

# REPUTACIÓN GRUPAL ENDÓGENA

## Un Modelo Formal

### ENDOGENOUS GROUP REPUTATION A Formal Model

CARLOS RODRÍGUEZ SICKERT Y EDUARDO VALENZUELA CARVALLO  
*Pontificia Universidad Católica de Chile, Instituto de Sociología. Chile*  
crodrigs@puc.cl evalenzc@puc.cl

#### RESUMEN

Cuando la inversión en reputación individual no puede resolver los vacíos contractuales, la reputación del grupo adquiere un rol central en la consecución de la cooperación social. En este trabajo se estudia en un modelo formal, el vínculo entre la presión social, la formación de la reputación del grupo y la emergencia intergrupala de la confianza. Específicamente, en el marco de una relación de confianza modelada como un juego asimétrico, se considera la existencia de mecanismos de sanción comunitarios en el grupo cuya confiabilidad es requerida. Se demuestra que para una proporción lo suficientemente alta de agentes cooperadores dispuestos a invertir recursos en la sanción de pares no cooperadores, la estrategia óptima de un agente racional de preferencias egoístas será responder con honor a la confianza depositada y, por ende, el equilibrio bayesiano perfecto del juego es uno en el que emerge la confianza intergrupala. El tamaño de dicha proporción crítica depende negativamente de la eficiencia del mecanismo de sanción y positivamente del tamaño de los incentivos perversos que enfrenta el agente cuya confianza es requerida. Aún cuando el poder disuasivo de la presión social no sea suficiente, la confianza intergrupala puede emerger si los beneficios potenciales de la cooperación compensan el daño eventual asociado a la traición de la confianza.

#### PALABRAS CLAVE ADICIONALES

Teoría de Juegos, Confianza, Normas Sociales.

#### ABSTRACT

When investment in individual reputation cannot solve contract incompleteness, group reputation becomes crucial to achieve social cooperation. In this article we develop a formal model in which the link between social pressure, group reputation formation and between groups trust is studied. Specifically, we model a transaction which involves trust as an asymmetric game. Additionally, we consider the operation of community-enforced sanctions within the group whose trustworthiness is required. We show that for a proportion high enough of honourable agents willing to sanction non-honourable peers, the optimal strategy of a selfish rational agent is to honour trust when placed in him and, therefore, the perfect Bayesian equilibrium is one in which inter-group trust emerges. The required proportion of sanctioning agents depends negatively on the efficacy of the sanctioning technology and positively on the size of the opportunistic incentives faced by the agents whose trustworthiness is required. Even if the deterrence effect of social pressure is not strong enough, trust can emerge if the potential benefits from cooperation compensate the eventual harm associated with abused trust.

#### ADDITIONAL KEYWORDS

Game Theory, Trust, Social Norms.

## INTRODUCCIÓN

La confianza depende de contextos y situaciones particulares de acción: no se tiene confianza siempre y en toda circunstancia. Para Baron, “la virtud de la confianza es la tendencia a comportarse como si uno creyera que otros se comportarán de manera correcta y la tendencia a valorar dicho comportamiento en uno mismo y en otros. Una virtud es un tipo de norma social que aceptamos en otros y por implicación en nosotros mismos” (Baron, 1988: 411). Esta disposición se comprende habitualmente como confianza ciega (*blind trust*), que opera bajo el presupuesto de una disposición moral que obliga a las partes a comportarse de la manera esperada. La expectativa de que el otro sólo puede comportarse de forma correcta queda confinada, sin embargo, al ambiente de las relaciones personales, familiares o de estrecha amistad, que definen el ámbito de los conocidos y próximos donde la confianza se da por sentada. En sociedades de baja confianza social, esta clase de confianza personal -confianza ciega- suele estar muy extendida y provee casi la única reserva y base de la cooperación social. Por la misma razón, se procura extender lo más posible el campo de las relaciones personales a través de un mecanismo llamado “sociabilidad” que permite hacer conocido lo desconocido o próximo lo que es extraño, generalmente mediante el reforzamiento de una relación de familiaridad entre quienes originalmente no son familiares (Valenzuela y Cousiño, 2000). La sociabilidad es el esfuerzo por disolver la *anonimia* y extrañeza en una relación hasta el punto que puede ser incorporada dentro del modelo de la confianza ciega: ocurre con frecuencia en las relaciones horizontales de amistad fuertemente dotadas de un sentido doméstico y de exigencias de cercanía y familiaridad que no le pertenecen propiamente, y también en las relaciones verticales que siguen el formato de las relaciones de servidumbre doméstica y de los favores personales que aseguran la lealtad y confianza ciega entre empleadores y empleados (“personas de confianza”). En sociedades de baja confianza, ésta se entiende ante todo como familiaridad y se reconoce precisamente en la capacidad de tratar familiarmente a alguien, por su nombre o apodo, de conocer a los suyos y de solicitar favores personales. En estos contextos, la confianza no es un atributo o disposición que pueda imputarse a las relaciones entre extraños ni puede concebirse adecuadamente en un marco de impersonalidad: es una propiedad inseparable de las relaciones de proximidad y cercanía que es propia de las relaciones personales.

La sociología y la economía se interesan, sin embargo, por las relaciones de confianza más allá del ámbito familiar, en contextos de anonimidad y extrañeza que permiten fundar relaciones de cooperación más amplias. Las características de las sociedades modernas donde el intercambio social anónimo es extensivo, y las cadenas de cooperación se alargan, así lo demandan. Es, en este sentido, en el que entenderemos la confianza social como confianza entre extraños. En esta línea, Luhmann (1980) argumenta que, sin ese tipo de confianza sólo pueden ser posibles formas muy simples de cooperación e incluso la acción individual es tan sensible a la interrupción como para ser planeada, sin confianza, más allá de la inmediatez. En la disciplina económica, Arrow (1972) presenta un punto de vista similar al argumentar que casi toda transacción comercial involucra un

elemento de confianza implicando, de este modo, que la base de la actividad económica está bajo riesgo en su ausencia.

En toda relación con extraños, en efecto, prevalece el riesgo que conlleva la exposición voluntaria del agente ante otro que puede sacar provecho o ventaja de esa exposición. La presunción de que el receptor de la confianza actuará correctamente no está asegurada ni va de suyo como en el modelo de la *blind trust* y debe ser objeto de cierta fundamentación y cálculo. Como sugiere Dasgupta “uno no confía que una persona va a hacer algo porque ésta dice que lo hará. Se confía en él/ella porque, conociendo lo que se conoce de su disposición, la información que posee, su habilidad, sus opiniones disponibles y sus consecuencias, se espera que elegirá hacerlo” (Dasgupta, 1988: 50). Esto es lo que Williamson (1993) denomina confianza calculada (*calculative trust*). La protección institucional que se obtiene mediante la ley y la juridificación de las relaciones sociales suele ser un mecanismo que sostiene las transacciones y limita el riesgo contenido en las relaciones entre extraños. No obstante, la confianza, como han argumentado autores como Coleman (1990), se exige precisamente de casos en que no existe una tercera parte que garantice institucionalmente las promesas y los compromisos. El Estado, en este esquema, disminuye los espacios en que se requiere confianza y la abundancia de regulaciones y protecciones legales se considera usualmente un signo de la ausencia de confianza social, así como la inclinación a formalizar las relaciones sociales se considera algo característico de sociedades de baja confianza social.

La confianza requiere de una fundamentación que se sitúa por encima de las obligaciones morales que surgen de la esfera familiar y por debajo de las regulaciones legales que provienen de la esfera institucional. De acuerdo con la teoría de la elección racional, la emergencia de la confianza en los sistemas sociales se sostiene en la estructura de incentivos de largo plazo que las relaciones económicas involucran. Más aún, autores como Hardin (1998: 2001) argumentan que la confianza (la confiabilidad para ser más precisos) y la reputación individual son conceptos inextricablemente unidos. En las palabras de Hardin, “En las relaciones arquetípicas de confianza, la parte interesada tiene un incentivo para ser confiable, el incentivo se funda en la valoración que se le adjudica a mantener la relación en el futuro, “[es que] mis expectativas están fundadas en una comprensión de tus intereses específicos respecto a mí” (Hardin, 2001: 3). En el marco de una relación de largo plazo, la inversión en reputación individual, como se presenta en el análisis desarrollado por Dasgupta (1988) y Kreps (1990), puede producir comportamiento confiable incluso entre agentes oportunistas si existe un grupo de agentes intrínsecamente confiables (agentes que harían honor a la confianza depositada, aun cuando no existan beneficios de largo plazo).

Está fuera de discusión la importancia del rol que desempeñan los incentivos de largo plazo para invertir en reputación individual en contextos que requieren de confianza. Sin embargo, no es correcto argumentar que en estos contextos, la cooperación social sólo puede ser alcanzada a través de mecanismos basados en la inversión en reputación individual. Cuando no existen las condiciones para invertir en una reputación individual, es la reputación del grupo la que determina si la confianza va a emerger o no. Éste es el

caso, por ejemplo, de las transacciones que no son repetidas o cuando son recurrentes pero tienen lugar entre eventuales depositadores de la confianza (pacientes por ejemplo) y receptores de la confianza (médicos, siguiendo el ejemplo) que son anónimos miembros de sus comunidades respectivas. Tal como argumenta Dasgupta (1988), si agentes anónimos no pueden encadenarse en un juego repetido o la información pasada no puede ser registrada, la reputación se transforma en un bien público puro, es decir, no existe rivalidad en el consumo ni nadie puede ser excluido de su consumo<sup>1</sup>. La provisión del bien público tiene éxito cuando la reputación del grupo de los eventuales receptores de la confianza es lo suficientemente alta como para inducir a los eventuales depositadores de la confianza para transar con ellos. Greif (1997), por ejemplo, argumenta que el sistema de responsabilidad comunal (*Community Responsibility System*) que operó entre los mercaderes de la Edad Media tardía permitía sostener el intercambio interpersonal mediante la construcción de reputación grupal.

En este artículo, seguimos el enfoque de la “reputación como un bien público” y formalmente investigamos si la presión social entre los miembros de la comunidad de eventuales receptores de la confianza puede potenciar la reputación del grupo a los niveles requeridos para la emergencia de la confianza. Nos alejamos del modelo de reputación colectiva desarrollado por Tirole (y, por ende, del enfoque de Greif que se construye sobre el de Tirole) al no considerar la operación de mecanismos directos de exclusión por parte de sus clientes o mediante mecanismos internos de exclusión implementados jerárquicamente dentro de las organizaciones a las que pertenecen. En nuestro modelo, el que la confianza entre los miembros de distintos grupos emerja o no va a depender de la capacidad del grupo cuya confiabilidad es requerida para involucrar a todos sus miembros en la provisión de este peculiar bien público (la reputación grupal) mediante mecanismos de presión social (por ejemplo, a través del ostracismo). Sobre la base de este modelo, pretendemos entender por qué ciertos grupos son capaces de construir la reputación grupal requerida para que se establezcan relaciones cooperativas con miembros fuera de su comunidad y otros no lo son.

El resto del artículo está organizado de la siguiente forma. En la primera sección, desarrollamos un marco analítico, el Juego de la Reputación Grupal, para representar estructuras de interacción social que (i) requieren confianza y (ii) donde sanciones sociales dentro del grupo cuya confiabilidad es requerida puedan ser implementadas. En la segunda sección (2), primero caracterizamos el desenlace del Juego de la Reputación Grupal asumiendo que los agentes se comportan como el tradicional *homo economicus*, para luego considerar una población híbrida en la que un subgrupo de la comunidad de receptores de la confianza presenta disposiciones recíprocas.

---

<sup>1</sup> Tirole (1996) demuestra que basta que el comportamiento pasado sea imperfectamente observado para que la reputación del grupo afecte el bienestar y los incentivos del agente. En ese escenario, sólo la exclusión parcial es posible.

## UN MODELO FORMAL

Consideramos dos poblaciones disjuntas: la compuesta por los potenciales depositadores de su confianza (*DCs*) y los eventuales receptores de ésta (*RCs*). En una primera etapa, dos miembros (uno de cada población) aleatoriamente pareados participan en el Juego de la Confianza. En una segunda etapa, después de observar el juego de sus pares, los receptores de la confianza juegan el Sub-juego de la Presión Social. El juego global, que involucra tanto el Juego de la Confianza como el Sub-Juego de la Presión Social será denominado como el Juego de la Reputación Grupal.

### El Juego de la Confianza

Una relación de confianza es una estructura de interacción particular caracterizada por la presencia de:

1. *Beneficios potenciales mutuos* asociados a un desenlace cooperativo.
2. *Participación voluntaria* por parte *DC(s)*, i.e., existe una opción de salida relevante para el (los) *DC(s)*.
3. *Incentivos perversos* enfrentados por el (los) *RC(s)*.

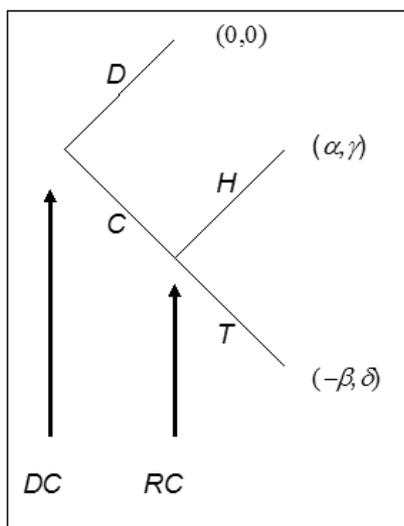
Estos tres atributos, si bien acercan nuestro enfoque al postulado en la definición de Coleman (1990) difieren de ésta al requerir que exista además una opción de salida relevante para el *DC*. Es el carácter voluntario de la exposición lo que, en nuestro modelo, activa la obligación moral específica —de naturaleza recíproca— que surge en una relación de confianza (que será modelada más adelante). En su más simple representación, un Juego de la Confianza involucra dos jugadores: el *DC* y el *RC*. En este Juego de la Confianza, el *DC* decide si participar o no en una transacción con el potencial *RC* que decide si hacer honor a la confianza que ha sido depositada en él, o traicionarla. La Figura 1 presenta el Juego de la Confianza en su forma extensiva<sup>2</sup>, donde los parámetros  $\alpha$ ,  $\beta$ ,  $\delta$ ,  $\gamma$  son positivos y, además, se cumple  $\delta > \gamma$ . Las acciones de confiar y desconfiar las denotamos por *D* y *C* respectivamente. Las acciones de traicionar la confianza y hacerle honor a ésta las denotamos por *T* y *H* respectivamente.

El primer atributo de una relación de confianza —la existencia de beneficios potenciales mutuos— está garantizada por la existencia de un desenlace, aquél en el que el *DC* confía y el *RC* responde con honor a la confianza depositada, que es *un óptimo*

---

<sup>2</sup> Una representación similar de una relación de Confianza es presentada originalmente en Dasgupta (1988) y Kreps (1990). En la literatura más reciente no existe un consenso definitivo respecto a qué juego representa de mejor forma una relación de confianza. Sin embargo, diversos autores, Williamson (1993), Snijders y Keren (1999) y James (2002), cuando se refieren a un juego de la confianza, toman la formulación inicial —o una versión simplificada— propuesta por Dasgupta y Kreps como su representación estándar.

Figura 1.  
El juego de la Confianza



*paretiano* respecto al desenlace en el que la transacción no se lleva a cabo. El segundo atributo —la voluntariedad del acto de confianza— está garantizado por la existencia de una opción de salida (la acción *D*). En este esquema no decimos que confiamos en disposiciones morales de los agentes para que no elijan acciones en nuestro desmedro, cuando la posibilidad de ejecutar esas acciones no ha sido originada por nuestra exposición voluntaria. La condición  $-\beta < 0$  garantiza que ésta sea una opción de salida relevante, pues de ejercerla alcanza una posición superior a la que se encontraría frente a una eventual traición (esto no se daría si, por ejemplo, un gángster nos instara a confiar en él con una pistola en nuestra sien). Finalmente, el tercer atributo —la existencia de incentivos perversos en el RC— está garantizada por la condición  $\delta > \gamma$  que implica que el RC alcanza una posición material superior cuando decide traicionar la confianza.

### El subjuogo de la presión social

Cada RC en el subjuogo de la presión social decide si monitorear o no las acciones de una proporción  $p \in [0,1]$  de sus pares. El proceso de monitoreo revela la acción de sus colegas en el juego de la confianza. Si se elige monitorear a un grupo de colegas, se debe decidir si castigar o no a cada uno de ellos. Información acerca del proceso de monitoreo en rondas previas del juego no está disponible. De ahí que argumentos

de inversión en reputación individual como el desarrollado por Kreps no se apliquen en este contexto.

Los costos de aplicar la sanción para un  $RC_i$  independiente de su comportamiento hacia el agente que depositó la confianza en él, están dados por  $\tau k_i + c$ ; donde  $c$  es el costo fijo asociado al proceso de supervisión y  $\tau k_i$  el costo marginal de sancionar a una proporción  $\eta$  de los colegas supervisados por él.

La desutilidad asociada al costo impuesto por la presión social está dada por  $-\lambda' l_i$  donde  $l_i$  es la proporción de colegas (de la población global) que sancionan al  $RC_i$ , es decir, si la comunidad completa decide sancionar al  $RC_i$ , él sufre una desutilidad dada por  $-\lambda'$ . Es conveniente definir  $\lambda = \lambda' \eta$ , el valor esperado de la sanción de un agente cuando la población completa de  $RCs$  ha decidido supervisar a sus colegas y la acción implementada por este agente es sancionada por toda la comunidad.

Para el juego global asumimos dos restricciones adicionales. En primer lugar, se asume  $\delta - \lambda < \gamma$ . Esto es, no es óptimo traicionar la confianza bajo el peor escenario posible para el  $RC$  que efectúa la traición: cuando todos los colegas han decidido supervisarse mutuamente y todos han decidido castigar la traición a los miembros de la comunidad de  $DCs$ . Asimismo, asumimos  $\delta - \beta < \gamma - c + \alpha$ , esto es, aun en la existencia de mecanismos de compensación potenciales, el desenlace  $(C, H)$  es *Pareto superior* al desenlace  $(C, T)$  incluso cuando los  $RCs$  deben asumir el costo fijo del proceso de monitoreo.

## ANÁLISIS

### Oportunismo en el Juego de la Reputación Grupal

La solución al Juego de la Reputación Grupal provista por la teoría de juegos estándar es inambigua, pero decepcionante. Esta predicción la expresamos en la siguiente proposición.

**Proposición 1.** *Si el Juego de la Reputación Grupal se juega entre individuos racionales motivados exclusivamente por su bienestar material y este hecho, la estructura del juego y todos los parámetros del juego son Conocimiento Común, entonces el único equilibrio perfecto de Nash en subjuegos (EPNS) del juego es  $(D, T)$ : todo  $RC$  en el que es depositada la confianza la traiciona y no supervisa a colega alguno; todo  $DC$  desconfía.*

### Prueba

En la comunidad de los  $RCs$  existen cuatro tipos de estrategia: responder con honor a la confianza y supervisar a sus pares ( $HM$ ); responder con honor a la confianza y no monitorear a sus pares ( $H$ ); traicionar la confianza y supervisar a sus pares; traicionar la confianza y no supervisar a sus pares. Los pagos esperados asociados con los tipos

de estrategia para el  $RC_i$  van a depender de la forma que adquiera la presión social. De este modo tenemos que los pagos están dados por

$$(1) \quad \begin{aligned} \pi_i^{HM} &= \gamma - \lambda l_H - \tau k_i - c \\ \pi_i^H &= \gamma - \lambda l_H \\ \pi_i^{TM} &= \delta - \lambda l_T - \tau k_i - c \\ \pi_i^T &= \delta - \lambda l_T \end{aligned}$$

donde  $l_H$  y  $l_T$  son la proporción de  $RCs$  que castigan a los colegas que responden con honor y traicionan la confianza respectivamente. Como se tiene

$$\pi_i^{HM} < \pi_i^H \wedge \pi_i^{TM} < \pi_i^T$$

ni supervisar ni sancionar van a tener lugar,  $l_H = l_T = k_i = 0$ . De este modo, la estrategia óptima en el subjuego de la presión social es no supervisar y, por ende, no sancionar a sus colegas independientemente del comportamiento en el que ellos incurran.

En la comunidad de los  $DCs$  los pagos esperados de las dos estrategias disponibles confiar ( $C$ ) y desconfiar ( $D$ ), están dados por

$$(2) \quad \begin{aligned} \pi^T &= \alpha (q_{FM} + q_F) - \beta (1 - q_F - q_{FM}) \\ \pi^D &= 0 \end{aligned}$$

Debido a que los  $DCs$  anticipan la ausencia de supervisión y sanciones, esto es,  $q_{FM} = q_F = 0$ , la utilidad esperada de confiar es  $\pi^T = -\beta$  y la estrategia óptima del  $RC$  es desconfiar. El conocimiento común de la racionalidad y el interés propio determinan la reputación del grupo, medida como la probabilidad estimada de responder con honor a la confianza, a su nivel mínimo posible, o sea, cero. El supuesto de conocimiento común implica que la reputación del grupo coincide con la confiabilidad efectiva de éste.

De este modo, dada una población homogéneamente compuesta por agentes motivados por el interés propio, la predicción de la teoría de juegos estándar es la desconfianza entre los miembros de distintos grupos. El argumento presentado en Kandori (1992), donde una extensión del teorema *folk* es presentada para justificar la emergencia de una norma social sobre la base del interés propio no es válida en nuestro modelo. Su resultado está basado en la capacidad de registrar la historia completa del juego, que bajo nuestro enfoque no está disponible.

Es importante notar que el comportamiento social en el Juego de la Reputación Grupal bajo este escenario es exactamente el mismo que en el Juego de la Confianza cuando mecanismos de sanción mutua en la comunidad de los RCs no se encuentran disponibles. En el corazón de este desenlace negativo descansa un problema de credibilidad que involucra dos componentes. En primer lugar, aun cuando un RC motivado por el interés propio estaría dispuesto a renunciar a los beneficios asociados a los incentivos perversos que enfrenta por el depósito de la confianza en su persona, su incapacidad para comprometer creíblemente el responder con honor a este acto previene la implementación del desenlace óptimo de Pareto ( $C,H$ ). En segundo lugar, la amenaza de sancionar a sus pares y de este modo disuadir su eventual comportamiento oportunista tampoco es creíble. Por lo tanto, si la población completa de RCs estuviera compuesta por agentes oportunistas, transacciones *spot* o repetidas pero sin capacidad de ser registradas que cumplieran con los atributos de una relación de confianza no tendrían lugar aunque existieran mecanismos de sanción cívicos que pudieran operar en el interior de la comunidad de RCs, pues estos no serían implementados.

### El Juego de la Reputación Grupal en Presencia de Disposiciones Recíprocas

En el problema analizado en este artículo puede argumentarse que tanto en el Juego de la Confianza como en el subjuego de la Presión Social el comportamiento de los agentes puede estar motivado por preferencias recíprocas. Annete Baier (1986) ha reconocido el papel crucial que desempeña el acto mismo de confiar en alguien que genera la obligación moral de devolver esa confianza, según el modelo del intercambio de dones (*gift exchange*) de Marcel Mauss (1925) que obliga a recibir y devolver el don que alguien ofrece. Cualquier relación con extraños se puede iniciar simplemente con un acto de confianza en las promesas y compromisos que alguien declara que se presume serán honradas en virtud de la misma confianza que se ha depositada en ellas. La probabilidad de responder adecuadamente a la confianza dada establece la confiabilidad (*trustworthiness*) de una persona. Así como el comportamiento en el Juego de la Confianza está influenciado por disposiciones recíprocas de carácter positivo; en el caso del subjuego de la Presión Social, el comportamiento de los agentes está influido por disposiciones recíprocas de carácter negativo. Se castiga el no participar de la producción de un bien público, la reputación del grupo.

En esta sección consideramos un modelo estilizado de la estructura moral de la comunidad de los (RCs). Específicamente denotamos por  $\hat{q}$  la proporción de agentes motivados por sentimientos recíprocos de modo que responderían con honor si les fuera depositada la confianza y sancionarían a aquellos colegas que traicionen la confianza depositada por los miembros de la comunidad de DCs. Bowles y Gintis (2002) se refieren a este tipo de agentes cooperadores dispuestos a sancionar a agentes no cooperadores como *reciprocadores* (si se permite el neologismo) *fuertes* (*strong reciprocators*). Se asume que el resto de la población, es decir, la proporción  $(1 - \hat{q})$  está motivada exclusivamente por el interés propio. El modelo es estilizado pues configura una población híbrida, pero

dicotómica. El agente está motivado por disposiciones recíprocas que le hacen responder con honor a la confianza y castigar a sus colegas traidores o, alternativamente, se comporta como un agente económico tradicional. En un modelo más realista, debiese considerarse un continuo de tipos. De este modo, en nuestro modelo la Naturaleza decide los dos posibles tipos que caracterizan la población de RCs: *homo reciprocans* (HR) y *homo economicus* (HE). Se asume que cuando un DC se enfrenta a un potencial RC, el primero no puede observar el tipo del segundo.

En este escenario, el desenlace del Juego de la Reputación Grupal puede ser caracterizado por el vector  $(q, p)$  donde  $q$  denota la proporción de agentes motivados recíprocamente más la proporción de agentes cuya respuesta positiva a la confianza depositada por ellos está inducida por el poder disuasivo de la presión social y  $p$  representa la proporción de DCs que efectivamente depositarían su confianza al enfrentarse a un miembro de la comunidad de RCs. En la siguiente proposición el desenlace predicho del juego está presentado para diferentes valores de  $\hat{q}$ .

**Proposición 2** *En un equilibrio Bayesiano perfecto (EBP) del Juego de la Reputación Grupal con tipos múltiple (RGM), los HR responden con honor la confianza y sancionan a los traidores. Para*

$$(3) \quad \hat{q} \geq \frac{\delta - \gamma}{\lambda}$$

el EBP del Juego RGM es  $(q, p) = (1, 1)$ , es decir, la confianza emerge en forma completa (todos los DCs depositan la confianza) y perfecta (todos los receptores responden con honor a ella); para

$$(4) \quad \frac{\beta}{\alpha + \beta} \leq \hat{q} < \frac{\delta - \gamma}{\lambda},$$

el EBP es  $(q, p) = (\hat{q}, 1)$ , esto es, la confianza emerge en forma completa pero imperfecta (la confianza es traicionada en una proporción  $(1 - \hat{q})$ ); y, finalmente, para

$$(5) \quad < \min \left\{ \frac{\delta - \gamma}{\lambda}, \frac{\beta}{\alpha + \beta} \right\},$$

el EBP es  $(q, p) = (\hat{q}, 0)$ , o sea, el desenlace del juego es la desconfianza.

### Prueba

Los resultados presentados en la proposición son válidos bajo el supuesto de que los parámetros, incluida la razón de *HR* con respecto a los *HE*, son conocimiento común. La reputación grupal, de este modo, coincide con el nivel de confiabilidad (y su disposición a sancionar a pares traidores) efectivo de la comunidad de los (*RCs*).

Si la condición expresada en la ecuación (3) se cumple, la confianza entre los miembros de los grupos de *DCs* y *RCs* emergerá en forma completa y perfecta, pues la proporción de *HRs* hace que traicionar la confianza no sea rentable para un *HE*. Los pagos esperados para un *RC* de tipo *HE* de las estrategias responder con honor a la confianza depositada y traicionarla están dados por

$$(6) \quad \begin{aligned} \pi^H &= \gamma \\ \pi^T &= \delta - \lambda \end{aligned}$$

La traición será disuadida si se cumple  $\pi^H \geq \pi^T$  condición equivalente a la condición (3).

Alternativamente, si se cumple la condición (4), traicionar la confianza sigue siendo provechoso para un *RC* de tipo *HE*. Sin embargo, la acción óptima del *DC* es confiar. Dado que en una proporción  $(1 - \hat{q})$  de las transacciones la confianza será traicionada nos referiremos a este desenlace como la emergencia completa, pero imperfecta de la confianza intergrupal.

Finalmente, si la proporción de agentes reciprocamente motivados no es lo suficientemente alta como para disuadir a los *HE* y tampoco hace rentable depositar la confianza, el desenlace es la desconfianza intergrupal.

Es importante notar que en este modelo conductista, en contraste con el análisis ortodoxo, la disponibilidad de mecanismos de presión social sí cambia el desenlace del juego. En la ausencia de castigos, la confianza completa y perfecta sólo emergería para ( $\hat{q} = 1$ ). En cambio, en la presencia de castigo, se requiere de una condición menos fuerte,

$$(7) \quad \hat{q} \geq \frac{\delta - \gamma}{\lambda}$$

para que emerja la confianza en el Juego de la Reputación Grupal. En este caso, la comunidad misma es la generadora de orden social.

### CONCLUSIONES

De los resultados del análisis de una población híbrida moralmente, se puede observar cómo una vez que una norma que sanciona el abuso de la confianza depositada por miembros de otra comunidad, su poder disuasivo puede hacer confiable al grupo como

un todo, aun cuando no todos han internalizado la norma. En este escenario el comportamiento moral es obligado por la comunidad que es capaz de reemplazar el rol de las instituciones.

Para grupos eficaces en su capacidad de establecer relaciones provechosas con miembros de otras comunidades, su confiabilidad y, por ende, su reputación es observada como un bien público que debe ser resguardado y protegido celosamente. En general operan sobre la base de la interdependencia de sus miembros y poseen un interés específico asociado a la reputación social sobre la cual descansan sus actividades. Por ejemplo, médicos que se derivan pacientes mutuamente y que dejan de hacerlo en caso que alguno falle o sujetos de microcrédito que condenan al ostracismo a agentes que no cumplan los compromisos con las organizaciones prestatarias.

En sociedades donde la confianza social es baja, en efecto, la capacidad de sancionar el comportamiento públicamente deshonoroso suele faltar por completo y la esfera pública no alcanza nunca el estatuto de una comunidad moralmente calificada. La tolerancia o indiferencia frente a conductas depredatorias, la aceptación del *free-rider* y el cierre de los grupos públicamente relevantes en torno a vínculos de lealtad bloquean permanentemente las posibilidades de cooperación en ambientes de anonimía y extrañeza.

Para que las relaciones de confianza se consoliden en el tiempo es fundamental que estén dadas las condiciones para una eficaz autorregulación por parte de los grupos cuya confiabilidad es requerida. Las instituciones deben orientarse a construir una reputación grupal y no a usufructuar del prestigio histórico cuando exista. La reputación grupal, al estar compuesta por la agregación de las reputaciones de los actores que conforman el grupo, se construye en el día a día. Para que esto sea posible, deben existir las estructuras que fortalezcan la comunidad moral de las organizaciones o agrupaciones: colegios profesionales o comités de ética dentro de las organizaciones.

## REFERENCIAS BIBLIOGRÁFICAS

- ARROW, K. (1972), "Gifts and exchanges", *Philosophy and Public Affairs*, vol. 1, n° 4, pp. 342-62.
- BAIER, A. (1986), "Trust and antitrust", *Ethics*, vol. 96, n° 2, pp. 231-60.
- BARON, J. (1998), "Trust: Beliefs and morality", en L. Putterman y A. Ben Ner (ed.), *Economics, Values, and Organization*, pp. 108-18, Cambridge University Press, Cambridge.
- BOWLES, S. y H. GINTIS (2002), Social capital and community governance, *Economic Journal*, vol. 112 n° 483, pp. 419-36.
- COLEMAN, J. (1990), *Foundations of Social Theory*, Harvard University Press, Cambridge, Massachusetts.
- DASGUPTA, P. (1988), Trust as a commodity, en D. GAMBETTA (editor), *Trust, making and breaking cooperative relations*, Blackwell, Nueva York, pp. 49-72.

- GREIF, A. (1997), On the social foundations and historical development of institutions that facilitate impersonal exchange: From the community responsibility system to individual legal responsibility in pre-modern Europe. *Stanford University, Department of Economics Working, Papers*, vol. 97(016), pp. 1-40.
- HARDIN, R. (1998), "Trust", en NEWMAN, P. (ed.), *The New Palgrave Dictionary of Economics and the Law*. Macmillan, Basingstoke, pp. 623-627.
- (2001), "Conceptions and explanations of trust", en K. COOK (ed.), *Trust in Society*, Russell Sage, Nueva York, pp. 3-39.
- JAMES, H. (2002), The trust paradox: A survey of economic inquiries into the nature of trust and trustworthiness, *Journal of Economic Behavior and Organization*, vol. 47, n° 3, pp. 291-307.
- KANDORI, M. (1992), "Social norms and community enforcement", *Review of Economic Studies*, n° 59, n°1, pp. 63-80.
- KREPS, D. (1990), "Corporate culture and economic theory", en SHEPSON, K. y J. ALT (eds.), *Perspectives on Positive Political Economy*, Nueva York, Cambridge University Press, pp. 90-143.
- LUHMANN, N. (1980) [1973], *Trust and Power*, Nueva York, Wiley.
- MAUSS, M. (1990) [1975], *The Gift: The Form and Reason for Exchange in Archaic Societies*, Londres, Routledge.
- SNIJDERS, CH. y G. KEREN (1999), "Determinants of trust", en EREV, I., D. BUDESCU y R. ZWICK (eds.), *Games and Human Behaviour*, pp. 355-385, Lawrence Erlbaum, Mahwah.
- TIROLE, J. (1996) A theory of collective reputations (with applications to the persistence of corruption and to firm quality), *Review of Economic Studies*, vol. 63, n°1, pp. 1-22.
- VALENZUELA, E. y C. COUSIÑO (2000), "Sociabilidad y asociatividad: un ensayo de sociología comparada", *Estudios Públicos*, n° 77.
- WILLIAMSON, O. (1993), "Calculativeness, trust and economic organisation", *Journal of Law and Economics*, vol. 36, pp. 487-500.

**RECIBIDO: 7/12/05**  
**ACEPTADO: 05/10/06**