer lugar tomamos como variable dependiente las respuestas que dieron B y C sobre la amabilidad de A en el TA. El modelo para la atribución de amabilidad realizada por B de las acciones de A representa una mejora en el ajuste del modelo en comparación con el modelo vacío (Diferencia en desvío=5.21,  $gl=1$ ,  $p=0.02$ ). Nótese que el coeficiente tiene el signo negativo esperado (Coeficiente= -0.062, ES = 0.027). Como habíamos predicho, los jugadores B consideran menos amables a los A a medida que pasan las rondas. Este resultado también se repite para las atribuciones de amabilidad de los jugadores C (Diferencia en desvío = 6.20,  $gl=1$ ,  $p=0.02$ ).

Dado que los sujetos interactúan repetidamente, es interesante preguntarse si los jugadores en el rol A aprenden a lo largo del experimento. Es posible que a mayor corrupción,

es decir, a mayor nivel de decisiones *in* de B y C, A se vuelva menos optimista con respecto a obtener un beneficio eligiendo la opción *in*. En consecuencia, podría ser el caso que los jugadores A elijan “pasar el juego” en menos ocasiones a medida que pasan las rondas (como sugiere la Figura 7), lo que además derivaría naturalmente en una disminución en el puntaje de amabilidad de A. Sin embargo, los análisis realizados no confirman esta intuición. En un análisis similar a los realizados anteriormente tomamos como variable dependiente la probabilidad de A de elegir *in*, y la regresamos en función de la variable ronda. Si bien el coeficiente resultó con el signo esperado, no alcanzó significatividad al 5% (coeficiente = -0.05, ES = 0.03).

En pocas palabras, los resultados confirman que no hay un efecto principal de nuestra manipulación sobre la proporción de la elección de *in*. Sin embargo, incluir una medida de la percepción de la amabilidad de A revela que las respuestas de los sujetos en los diferentes tratamientos están parcialmente moderadas por las atribuciones de amabilidad. Por lo tanto, el comportamiento individual es sensible a la forma en la que los sujetos juzgan la bondad de los demás. No obstante, el efecto de las percepciones de amabilidad va en la dirección opuesta a la esperada.

## 7. Consideraciones finales

¿Qué hemos aprendido de nuestro estudio acerca de las preferencias sociales y la corrupción?

El objetivo principal del presente trabajo consistió en evaluar si la presencia de un jugador activo, que puede señalar intenciones amables y que además sufre las externalidades negativas generadas por la corrupción, podía resultar en que los jugadores B

y C eligieran en menor medida la acción *in*. Para ello, consideramos un juego de corrupción de tres jugadores en el que modelamos explícitamente la obligación fiduciaria entre el jugador A y el jugador B. Además evaluamos si las atribuciones de amabilidad podían explicar el comportamiento en nuestro juego. En primer lugar, la intuición que nos llevó a realizar el presente estudio fue que la presencia de un jugador activo que por medio de su acción pudiera señalar intenciones amables, debía tener un efecto disuasivo sobre B y C, disminuyendo la probabilidad de que B y C elijan la opción corrupta. Sin embargo, los datos no confirman esta afirmación. La presencia de un jugador A que actúa como Principal Benévolo no tiene un impacto significativo en las decisiones de los otros jugadores de elegir *in* o *out*. Aún más desconcertante es el hecho de que los jugadores en el rol B eligieron la opción que maximizaba su ganancia aun cuando percibieron mayor amabilidad de parte de A.

¿Qué implican nuestros resultados para la corrupción real? Los estudios previos de Abbink et al. (2002) y Barr y Serra (2009) interpretaron sus resultados como evidencia de que el mecanismo de la reciprocidad permite la consolidación de las relaciones corruptas. Sin embargo, tal como indican nuestros datos, en nuestro juego las atribuciones de amabilidad parecen estar desacopladas de la reciprocidad. A pesar de que los sujetos tuvieron en cuenta la amabilidad de los otros participantes, sus decisiones no se siguieron de allí, lo que parece socavar el mecanismo de reciprocidad propuesto por estudios anteriores. Este resultado es importante ya que desafía un supuesto importante de las teorías de la reciprocidad basadas en la intención (Rabin, 1993; Falk & Fischbacher, 2006). Es claro que la falta de respaldo a un *mecanismo de reciprocidad* entre Principal y Agente en nuestro experimento no quita que el mecanismo pueda tener un efecto más potente en un juego

repetido en el cual los participantes siempre interactúan con la misma pareja. Es un resultado conocido en la literatura experimental que la confianza y la reciprocidad aumentan en juegos repetidos (Gächter & Falk, 2002). Es posible que en juegos donde hay un incentivo monetario explícito para no reciprocitar, la reciprocidad sea más un subproducto de incentivos propios de un juego repetido (así por ejemplo en Abbink et al., 2002, y Jaquemet, 2012), mientras que en escenarios de una sola ronda, la corrupción puede explicarse a partir del mero interés egoísta.

Ya hemos mencionado que una posible explicación de este resultado aparentemente paradójico gira en torno al hecho, que ya ha sido destacado por otros autores (ver, Dufwenberg & Gneezy, 2000; Charness & Rabin, 2002; Offerman, 2002), de que la reciprocidad positiva suele ser una motivación débil en juegos de confianza y similares. Otra explicación para la falta de efecto de un mecanismo de reciprocidad entre A y B, puede deberse al hecho de que al menos una fracción de los participantes puede haber considerado que la acción de A estaba estratégicamente motivada, dado que pasando el juego a B, es posible interpretar que la motivación de A pueda estar basada en la expectativa de obtener mayores ganancias. Algunos estudios muestran que la reciprocidad entre los jugadores es mayor en juegos de confianza cuando es posible descartar este tipo de motivaciones estratégicas (ver Stanca et al., 2009). En el futuro, los estudios podrían centrarse en descartar las motivaciones estratégicas para poder diferenciar de manera adecuada el efecto de la amabilidad de otras motivaciones posibles. Como en el estudio de Stanca et alia, esto puede lograrse manipulando el conjunto de información de los jugadores.

En línea con investigaciones recientes (Gneezy et al, 2014), aunque en un contexto un tanto diferente, nuestros resultados también sugieren que la corrupción está impulsada

principalmente por la codicia o el interés propio. Recientemente Gneezy et al. (2014) han afirmado que desde la perspectiva de establecer políticas públicas es muy relevante saber si la principal motivación para el soborno es la reciprocidad o la codicia (en el sentido de la recompensa de maximización), porque si "greed drives bribery, policy interventions should focus on preventing bribes that are contingent on a certain outcome. On the other hand, if reciprocity motivates bribery, policy interventions should focus on making reciprocity more difficult, for example, by decreasing personal contact through anonymity and staff rotation."

Desde la perspectiva de diseñar políticas anti-corrupción, nuestros resultados no son buenas noticias (cfr. Jacquemet, 2012). En principio sugieren que señalar amabilidad mediante acciones no es tan sencillo como pudiera suponerse, y otras motivaciones pueden contrarrestar fácilmente el rol de la amabilidad. Es posible que identificar los mecanismos responsables del surgimiento de la corrupción sea más complejo debido a que la corrupción puede ser el resultado del trabajo conjunto de múltiples mecanismos. Sin embargo, la incompatibilidad entre nuestros resultados y las predicciones de las teorías de preferencias sociales en un juego simple como el nuestro es desconcertante.

## Referencias bibliográficas del Capítulo 2

- Abbink, K. (2005) Fair Salaries and the Moral Costs of Corruption, *Advances in Cognitive Economics*, B. N. Kokinov, Sofia, NBU Press.
- Abbink, K. (2006) Laboratory Experiments on Corruption, in Rose-Ackerman, S. ed., *International Handbook on the Economics of Corruption*. Cheltenham UK; Northampton, MA, USA: Edward Elgar, pp. 418-437.
- Abbink, K., Irlenbusch, B., Renner, E. (2002) An Experimental Bribery Game, *Journal of Law, Economics and Organization*, 18(2): 428-454.
- Abbink, K. and Danila Serra (2012) Anticorruption Policies: Lessons from the Lab, en *New Advances in Experimental Research on Corruption* (Research in Experimental Economics, Volume 15), ed. Danila Serra and Leonard Wantchekon, 77–115, Emerald Group Publishing Ltd.
- Aidt, Toke (2003) Economic analysis of corruption: a survey, *The Economic Journal*, 113, 491, 632–652.
- Akerlof, G. A. (1982) Labor Contracts as Partial Gift Exchange, *Quarterly Journal of Economics*, 97(4), 543–569.
- Alatas, V., Cameron, L., Chaudhuri, A., Erkal, N., Gangadharan, L. (2009), Subject Pool Effects in a Corruption Experiment: A Comparison of Indonesian Public Servants and Indonesian Students, *Experimental Economics*, 12(1): 113-132.
- Andvig, J. C. (2005) Experimental Research and Corruption: A Survey of Budding Research, in Transparency International (ed.), *Global Corruption Report*. Cambridge UK: Cambridge University Press, pp. 255-267.
- Banerjee, A. (1997) A theory of misgovernance, *Quarterly Journal of Economics*, 112(4), 1289–1332.
- Banerjee, A., Hanna, R., and Mullainathan, S. (2012) Corruption, *MIT working paper series*, working paper 12-08.

- Banfield, E. C. (1975) Corruption as a feature of Governmental organization, *Journal of Law and Economics*, vo. 18, n°3.
- Banuri, S. and Eckel, C. (2012) Cracking Down on Bribery, trabajo en preparación.
- Bartling B. & Fischbacher U. (2012) Shifting the Blame: On Delegation and Responsibility, *Review of Economic Studies*, 79(1), 67-87.
- Becker, G. S. (1968) Crime and Punishment: An Economic Approach, *Journal of Political Economy*, vol. 76, p. 169-217.
- Bobkova, N. & Egbert, H. (2012) Corruption investigated in the lab: A Survey of the Experimental Literature, *International Journal of Latest Trends in Finance and Economic Sciences*, vol. 2 n° 4.
- Bowles, S. & Polanía-Reyes (2012) Economic incentives and social preferences: substitutes or complements, *Journal of Economic Literature*, vol. 50 (2), p. 368-425.
- Brown, Martin, Armin Falk, Ernst Fehr (2003) Relational Contracts and the Nature of Market Interactions, *Econometrica*, 2004, 72 (3), 747-780.
- Camerer, C., Loewenstein, G. & Rabin, M. (eds) (2004), *Advances in Behavioral Economics*, New York: Princeton University Press.
- Camerer, C. & Thaler, R (1995) Anomalies: Ultimatums, Dictators and Manners, *The Journal of Economic Perspectives*, 9 (2): 201-219
- Charness, G. & Rabin, M. (2002) Understanding social preferences with simple tests, *Quarterly Journal of Economics*, vol. 117, n°3, 817-869.
- Coffman, L. (2011) Intermediation reduces punishment and reward, *American Economic Journal: Microeconomics*, 3, 77-106.
- Croson, R. & Konow, J., (2009) Social preferences and moral biases, *Journal of Economic Behavior & Organization*, 69, 201–212.
- Dal Bó, E. (2006) Regulatory Capture: A Review, *Oxford Review of Economic Policy*, vol. 22 no. 2, 203-225.

- Della Porta, D. & Vanucci, A. (2012) *The Hidden order of corruption. An Institutional Approach*, Ashgate.
- Drugov, M., Hamman, J. & Serra, D. (2014) Intermediaries in corruption: an experiment, *Experimental Economics*, 17:78–99
- Dufwenberg, M. & Gneezy, U. (2000) Measuring Beliefs in an Experimental Lost Wallet Game, *Games and Economic Behavior*, 30, 163-182.
- Dufwenberg, M. & Kirchsteiger, G. (2000) Reciprocity and wage undercutting, *European Economic Review*, 44(4), 1069-1078
- Eek, D., & Gärling, T. (2006). Prosocials prefer equal outcomes to maximizing joint outcome. *British Journal of Social Psychology*, 45, 321-337.
- Elster, J. (1989) *The Cement of Society*, Cambridge, Cambridge University Press.
- Engel C., Zhurakhovska L. (2012) Harm on an Innocent Outsider as a Lubricant of Cooperation – An Experiment, issue 2012/02, Bonn, Max Planck Institute for Research on Collective Goods.
- Falk, A. (2007) Gift exchange in the field, *Econometrica*, 75 (5), 1501-1511.
- Falk, A. & Fischbacher, U. (2006) A Theory of reciprocity, *Games and Economic Behavior*, 54(2), 293-315.
- Falk, A., Fehr, E. & Fischbacher, U. (2008) Testing Theories of Fairness - Intentions Matter, *Games and Economic Behavior* 62: 287-303.
- Falk, A. & Heckman, J. (2009) Lab Experiments are a major source of knowledge in the social sciences, *Science*, vol. 326 no. 5952, pp. 535-538.
- Fehr, E. & K. Schmidt (1999) A Theory of Fairness, Competition, and Cooperation, *Quarterly Journal of Economics*, 114: 817-868.
- Fehr, E., S. Gächter, and G. Kirchsteiger (1997) Reciprocity as a Contract Enforcement Device: Experimental Evidence, *Econometrica*, 65(4), 833–860.

- Fehr, E., G. Kirchsteiger, and A. Riedl (1993) Does Fairness Prevent Market Clearing? An Experimental Investigation, *Quarterly Journal of Economics*, 108(2), 437–459.
- Fershtman, C., Gneezy, U. y List, J. (2012) Equity aversion: social norms and the desire to be ahead, *American Economic Journal: Microeconomics* 2012, 4(4): 131–144
- Fischbacher, U. (2007) z-Tree. Zurich Toolbox for Ready-made Economic Experiments, *Experimental Economics*, 10: 171-178.
- Gambetta, Diego (2002) Corruption: an analytical map, en Kotkin, Stephen y Sajó, Andrés (Eds.) *Political corruption in transition: A skeptic's guide*, CEU Central European University Press, Budapest, pp. 33-56.
- Gächter, S. & Falk, A. (2002) Reputation and reciprocity: Consequences for the Labour relation, *The Scandinavian Journal of Economics*, 104 (1), 1-26.
- Greiner, B. (2004) An online recruitment system for economic experiments (MPRA paper).
- Goldsmith, A. (1999) Slapping the Grasping Hand: Correlates of Political Corruption in Emerging Markets, *American Journal of Economics and Sociology* 58 (4), pp. 865–883.
- Goldstein H (2010). *Multilevel Statistical Models*. Wiley: Chichester.
- Graeff, P. (2007) Why should one trust in corruption? The linkage between corruption, norms and social capital, en Lambsdorff, J. G. (2007) *The Institutional Economics of Corruption and Reform: Theory, Evidence and Policy*, Cambridge: Cambridge University Press, pp. 40-58.
- Haidt, J. (2001) The emotional dog and its rational tail: a social intuitionist approach to moral judgment, *Psychological Review*, 108 (4), 814-834.
- Hamman, J. R., Loewenstein, G. & Weber R. (2010) Self-interest through delegation: An additional Rationale for the Principal-Agent Relationship, *American Economic Review*, 100 (4), 1826-1846.
- Harstad, B. & Svensson, J. (2011) Bribes, Lobbying and development, *American Political Science Review*, 105 (1), 46-63.

- Hox, J. J. (1995) *Applied Multilevel Analysis*, TT-Publikaties, Amsterdam.
- Huntington, S. (1968) *Political Order in Changing Societies*, New Haven, Yale University Press.
- Jacquemet, N. (2012) *Corruption as betrayal: Experimental Evidence*, trabajo en preparación.
- Jain, Arvind K. (2001) Corruption: A review, *Journal of Economic Surveys*, vol. 15, num. 1.
- Kahnemann, D., Knetsch, J. and Thaler, R. (1986) Fairness as a Constraint on Profit Seeking: Entitlements in the Market, *The American Economic Review*, 76(4), pp. 728-741.
- Kaufmann, D. (2005) Myths and Realities of Governance and Corruption, In *Global Competitiveness Report 2005-06*, 81–98. World Economic Forum.
- Klitgaard, R. (1988), *Controlling Corruption*, Berkeley, CA: University of California Press.
- Lambsdorff, J. G. (2002a) Corruption and Rent-Seeking, *Public Choice*, Vol. 113:1/2, pp. 97-125
- Lambsdorff, J. G. (2002b) Making corrupt deals: Contracting in the shadow of the law, *Journal of Economic Behavior and Organization*, Vol. 48 (2002) 221–241
- Lambsdorff, J. G. & Teksoz, U. (2004) *Corrupt Relational Contracting*, en *Corruption and the New Institutional Economics*, ed. by J. Lambsdorff, M. Taube and M Schramm, Routledge: 2004
- Lambsdorff, J. G. (2007) *The Institutional Economics of Corruption and Reform: Theory, Evidence and Policy*, Cambridge: Cambridge University Press.
- Lambsdorff, J. G., Frank, B. (2010) Bribing versus Gift-Giving: An Experiment, *Journal of Economic Psychology*, 31(3): 347-357.
- Lambsdorff, J. G. (2012) Behavioral and experimental economics as a guidance to anticorruption, en Serra, D. & Wantchekon, L. (Eds.) (2012) *New advances in the experimental research on corruption*

- Leckie, G. & Charlton, C. (2012) runmlwin: A program to run the MLwiN Multilevel Modeling Software from within Stata, *Journal of Statistical Software*, vol. 52, Issue 11.
- Leff, N. (1964) Economic Development through Bureaucratic Corruption, *American Behavioral Scientist* 8 (3) (November): 8–14.
- McCabe, K., Rigdon, M. & Smith, V. (2003) Positive reciprocity and intentions in trust games, *Journal of Economic Behavior and Organization*, 52: 267-275
- Malmendier, U., te Velde, V., Weber, R. (October 2013) Rethinking Reciprocity, manuscrito en preparación para the *Annual Review of Economics*.
- Mullainathan, Sendhil (2004) Development economics through the lens of psychology. Washington, DC: World Bank.
- Neugebauer, T., J. Perote, et al. (2009) Selfish-biased Conditional Cooperation. On the Decline of Contributions in Repeated Public Goods Experiments, *Journal of Economic Psychology* 30(1): 52-60.
- Nichols, Philip M. (2004) Corruption as an Assurance Problem, *American University International Law Review* 19, no. 6, 1307-1349.
- Nisbett, R. and Wilson, T. (1977) Telling more than we can know: Verbal reports on mental processes, *Psychological Review*, vol. 85 (3), pp. 231-259.
- Offerman, T. (2002) Hurting hurt more than helping helps, *European Economic Review*, 46, 1423-1437
- Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2002) Reliable estimation of generalized linear mixed models using adaptive quadrature, *The Stata Journal*, 2, number 1, pp. 1-21.
- Rabin, M. (1993) Incorporating Fairness into Game Theory and Economics, *American Economic Review*, 83, 1281-1302.
- Rasbash, J., Charlton C., Browne W.J., Healy M., y Cameron, B. (2009). MLwiN Version 2.1. Centre for Multilevel Modelling, University of Bristol: Bristol, UK.

- Renner, E. (2004) Wie lässt sich Korruption wirksam bekämpfen?, *Vierteljahreshefte zur Wirtschaftsforschung*, 73(2): 292-300.
- Rose-Ackerman, Susan (1975) The Economics of Corruption, *Journal of Public Economics*, IV, 187-203.
- Rose-Ackerman, Susan (1978) *Corruption: A Study of Political Economy*, New York: Academic Press.
- Schleifer, A., Vishny, R. W. (1993) Corruption, *The Quarterly Journal of Economics*, Vol. 108, No. 3, pp. 599-617.
- Sen, A. K. (1967) Isolation, Assurance and the Social Rate of Discount, 81 *The Quarterly Journal of Economics*. 112, 114
- Serra, D. (2012) Combining top-down and bottom-up accountability: Evidence from a bribery experiment, *Journal of Law, Economics and Organization*, 28(3):569-587.
- Serra, D. & Wantchekon, L. (Eds.) (2012) New advances in the experimental research on corruption (Research in Experimental Economics, Volume 15), Emerald, UK.
- Schulze, G., Frank, B. (2003) Deterrence versus Intrinsic Motivation: Experimental Evidence on the Determinants of Corruptibility, *Economics of Governance*, 4(2): 143-160.
- Snijders, T.A.B., Bosker, R.J. (1999) *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, Sage, Thousand Oaks, CA.
- Sobel, J. (2005) Interdependent preferences and reciprocity, *Journal of Economic Literature*, Vol. XLIII, pp. 392–436.
- Spiller, P. (1990) Politicians, Interest groups, and regulators: A Multiple-Principle Agency Theory of Regulation, or “Let them be bribed”, *Journal of Law and Economics*, Vol. 33, No. 1, pp. 65-101.
- Stanca, L., Bruni, L., Corazzini, L. (2008) Testing Theories of Reciprocity: Do Motivations Matter? *Journal of Economic Behavior and Organization*, doi:10.1016/j.jebo.2009.04.009.

- Van Beest, I, Van Dijk, E, De Dreu, CKW, & Wilke, H. (2005) Do-no-harm in coalition formation: Why losses inhibit exclusion and promote fairness cognitions, *Journal of Experimental Social Psychology* 41: 609–617.
- van Veldhuizen, R. (2013) The Influence of Wages on Public Officials' corruptibility: a Laboratory Experiment, *Journal of Economic Psychology*, 32: 341-356.
- Vetter, Stefan (2013) Delegating decision rights for anticipated rewards as an alternative to corruption: an experiment, *European Journal of political Economy*, 31, 188-204.
- Vidal i Blanes, J., Draka, M. & Fons-Rosen, C. (2012) Revolving Door Lobbyists, *American Economic Review*, 102(7): 3731–3748.
- Warren, Mark. E. (2004) What Does Corruption Mean in a Democracy? *American Journal of Political Science*, Vol. 48, No. 2, pp. 328-343.

## Apéndice 1

### Cuestionario de Maquiavelismo Mach IV

Por favor, expresa cuán de acuerdo/desacuerdo estás con las siguientes afirmaciones (hacé un círculo sobre tu elección).

1) Nunca le digas a nadie la verdadera razón por la cual hiciste algo, a menos que hacerlo te sea de utilidad.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

2) La mejor manera de manejar a la gente es decirle lo que quiere escuchar.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

3) Uno debe actuar sólo cuando hacerlo es moralmente correcto.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

4) La mayoría de las personas son básicamente buenas y amables.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

5) Lo menos riesgoso es presuponer que todas las personas tienen un costado malicioso y que saldrá a relucir ni bien se les presente la primera oportunidad.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

6) La honestidad es siempre la mejor política.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

7) No existe ninguna excusa para mentirle a otra persona.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

8) En general, las personas no se esfuerzan a menos que se las obligue a hacerlo.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

9) En definitiva, es mejor ser humilde y honesto que ser importante y deshonesto.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

10) Cuando le pedís a alguien que haga algo por vos, es mejor decir las verdaderas razones por las que lo querés, que dar razones que tengan más peso.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

11) La mayoría de las personas exitosas llevan vidas limpias y moralmente intachables.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

12) Quien confía completamente en cualquiera, busca problemas.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

13) La mayor diferencia entre la mayoría de los criminales y el resto de la gente es que los criminales son lo suficientemente estúpidos como para dejarse atrapar.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

14) La mayoría de las personas son valientes.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

15) Es sensato adular a personas importantes.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

16) Es posible ser bueno en todos los aspectos.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

17) El Sr. Barnum estaba equivocado cuando dijo que nace un idiota por minuto.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

18) Es difícil salir adelante sin tomar algún que otro atajo.

Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

19) A las personas que padecen una enfermedad incurable se les debería dar la posibilidad de decidir si quieren que se ponga fin a su vida sin dolor.

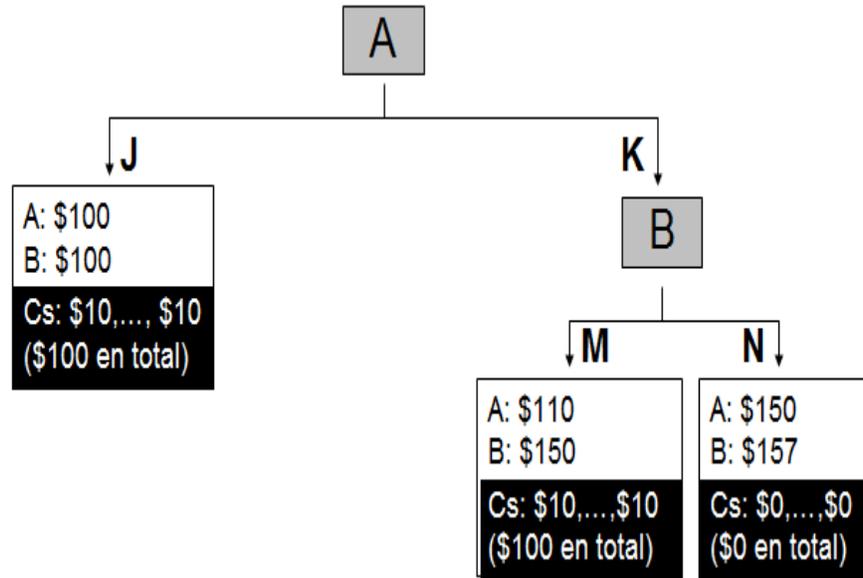
Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

20) La mayoría de la gente olvida más fácilmente la muerte de sus padres que la pérdida de sus bienes.

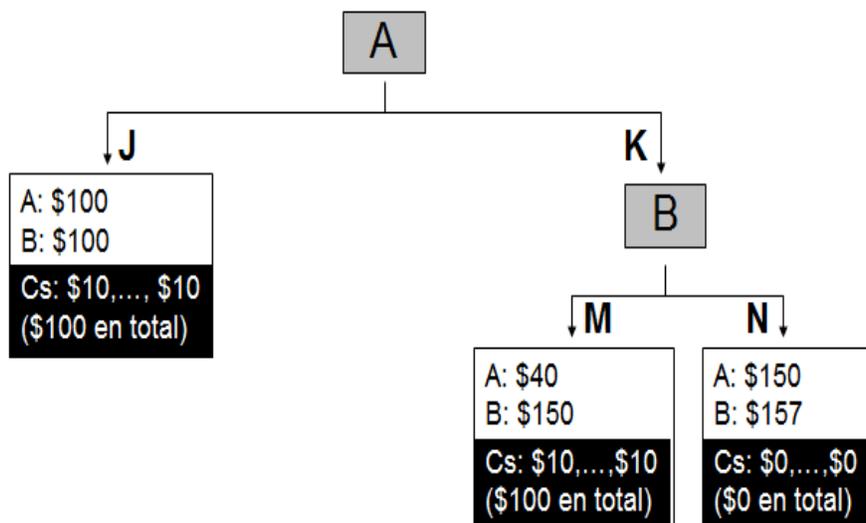
Fuertemente en desacuerdo – En desacuerdo – Neutral – De acuerdo – Fuertemente de acuerdo

## Apéndice 2

Árbol de decisión del Escenario control 6:



Árbol de decisión del Escenario control 7:



### Apéndice 3

**Tabla 6.** Estimación de los coeficientes de regresión por medio de IGLS (*Iterative generalized least squares*) de MlwiN. Atribuciones de amabilidad de B hacia A como variable dependiente.

Variable dependiente: Amabilidad de A			
	Modelo Nulo	Modelo 1	Modelo 2
Intercepto	4.500 (0.191)	4.499 (0.197)	2.520 (0.149)
Ronda		-0.062 (0.027)*	0.010 (0.013)
Acción A			3.689 (0.154)*
Dif. en desvío		<b>5.219</b>	<b>255.293</b>
<i>Parte aleatoria</i>			
Varianza nivel 2	0.319 (0.273)	0.396 (0.287)	0.268 (0.114)
Varianza nivel 1	4.361 (0.495)	4.180 (0.475)	0.938 (0.107)
N° de obs.	180	180	180

Nota: La diferencia en desvíos en negrita indica que el modelo representa una mejora con respecto al modelo anterior.

\* Indica  $p < 0.05$  (test de dos colas)

**Tabla 7.** Estimaciones de los coeficientes con IGLS (*Iterative generalized least squares*) de MlwiN. La variable dependiente es la atribución de amabilidad de C hacia A.

Variable dependiente: Amabilidad de A			
	Modelo Nulo	Modelo 1	Modelo 2
Intercepto	4.527 (0.165)	4.521 (0.171)	2.421 (0.135)
Ronda		-0.071 (0.028)*	-0.001 (0.011)
Acción A			3.900 (0.127)*
Dif. en desvío		<b>6.208</b>	<b>317.862</b>
<i>Parte Aleatoria</i>			
Varianza nivel 2	0.009 (0.208)	0.095 (0.221)	0.266 (0.101)
Varianza nivel 1	4.818 (.548)	4.574 (0.521)	0.657 (0.075)
N° de obs.	180	180	180

Nota: La diferencia en desvíos en negrita indica que el modelo representa una mejora con respecto al modelo anterior.

\* Indica  $p < 0.05$  (test de dos colas)

**Table 8.** Estimaciones de los coeficientes con IGLS (Iterative generalized least squares) de MIwiN.

	Variable dependiente Amabilidad de B				
	Modelo Nulo	Modelo 1	Modelo 2	Modelo 3	Modelo 4
Intercepto	5.19 (0.140)*	5.192 (0.139)*	3.792 (0.195)*	3.748 (0.227)*	3.896 (0.242)*
Ronda		0.002 (0.012)	-0.004 (0.010)	-0.004 (0.010)	-0.002 (0.010)
Acción B			1.662 (0.168)*	1.661 (0.168)*	1.484 (0.200)*
Tratamiento				0.100 (0.269)	-0.409 (0.412)
Acción B*Trat					0.606 (0.376)
Dif. en Desvío		0.022	<b>82.931</b>	0.139	2.558
<i>Parte Aleatoria</i>					
Varianza nivel 2	0.806 (0.204)	0.808 (0.204)	0.806 (0.190)	0.805 (0.190)	0.769 (0.183)
Varianza nivel 1	1.277 (0.117)	1.277 (0.117)	0.919 (0.084)	0.919 (0.084)	0.917 (0.084)
N° de obs.	296	296	296	296	296

Nota: La diferencia en desvíos en negrita indica que el modelo representa una mejora con respecto al modelo anterior.

\* Indica  $p < 0.05$  (test de dos colas)

## Referencias Bibliográficas completas

- Abbink, K. (2005) Fair Salaries and the Moral Costs of Corruption, *Advances in Cognitive Economics*, B. N. Kokinov, Sofia, NBU Press.
- Abbink, K. (2006) Laboratory Experiments on Corruption, in Rose-Ackerman, S. ed., *International Handbook on the Economics of Corruption*. Cheltenham UK; Northampton, MA, USA: Edward Elgar, pp. 418-437.
- Abbink, K., Irlenbusch, B., Renner, E. (2002) An Experimental Bribery Game, *Journal of Law, Economics and Organization*, 18(2): 428-454.
- Abbink, K. and Danila Serra (2012) Anticorruption Policies: Lessons from the Lab, en *New Advances in Experimental Research on Corruption* (Research in Experimental Economics, Volume 15), ed. Danila Serra and Leonard Wantchekon, 77–115, Emerald Group Publishing Ltd.
- Aidt, Toke (2003) Economic analysis of corruption: a survey, *The Economic Journal*, 113, 491, 632–652.
- Akerlof, G. A. (1982) Labor Contracts as Partial Gift Exchange, *Quarterly Journal of Economics*, 97(4), 543–569.
- Akerlof, G. & Kranton, R. (2000) Economics and Identity, *The Quarterly Journal of Economics*, 115 (3), p. 715-753.
- Alatas, V., Cameron, L., Chaudhuri, A., Erkal, N., Gangadharan, L. (2009), Subject Pool Effects in a Corruption Experiment: A Comparison of Indonesian Public Servants and Indonesian Students, *Experimental Economics*, 12(1): 113-132.
- Amir, O., Ariely, D. & Mazar, N. (2008) The Dishonesty of Honest People: A Theory of Self-Concept Maintenance, *Journal of Marketing Research*, Vol. 45: 633-634.
- Andvig, J. C. (2005) Experimental Research and Corruption: A Survey of Budding Research, in Transparency International (ed.), *Global Corruption Report*. Cambridge UK: Cambridge University Press, pp. 255-267.

- Armantier, O. & Boly, A. (2013) Comparing Corruption in the Lab and in the Field in Burkina Faso and in Canada, *Economic Journal*, 123(573), p. 1168-1187.
- Banerjee, A. (1997) A theory of misgovernance, *Quarterly Journal of Economics*, 112(4), 1289–1332.
- Banerjee, A., Hanna, R., and Mullainathan, S. (2012) Corruption, *MIT working paper series*, working paper 12-08.
- Banfield, E. C. (1975) Corruption as a feature of Governmental organization, *Journal of Law and Economics*, vo. 18, n°3.
- Banuri, S. and Eckel, C. (2012) Cracking Down on Bribery, trabajo en preparación.
- Bardsley N. (2005) Experimental Economics and the Artificiality of Alteration, *Journal of Economic Methodology*, 12, 239-251.
- Barr, A., & Serra, D. (2009) The effects of externalities and framing on bribery in a petty corruption experiment, *Experimental Economics*, 12, 488-503.
- Bartels, D., & Pizarro, D. (2011) The mismeasure of morals: Antisocial personality traits predict utilitarian responses to moral dilemmas. *Cognition*, 121, 154-161.
- Bartling B. & Fischbacher U. (2012) Shifting the Blame: On Delegation and Responsibility, *Review of Economic Studies*, 79(1), 67-87.
- Becker, G. S. (1968) Crime and Punishment: An Economic Approach, *Journal of Political Economy*, vol. 76, p. 169-217.
- Benabou, R. & Tirole, J. (2004) Willpower and Personal Rules, *Journal of Political Economy*, 112 (4), 848-887.
- Berg, J., Dickhaut, J., & McCabe, K. (1995) Trust, reciprocity and social history, *Games and Economic Behavior*, 10 (1), 122–142.
- Bicchieri (2006) *The Grammar of Society. The Nature and Dynamics of social norms*, Cambridge, Cambridge University Press.

- Bobkova, N. & Egbert, H. (2012) Corruption investigated in the lab: A Survey of the Experimental Literature, *International Journal of Latest Trends in Finance and Economic Sciences*, vol. 2 n° 4.
- Bolton, G. & Ockenfels, A. (2000) ERC: A Theory of Equity, Reciprocity, and Competition, *American Economic Review*, 90 (1), 166-193.
- Bowles, S. & Polanía-Reyes (2012) Economic incentives and social preferences: substitutes or complements, *Journal of Economic Literature*, vol. 50 (2), p. 368-425.
- Brown, Martin, Armin Falk, Ernst Fehr (2003) Relational Contracts and the Nature of Market Interactions, *Econometrica*, 2004, 72 (3), 747-780.
- Camerer, C., Loewenstein, G. & Rabin, M. (eds) (2004), *Advances in Behavioral Economics*, New York: Princeton University Press.
- Camerer, C. & Thaler, R (1995) Anomalies: Ultimatums, Dictators and Manners, *The Journal of Economic Perspectives*, 9 (2): 201-219
- Charness, G. & Rabin, M. (2002) Understanding social preferences with simple tests, *Quarterly Journal of Economics*, vol. 117, n°3, 817-869.
- Christie, R., & Geis, F. L. (1970) *Studies in Machiavellianism*. Academic Press: New York.
- Coffman, L. (2011) Intermediation reduces punishment and reward, *American Economic Journal: Microeconomics*, 3, 77-106.
- Coleman, J. (1990) *Foundations of Social Theory*, Cambridge: Harvard University Press.
- Croson, R. & Konow, J., (2009) Social preferences and moral biases, *Journal of Economic Behavior & Organization*, 69, 201–212.
- Czibor, A. & Bereczkei, T (2012) Machiavellian people's success results from monitoring their partners, *Personality and Individual Differences*, 53 (3), p. 202–206.
- Dal Bó, E. (2006) Regulatory Capture: A Review, *Oxford Review of Economic Policy*, vol. 22 no. 2, 203-225.

- Dawes, C., Fowler, J., Johnson T., McElreath, R. & Smirnov, O. (2007) Egalitarian motives in humans, *Nature* 446 (7137), 794-796.
- Della Porta, D. & Vanucci, A. (2012) *The Hidden order of corruption. An Institutional Approach*, Ashgate.
- Drugov, M., Hamman, J. & Serra, D. (2014) Intermediaries in corruption: an experiment, *Experimental Economics*, 17:78–99
- Dufwenberg, M. & Gneezy, U. (2000) Measuring Beliefs in an Experimental Lost Wallet Game, *Games and Economic Behavior*, 30, 163-182.
- Dufwenberg, M. & Kirchsteiger, G. (2000) Reciprocity and wage undercutting, *European Economic Review*, 44(4), 1069-1078.
- Dufwenberg, M., Gächter, S., & Hennig-Schmidt, H. (2011) The framing of games and the Psychology of play, *Games and Economic Behavior*, 73, 459-478.
- Eek, D., & Gärling, T. (2006). Prosocials prefer equal outcomes to maximizing joint outcome. *British Journal of Social Psychology*, 45, 321-337.
- Elliott, C. S., Hayward, D. M., & Canon, S. (1998) Institutional framing: Some experimental Evidence, *Journal of Economic Behavior & Organization*, 35, 455–464.
- Engel C., Zhurakhovska L. (2012) Harm on an Innocent Outsider as a Lubricant of Cooperation – An Experiment, issue 2012/02, Bonn, Max Planck Institute for Research on Collective Goods.
- Falk, A. (2007) Gift exchange in the field, *Econometrica*, 75 (5), 1501-1511.
- Falk, A. & Fischbacher, U. (2006) A Theory of reciprocity, *Games and Economic Behavior*, 54(2), 293-315.
- Falk, A., Fehr, E. & Fischbacher, U. (2008) Testing Theories of Fairness - Intentions Matter, *Games and Economic Behavior* 62: 287-303.
- Falk, A. & Heckman, J. (2009) Lab Experiments are a major source of knowledge in the social sciences, *Science*, vol. 326 no. 5952, pp. 535-538.

- Fehr, E. & Fischbacher, U. (2003) The Nature of Human Altruism, *Nature* 425, 785-791.
- Fehr, E. & Fischbacher, U. (2004) Third-party punishment and social norms, *Evolution and Human Behavior*, 25, 63-87.
- Fehr, E. & Gächter (2002) Altruistic punishment in humans, *Nature*, 415: 137-140.
- Fehr, E., S. Gächter, and G. Kirchsteiger (1997) Reciprocity as a Contract Enforcement Device: Experimental Evidence, *Econometrica*, 65(4), 833–860.
- Fehr, E., G. Kirchsteiger, and A. Riedl (1993) Does Fairness Prevent Market Clearing? An Experimental Investigation, *Quarterly Journal of Economics*, 108(2), 437–459.
- Fehr, E. & Schmidt, K. (1999) A Theory of Fairness, Competition, and Cooperation, *Quarterly Journal of Economics*, 114: 817-868.
- Fershtman, C., Gneezy, U. y List, J. (2012) Equity aversion: social norms and the desire to be ahead, *American Economic Journal: Microeconomics* 2012, 4(4): 131–144
- Fischbacher, U. (2007) z-Tree. Zurich Toolbox for Ready-made Economic Experiments, *Experimental Economics*, 10: 171-178.
- Gambetta, Diego (2002) Corruption: an analytical map, en Kotkin, Stephen y Sajó, Andrés (Eds.) *Political corruption in transition: A skeptic's guide*, CEU Central European University Press, Budapest, pp. 33-56.
- Gächter, S. & Falk, A. (2002) Reputation and reciprocity: Consequences for the Labour relation, *The Scandinavian Journal of Economics*, 104 (1), 1-26.
- Greiner, B. (2004) An online recruitment system for economic experiments (MPRA paper).
- Goldsmith, A. (1999) Slapping the Grasping Hand: Correlates of Political Corruption in Emerging Markets, *American Journal of Economics and Sociology* 58 (4), pp. 865–883.
- Goldstein H (2010). *Multilevel Statistical Models*. Wiley: Chichester.

- Graeff, P. (2007) Why should one trust in corruption? The linkage between corruption, norms and social capital, en Lambsdorff, J. G. (2007) *The Institutional Economics of Corruption and Reform: Theory, Evidence and Policy*, Cambridge: Cambridge University Press, pp. 40-58.
- Gunnthorsdottir, A., McCabe, K., & Smith, V. (2002) Using the Machiavellianism instrument to predict trustworthiness in a bargaining game, *Journal of Economic Psychology*, 23, 49-66.
- Haidt, J. (2001) The emotional dog and its rational tail: a social intuitionist approach to moral judgment, *Psychological Review*, 108 (4), 814-834.
- Hamman, J. R., Loewenstein, G. & Weber R. (2010) Self-interest through delegation: An additional Rationale for the Principal-Agent Relationship, *American Economic Review*, 100 (4), 1826-1846.
- Harstad, B. & Svensson, J. (2011) Bribes, Lobbying and development, *American Political Science Review*, 105 (1), 46-63.
- Hertel, G., & Fiedler, K. (1994) Affective and cognitive influences in a social dilemma game, *European Journal of Social Psychology*, 24, 131–145.
- Hertel, G., & Fiedler, K. (1998) Fair and dependent versus egoistic and free: Effects of semantic and evaluative priming on the “ring measure of social values”, *European Journal of Social Psychology*, 28, 49 –70.
- Hodgson, G. M., & Jiang, S. (2007) The economics of corruption and the corruption of economics: An institutionalist perspective, *Journal of Economic Issues*, 41 (4), 1043-1061.
- Hox, J. J. (1995) *Applied Multilevel Analysis*, TT-Publikaties, Amsterdam.
- Huntington, S. (1968) *Political Order in Changing Societies*, New Haven, Yale University Press.
- Jacquemet, N. (2012) *Corruption as betrayal: Experimental Evidence*, trabajo en preparación.
- Jain, Arvind K. (2001) Corruption: A review, *Journal of Economic Surveys*, vol. 15, num. 1.

- Kahnemann, D., Knetsch, J. and Thaler, R. (1986) Fairness as a Constraint on Profit Seeking: Entitlements in the Market, *The American Economic Review*, 76(4), pp. 728-741.
- Kaufmann, D. (2005) Myths and Realities of Governance and Corruption, In Global Competitiveness Report 2005-06, 81–98. World Economic Forum.
- Klitgaard, R. (1988), *Controlling Corruption*, Berkeley, CA: University of California Press.
- Lambsdorff, J. G. (2002a) Corruption and Rent-Seeking, *Public Choice*, Vol. 113:1/2, pp. 97-125
- Lambsdorff, J. G. (2002b) Making corrupt deals: Contracting in the shadow of the law, *Journal of Economic Behavior and Organization*, Vol. 48 (2002) 221–241
- Lambsdorff, J. G. (2007) *The Institutional Economics of Corruption and Reform: Theory, Evidence and Policy*, Cambridge: Cambridge University Press.
- Lambsdorff, J. G., Frank, B. (2010) Bribing versus Gift-Giving: An Experiment, *Journal of Economic Psychology*, 31(3): 347-357.
- Lambsdorff, J. G. (2012) Behavioral and experimental economics as a guidance to anticorruption, en Serra, D. & Wantchekon, L. (Eds.) (2012) *New advances in the experimental research on corruption*.
- Lambsdorff, J. G. Taube, M., & Schramm, M. (Ed.) (2005) *The New Institutional Economics of Corruption*, New York, Routledge.
- Lambsdorff, J. G. & Teksoz, U. (2004) *Corrupt Relational Contracting*, en *Corruption and the New Institutional Economics*, ed. by J. Lambsdorff, M. Taube and M Schramm, Routledge: 2004
- Leckie, G. & Charlton, C. (2012) runmlwin: A program to run the MLwinN Multilevel Modeling Software from within Stata, *Journal of Statistical Software*, vol. 52, Issue 11.
- Leff, N. (1964) Economic Development through Bureaucratic Corruption, *American Behavioral Scientist* 8 (3) (November): 8–14.

- Liberman, V., Samuels, S. M., & Ross, L. (2004). The name of the game: predictive power of reputations versus situational labels in determining prisoners dilemma game moves, *Personality and Social Psychology Bulletin*, 30 (9), 1175–1185.
- Lindenberg, S., & Steg, L. (2007). Normative, gain and hedonic goal frames guiding environmental behavior, *Journal of Social Issues*, 63 (1), 117–137.
- McCabe, K., Rigdon, M. & Smith, V. (2003) Positive reciprocity and intentions in trust games, *Journal of Economic Behavior and Organization*, 52: 267-275
- Malmendier, U., te Velde, V., Weber, R. (October 2013) Rethinking Reciprocity, manuscrito en preparación para the *Annual Review of Economics*.
- Meyer, H. D. (1992) Norms and self-interest in ultimatum bargaining: The Prince's Prudence, *Journal of Economic Psychology*, 13: 215-232.
- Miller, S. (2010) Institutional Corruption. In Seumas Miller *The Moral Foundations of Social Institutions: A Philosophical Study* (pp. 154-175). New York: Cambridge University Press.
- Mullainathan, S. (2004) *Development economics through the lens of psychology*. Washington, DC: World Bank.
- Neugebauer, T., J. Perote, et al. (2009) Selfish-biased Conditional Cooperation. On the Decline of Contributions in Repeated Public Goods Experiments, *Journal of Economic Psychology* 30(1): 52-60.
- Nichols, Philip M. (2004) Corruption as an Assurance Problem, *American University International Law Review* 19, no. 6, 1307-1349.
- Nisbett, R. and Wilson, T. (1977) Telling more than we can know: Verbal reports on mental processes, *Psychological Review*, vol. 85 (3), pp. 231-259.
- Offerman, T. (2002) Hurting hurt more than helping helps, *European Economic Review*, 46, 1423-1437.
- Ostrom, E. (2000) Collective action and the evolution of social norms, *Journal of Economic Perspectives*, 14 (3), 137-158.

- Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2002) Reliable estimation of generalized linear mixed models using adaptive quadrature, *The Stata Journal*, 2, number 1, pp. 1-21.
- Rabin, M. (1993) Incorporating Fairness into Game Theory and Economics, *American Economic Review*, 83, 1281-1302.
- Rasbash, J., Charlton C., Browne W.J., Healy M., y Cameron, B. (2009). MLwiN Version 2.1. Centre for Multilevel Modelling, University of Bristol: Bristol, UK.
- Renner, E. (2004) Wie lässt sich Korruption wirksam bekämpfen?, *Vierteljahreshefte zur Wirtschaftsforschung*, 73(2): 292-300.
- Rose-Ackerman, Susan (1975) The Economics of Corruption, *Journal of Public Economics*, IV, 187-203.
- Rose-Ackerman, Susan (1978) *Corruption: A Study of Political Economy*, New York: Academic Press.
- Rose-Ackerman, S. (2006) Introduction and overview. En Susan Rose-Ackerman (Ed.) *International handbook on the economics of corruption*, (pp. xvi-xxxviii). Cornwall: MPG Books Ltd, Bodmin.
- Rose-Ackerman, S. (2010) The Law and Economics of Bribery and Extortion, *Annu. Rev. Law Soc. Sci.*, 6: 217–38.
- Schleifer, A., Vishny, R. W. (1993) Corruption, *The Quarterly Journal of Economics*, Vol. 108, No. 3, pp. 599-617.
- Sen, A. K. (1967) Isolation, Assurance and the Social Rate of Discount, 81 *The Quarterly Journal of Economics*. 112, 114
- Serra, D. (2012) Combining top-down and bottom-up accountability: Evidence from a bribery experiment, *Journal of Law, Economics and Organization*, 28(3):569-587.
- Serra, D. & Wantchekon, L. (Eds.) (2012) *New advances in the experimental research on corruption (Research in Experimental Economics, Volume 15)*, Emerald, UK.

- Schulze, G., Frank, B. (2003) Deterrence versus Intrinsic Motivation: Experimental Evidence on the Determinants of Corruptibility, *Economics of Governance*, 4(2): 143-160.
- Schweitzer, H. (2005) Corruption – its spread and decline. En Lambsdorff, Taube y Schramm (Ed.) *The New Institutional Economics of Corruption*, p. 40-58.
- Shleifer, A. & Vishny, R. (1993) Corruption, *The Quarterly Journal of Economics*, 108 (3), pp. 599-617
- Simpson, B. & Willer, R. (2008) Altruism and Indirect Reciprocity: The Interaction of Person and Situation in Prosocial Behavior, *Social Psychology Quarterly*, 71:37-52.
- Snijders, T.A.B., Bosker, R.J. (1999) *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, Sage, Thousand Oaks, CA.
- Spitzer, M., Fischbacher, U., Herrnberger, B., Grön, G., & Fehr, E. (2007) The neural signature of social norm compliance, *Neuron*, 56, 185-196.
- Sobel, J. (2005) Interdependent preferences and reciprocity, *Journal of Economic Literature*, Vol. XLIII, pp. 392–436.
- Spiller, P. (1990) Politicians, Interest groups, and regulators: A Multiple-Principle Agency Theory of Regulation, or “Let them be bribed”, *Journal of Law and Economics*, Vol.33, No. 1, pp.65-101.
- Stanca, L., Bruni, L., Corazzini, L. (2008) Testing Theories of Reciprocity: Do Motivations Matter? *Journal of Economic Behavior and Organization*, doi:10.1016/j.jebo.2009.04.009.
- Van Beest, I, Van Dijk, E, De Dreu, CKW, & Wilke, H. (2005) Do-no-harm in coalition formation: Why losses inhibit exclusion and promote fairness cognitions, *Journal of Experimental Social Psychology* 41: 609–617.
- van Veldhuizen, R. (2013) The Influence of Wages on Public Officials' corruptibility: a Laboratory Experiment, *Journal of Economic Psychology*, 32: 341-356.
- Vetter, Stefan (2013) Delegating decision rights for anticipated rewards as an alternative to corruption: an experiment, *European Journal of political Economy*, 31, 188-204.

Vidal i Blanes, J., Draka, M. & Fons-Rosen, C. (2012) Revolving Door Lobbyists, *American Economic Review*, 102(7): 3731–3748.

Warren, Mark. E. (2004) What Does Corruption Mean in a Democracy? *American Journal of Political Science*, Vol. 48, No. 2, pp. 328-343.

Wilson, D., Near, D. & Miller, R. (1996) Machiavellianism: a synthesis of the evolutionary and psychological literatures, *Psychological Bulletin*, 119(2):285-99.

