

# Incentivos Reputacionales para la Autorregulación: Un Análisis Experimental<sup>1</sup>

Javier Nuñez E. and José Luis Lima R.<sup>2</sup>

## ABSTRACT

Credence good industries are generally organized as Self Regulatory Organizations (SRO). As self regulation implies regulatory capture the SRO's incentives to control quality and report low quality provision are not guaranteed. We employ experimental techniques to examine the reputation-based incentives to provide good quality provision in a credence-good environment. We find strong support for effective reputation-based incentives under specific environments. We also find that most consumers learn the structure of the market in a way that convergence towards market equilibrium is achieved, despite the credence good assumption.

## RESUMEN

Las industrias de bienes de confianza generalmente se encuentran organizadas como Organizaciones Autorreguladas (OA). Sin embargo, la autorregulación implica una situación de captura regulatoria por lo cual los incentivos de la OA para controlar la calidad provista en el mercado y denunciar evidencia de provisión de mala calidad no están garantizadas. En este trabajo analizamos, utilizando técnicas experimentales, si crear una reputación de buena calidad entre los consumidores es suficiente para que la OA tengan los incentivos para hacer su trabajo. Nuestros resultados ofrecen un sólido soporte a la posibilidad de que la OA realice su trabajo por motivos reputacionales en el caso en que es teóricamente factible. También encontramos evidencia de que los consumidores pueden aprender la estructura del mercado y ajustar su comportamiento hacia los equilibrios teóricos, y de esta manera suplir su falta de experimentación directa de la mala calidad y control.

JEL Codes: C72 C92 D83 L15 L84

Key Words : Credence Good, Self-Regulated Organization, Reputational Incentives.

---

<sup>1</sup> We thank FONDECYT-Chile for financial support. José Luis Lima also thanks financial support from a MECESUP scholarship.

<sup>2</sup> Department of Economics, University of Chile, Diagonal Paraguay 257, Santiago, Chile. Corresponding e-mails: [jnunez@econ.uchile.cl](mailto:jnunez@econ.uchile.cl), [jlima@econ.uchile.cl](mailto:jlima@econ.uchile.cl)



## 1. INTRODUCCION

La autorregulación es una solución de regulación privada que se plantea al problema de provisión de calidad en mercados donde al consumidor le es difícil reconocer la calidad del servicio tanto antes como después de recibirlo (bien de confianza, [5]).<sup>3</sup> Debido al problema informacional para el consumidor y a que en una industria autorregulada son los mismos proveedores del servicio aquellos que vigilan la calidad que se provee en el mercado (captura regulatoria) existen serias dudas sobre su validez y sobre las condiciones bajo las cuales podría funcionar.

En Nuñez [9] se argumenta que la autorregulación podría funcionar si existe interés por parte de los proveedores por crear una reputación de buena calidad, de tal manera que los consumidores confíen y compren el bien en el mercado, aminorando de esta manera el problema informacional. Nuñez [10] establece teóricamente que aún bajo las limitaciones de bien de confianza y captura regulatoria es posible, aunque no asegurable, que la Autorregulación funcione por motivos reputacionales y se requiere de un análisis empírico para poder determinarlo con certeza. En este trabajo nosotros utilizamos técnicas experimentales para determinar si el motivo reputacional es una razón válida para que funcione la Autorregulación en mercados de bienes de confianza.

Como fue señalado en un principio, en un mercado de bienes de confianza los consumidores no pueden determinar la calidad de los servicios recibidos tanto antes como después de la compra; debido a esto existe un fuerte incentivo en los proveedores de engañar a sus clientes y proveerlos de servicios de baja calidad. Sin embargo, si los clientes anticipan éste comportamiento es posible que se termine en una situación tipo “Lemon market”, solo con servicios de mala calidad provistos en el mercado. Para evitar esta situación y brindar confianza a los consumidores de que van a recibir servicios de buena calidad los proveedores pueden organizarse como una Organización Autorregulada (OA), que realiza dos acciones principales [9]:<sup>4</sup>

- a) vigila activamente la calidad provista por sus miembros; y,
- b) castiga y denuncia públicamente cualquier evidencia grave de provisión de mala calidad o fraude.

La unión de un alto nivel de vigilancia unido a la amenaza de castigo induciría normalmente a los proveedores a brindar servicios de buena calidad, lo cual es la base para la creación de buena reputación ante los consumidores. Sin embargo, los consumidores, al no poder observar la calidad que reciben, tampoco pueden inferir el nivel de vigilancia que realiza la OA, por lo que necesitan de una señal creíble que les permita inferir ambos, señal que naturalmente es el resultado del proceso de vigilancia.<sup>5</sup> De esta manera, para establecer si el motivo reputacional es suficiente para que la

---

<sup>3</sup> Ejemplo de bienes de confianza son los servicios expertos (médicos, financieros, legales) y los de reparación de vehículos, donde es el mismo proveedor el que actúa como el experto que diagnostica las necesidades del cliente.

<sup>4</sup> Ejemplos de OA pueden encontrarse en los mercados financieros (Asociación de Bancos Suizos), médicos (Colegios médicos), publicitarios (Concejo Nacional de Publicidad Autorregulada en Chile), etc.

<sup>5</sup> Por un lado, la inversión en mecanismos de vigilancia no garantiza a los consumidores de que serán efectivamente utilizados por la OA (pueden ser medidas de distracción). Por otro lado, es razonable asumir que la OA no tendrá incentivos para inventar evidencia de fraude por el daño moral que pueda causar a sus miembros y por los problemas judiciales que le podría generar.

Autorregulación funcione se requiere establecer si brinda los incentivos suficientes para que la OA realice las dos acciones principales que se esperan de ella.<sup>6</sup>

No obstante, la denuncia de fraude puede ser interpretada de dos maneras distintas por los consumidores, con implicancias distintas sobre los incentivos para la OA:

- a) los consumidores pueden interpretar la denuncia como una señal de un alto nivel de vigilancia que induce una provisión de buena calidad en el mercado, encontrándose fraude por casualidad, lo que los incentiva a participar en el mercado. Si esto ocurre la OA tiene incentivos para vigilar a sus miembros y exponer evidencia de fraude.
- b) los consumidores pueden interpretar cualquier denuncia como una señal de un bajo nivel de vigilancia que induce a una provisión masiva de mala calidad en el mercado, por lo que no es raro encontrar fraude, lo que los inhibe a participar en el mercado. Bajo esta interpretación la OA no tiene incentivos para vigilar a sus miembros y exponer el fraude.

Técnicamente no es posible determinar cuál interpretación prevalecerá en el mercado debido a que en el juego de señalización subyacente a la denuncia e interpretación de una exposición de fraude no es posible aplicar algún refinamiento existente en la literatura (como el criterio intuitivo de Cho y Kreps (1989), por ejemplo [6]), o algún modelo de aprendizaje debido a la habilidad severamente limitada de los consumidores para observar vigilancia y reconocer fraude. Es esta limitación teórica la que requiere de un enfoque empírico que permita resolverla.

Sin embargo, un estudio de campo está severamente limitado debido a que la evidencia de fraude en cualquier mercado es difícil de establecer una vez que ha sucedido, y porque el economista comparte la misma incertidumbre de los consumidores al tratar de interpretar cualquier exposición de fraude por parte de la OA. Un método empírico alternativo es el uso de técnicas experimentales que proveen al investigador del control de importantes aspectos teóricos como el nivel de vigilancia, la ocurrencia de fraude y la probabilidad con que este es descubierto por la OA, por ejemplo. De esta manera para establecer si el motivo reputacional es suficiente para que funcione la Autorregulación analizaremos, mediante técnicas experimentales, cómo los consumidores interpretan o aprenden a interpretar las denuncias de fraude que reciben de la OA, y si dada esta interpretación el comportamiento de la OA es consistente con la presencia o ausencia de incentivos reputacionales para vigilar y exponer fraude.

Los resultados experimentales obtenidos permiten despejar la limitación teórica e identifican una clara oportunidad para que la autorregulación funcione por motivos reputacionales. Asimismo, se valida el modelo teórico de referencia como un predictor cualitativo o direccional del comportamiento observado en las sesiones experimentales, y establece la robustez de sus predicciones ante la ausencia de un conocimiento común inicial acerca de los parámetros teóricamente relevantes del juego.

Este trabajo está organizado de la siguiente manera: en la segunda sección introducimos el modelo teórico utilizado en las sesiones experimentales; en la tercera sección discutimos el diseño experimental utilizado (controles, tratamientos y procedimientos); en la cuarta sección discutimos los resultados experimentales, y en la quinta sección se resume y concluye el trabajo.

---

<sup>6</sup> Obviamente, la OA puede realizar otras acciones, como por ejemplo servir de centro de arreglo de disputas entre los proveedores.

## 2. MODELO TEORICO

### 2.1 Resumen del Modelo Base: Nuñez (2001) [10]

El modelo que utilizamos para estudiar los incentivos reputacionales para la OA está basado en el modelo de Autorregulación de Nuñez [10] bajo información asimétrica sobre los costos de vigilancia de la OA en industrias de bienes de confianza.

Nuñez describe un juego de dos etapas con los miembros proveedores de la OA actuando como “agentes” que se benefician linealmente del nivel de “fraude” o rebaja en la calidad del servicio que proveen, pero enfrentan un castigo monetario fijo en caso de que la directiva de la OA, encargada de la vigilancia y que actúa como el “principal”, descubra el fraude utilizando un mecanismo de vigilancia, de Alto o Bajo costo que viene determinado por la naturaleza. Intuitivamente se espera que, todo lo demás constante, un principal de Bajo costo de vigilancia realice un nivel de vigilancia mayor que un principal de Alto costo; de los agentes se espera que reduzcan su nivel de fraude óptimo en reacción a un incremento en el nivel esperado de vigilancia.

El bienestar del Público es afectado negativamente por el fraude por lo que sus percepciones sobre el tipo de costo del principal y nivel de fraude asociado afectan su comportamiento en el mercado (su decisión de compra y su predisposición a pagar por el servicio, por ejemplo). Este comportamiento determina de manera importante cuán valioso es para el principal el ser percibido como de Alto o Bajo costo. En este escenario el principal goza de una “ganancia reputacional” cuando es percibido por los consumidores como un tipo de Bajo costo en lugar de uno de Alto costo, debido a que el primero tiene un nivel de fraude asociado más bajo que el segundo por lo que los consumidores se sentirán más confiados en participar en el mercado y pagar por el servicio. La ganancia reputacional puede ser cuantificada como la diferencia entre las valoraciones que resultan del comportamiento de los consumidores para ambos tipos.

El juego se compone de dos etapas. Al inicio de la primera etapa la Naturaleza determina el tipo de costo de vigilancia del principal de acuerdo a probabilidades que son de conocimiento común. El principal observa su tipo y decide un nivel de vigilancia óptimo esperando una reacción óptima por parte del agente. El agente no observa el tipo de costo del principal y decide su nivel de fraude óptimo.

En la segunda etapa del juego, el nivel de fraude ( $x$ ) y vigilancia ( $y$ ) se realizan y el fraude es encontrado (y castigado) con probabilidad  $P(x,y)$ . En caso de que el principal encuentre fraude, este toma su segunda decisión: exponerlo o no a los consumidores. Los consumidores no pueden observar o determinar por sí mismos la existencia de fraude debido a la característica de bienes de confianza del mercado pero observan si el principal expone o no evidencia de fraude. Con esta información los consumidores actualizan sus creencias sobre el tipo de costo del principal y el nivel de fraude existente utilizando la Regla de Bayes. Solo si las creencias actualizadas de los consumidores favorecen al tipo de costo Bajo, el principal recibe una ganancia reputacional.

Nuñez demuestra la existencia de equilibrios bayesianos mezcladores en los que no hay exposición ni vigilancia y existe un nivel de fraude óptimo máximo. Un equilibrio bayesiano mezclador con exposición de fraude, un nivel de vigilancia positivo y bajos

niveles de fraude puede existir solamente cuando  $P(x(y^*), y^*)$  aumenta cuando el nivel de vigilancia óptimo ( $y^*$ ) aumenta. La intuición de esto último es que cuando ambos tipos deciden exponer fraude en equilibrio, los consumidores pueden actualizar sus creencias a favor del tipo de Bajo costo cuando ocurre una exposición, y a favor del tipo de Alto costo cuando no. Debido a que la ganancia reputacional es la misma para ambos tipos, no se puede hacer una comparación de pagos y aplicar algún refinamiento para juegos de señalización, dejando el modelo con muchos equilibrios que tienen que ser refinados de otra manera (empíricamente).

## 2.2 Estructura del Modelo para el experimento

Para efectos de testear experimentalmente el modelo base, simplificamos el juego de dos etapas en una sola etapa e imponemos los resultados de la primera. Esta simplificación permite enfocarnos en la etapa del juego donde los incentivos para exponer y hacer vigilancia son determinados (en la interpretación de una exposición de fraude) y facilita el control experimental y la posterior interpretación de la información. Adicionalmente nos permite mirar a los consumidores de la OA como el “Público” cuya opinión es de nuestro interés, abstrayéndonos del efecto del fraude sobre su bienestar (lo cual requeriría introducir en el experimento una acción de compra en el mercado), y permite la consideración tanto de clientes potenciales y activos en el análisis.

El fijar la primera etapa del juego nos deja con dos tipos de OA: una de “Alta vigilancia – Bajo Fraude” o tipo H, y una de “Baja vigilancia – Mucho Fraude” o tipo L. Suponemos que la probabilidad de que una OA sea tipo H es igual a 0.5. Cada tipo encuentra fraude cuando realiza vigilancia con probabilidades  $P(x_H^*, y_H^*) = P_H$  y  $P(x_L^*, y_L^*) = P_L$  respectivamente, donde  $y_H^* > y_L^*$ , y  $x_H^* < x_L^*$ .

La estructura reducida del juego es la siguiente: las probabilidades  $P_H$  y  $P_L$  son de conocimiento común; al inicio del juego, la Naturaleza determina el tipo de la OA y ésta lo observa, realiza su vigilancia y encuentra fraude con probabilidad  $P_i$ ,  $i = H, L$ . Si logra encontrar fraude decide exponerlo o no al público; si no lo encuentra no puede exponer nada (no puede inventarlo). El público no observa el tipo de la OA, solamente observa si hubo exposición o no. Si alguna evidencia de fraude es expuesta el público determina qué tan probable es que la OA sea tipo H utilizando la regla de Bayes:

$$P(H / \text{hubo exposición}) \equiv P(H / 1) = \frac{0.5 (P_H e_H)}{0.5 (P_H e_H) + 0.5 (P_L e_L)}$$

donde  $e_i \in \{ 0, 1 \}$ ,  $i = H, L$ , es la decisión de la OA tipo  $i$  de exponer fraude ( $e_i = 1$ ) o no ( $e_i = 0$ ). Por supuesto,  $P(L/1) = 1 - P(H/1)$ . En cambio, si ninguna evidencia de fraude es expuesta el público actualiza su creencia de que la OA es tipo H utilizando:

$$P(H / \text{no hubo exposición}) \equiv P(H / 0) = \frac{0.5 (1 - P_H e_H)}{0.5 (1 - P_H e_H) + 0.5 (1 - P_L e_L)}$$

Con las creencias actualizadas de esta manera, el público escoge entre tres posibles opiniones acerca del tipo de la OA:

- a) “creo que es tipo H”
- b) “creo que es tipo L”

c) “estoy indeciso”

Los pagos del público están determinados de tal manera de que tengan incentivos para construir una opinión: si la opinión del público coincide con el tipo de la OA gana  $U > 0$ , si no coincide gana cero, y si dice estar indeciso gana una lotería que le da  $U$  o cero con probabilidad 0.5. La idea detrás de esta estructura de pagos es que si el fraude afecta finalmente el bienestar del Público, ellos ciertamente se benefician de tener una idea correcta acerca del nivel de vigilancia y fraude de la OA al momento de decidir si adquieren o no el servicio en el mercado. La lotería permite un escape para aquellas opiniones que se hacen al azar en el laboratorio.

Los pagos para la OA son los siguientes: obtiene  $W > 0$  si el público opina que es tipo H, obtiene una ganancia de cero si opina que es tipo L (por lo que la ganancia reputacional es  $W$ ), y cuando el público está indeciso gana una lotería que le da  $W$  o cero con igual probabilidad. La **Tabla 1** resume la estructura de pagos utilizada.

Un supuesto final que no afecta los equilibrios bayesianos del juego pero que tiene relevancia para el experimento, cuyas sesiones se componen de muchas rondas de interacción con el mismo juego, es que el Público no puede verificar su opinión al final de cada ronda debido a la característica de bien de confianza del mercado (como tanto el fraude como la vigilancia no pueden ser observados por los consumidores entonces no puede observar el tipo de OA). Esto implica también que los consumidores no serán capaces de observar su ganancia efectiva sino hasta el final de cada sesión experimental.

### 2.3 Equilibrios en estrategias puras del modelo

En el modelo base un equilibrio donde ambos tipos de OA vigilan y exponen fraude, y existe bajo nivel de fraude en el mercado existe solo cuando  $P(x^*, y^*)$  se incrementa cuando el nivel de vigilancia aumenta. Los equilibrios donde ambos tipos de OA no vigilan ni exponen, y hay un máximo nivel de fraude en el mercado siempre existen indistintamente al signo que tome el cambio de  $P(x^*, y^*)$  cuando  $y^*$  aumenta. Por consiguiente en el modelo reducido, donde solamente existen dos combinaciones distintas de  $(x^*, y^*)$  que definen los dos tipos de OA y las probabilidades  $P_H$  y  $P_L$ , existe un equilibrio donde ambos tipos de OA exponen cuando  $(P_H - P_L) > 0$ , y siempre existen equilibrios donde ambos tipos de OA no exponen cuando  $(P_H - P_L)$  es  $> 0$  ó  $< 0$ .

La **Figura 1** resume todos los equilibrios mezcladores en estrategias puras donde ambos tipos de OA deciden no exponer fraude ( $e_H = e_L = 0$ ), asumiendo, sin pérdida de generalidad, que  $U = W = 100$ . Cuando no hay exposición de fraude en equilibrio, la creencia actualizada  $P(H/0)$  es igual al 0.5 inicial debido a que el Público no es capaz de discriminar entre una situación en que la OA no detecta fraude y otra en la que lo está encubriendo; como consecuencia el Público queda indiferente entre cualquier opinión.

La regla de Bayes no puede ser aplicada para obtener las creencias fuera de la senda de equilibrio (CFE), las cuales quedan indeterminadas. Por ejemplo, en la Figura 1a, la CFE que sostiene la reacción del público es que cualquier exposición de fraude (fuera de equilibrio) debe provenir de una OA tipo H con una probabilidad mayor a 0.5; en la Figura 1b la CFE es que una exposición (fuera de equilibrio) debe provenir de una OA tipo L con probabilidad mayor a 0.5; y en la Figura 1c la CFE es que es igualmente probable que una exposición provenga de cualquier tipo. Debido a esta

indeterminación, todos los equilibrios con no exposición de fraude son secuenciales en el sentido de Kreps y Wilson (1982). Sin embargo, no se puede aplicar ningún refinamiento a estos equilibrios debido a que tanto el público como la OA no poseen estrategias dominadas y no es posible hacer comparación alguna entre los pagos que reciben cada tipo de OA (ambos ganan el máximo de  $W$  y el mínimo de cero, por lo que, por ejemplo, el criterio intuitivo de Cho y Kreps (1987) no puede ser aplicado).

A pesar de que ninguno de estos equilibrios puede ser refinado, algunos de ellos no son consistentes con una “MALA” interpretación de la exposición del fraude. Una MALA interpretación ocurre cuando el público piensa que cualquier exposición accidental es más probable que provenga de una OA tipo L. La intuición detrás de la misma es que, por un lado, el nivel de fraude de una OA tipo L debe ser tan alto o generalizado que cualquier nivel positivo de vigilancia es suficiente para encontrarlo y, por otro lado, el nivel de fraude de una OA tipo H debe ser tan bajo que en realidad es difícil encontrarlo. Con una MALA interpretación en sus mentes el Público opinará que la OA es tipo L ante cualquier exposición de fraude y en consecuencia ambos tipos de OA no tendrán incentivos reputacionales para exponer (y, una etapa más atrás, para vigilar a sus miembros). En resumen, bajo una MALA interpretación se espera que ningún tipo de Organización exponga fraude si lo encuentra, y que el Público opine que la Organización es tipo L en caso de que, por error, lo exponga; el público puede tener cualquier opinión cuando no se expone fraude. Los únicos equilibrios en la Figura 1 que son consistentes con una MALA interpretación de la exposición del fraude son aquellos en 1b, 1d y 1e (el equilibrio 1b domina en el sentido de Pareto a los otros dos).

La **Figura 2** resume el caso de un equilibrio con exposición. En este equilibrio ambos tipos de OA exponen cualquier evidencia de fraude al Público (es decir,  $e_H = e_L = 1$ ). La creencia actualizada de qué tipo es más probable que exponga fraude favorece a la OA tipo H (como  $(P_H - P_L) > 0$ , las actualizaciones resultan en  $P(H/1) > P(L/1)$ ). En este caso la CFE está completamente determinada debido a que es más probable que una OA tipo L no exponga (las actualizaciones resultan en  $P(L/0) > P(H/0)$ ).

Este equilibrio es consistente con una “BUENA” interpretación de la exposición del fraude. Una BUENA interpretación ocurre cuando el público piensa que una exposición es más probable que venga de una OA tipo H. La idea detrás de esta interpretación es que, por un lado, el nivel de vigilancia de una OA tipo H debe ser tan grande que fácilmente descubre cualquier nivel existente de fraude, y que, por otro lado, el nivel de vigilancia de una OA tipo L debe ser tan bajo o inexistente que difícilmente puede encontrar fraude. Con una BUENA interpretación en sus mentes el público opinará que la OA es tipo H ante cualquier exposición de fraude y opinará que es tipo L cuando no ocurra exposición de fraude; en consecuencia ambos tipos de OA tendrán incentivos reputacionales para exponer (y, una etapa más atrás, para vigilar a sus miembros). En resumen, bajo una BUENA interpretación se espera que ambos tipos de Organización expongan fraude si lo encuentra, y que el Público opine que la Organización es tipo H en caso de que lo exponga, y que es tipo L cuando no lo exponga. En la **Tabla 2** se resume las proporciones de juego para cada acción posible de la OA y del Público consistentes con una MALA y BUENA interpretación de la exposición del fraude.

La validez predictiva del modelo requiere que en las sesiones experimentales los individuos libremente actúen de acuerdo a las acciones de equilibrio de una MALA interpretación cuando  $(P_H - P_L) < 0$  (es decir, se espera que jueguen las acciones de los



equilibrios 1b, 1d ó 1e, como se resume en la Tabla 2). En cambio, se espera esclarecer qué tipo de interpretación, MALA ó BUENA, surge del comportamiento de los individuos cuando  $(P_H - P_L) > 0$ ; si este comportamiento es consistente con una BUENA interpretación será evidencia a favor del motivo reputacional para que la autorregulación funcione en industrias de bienes de confianza.

### 3. DISEÑO EXPERIMENTAL

#### 3.1 Controles y Tratamientos

La primera variable de control lógicamente viene dada por el signo de  $(P_H - P_L)$ , que divide claramente las posibilidades de que surja o no una BUENA interpretación. La **Tabla 3** resume los valores de  $P_H$  y  $P_L$  utilizadas para este control.

En un ambiente más realista, las probabilidades de encontrar fraude  $P_H$  y  $P_L$  también forman parte de un largo proceso de aprendizaje social tanto para el público como para la OA. Si no existe inicialmente información acerca de estas probabilidades, qué interpretación puede emerger en el Público depende de sus conjeturas iniciales sobre el signo de  $(P_H - P_L)$ . Sin un referente externo, diferentes sujetos deberían tener diferentes conjeturas iniciales.

Esto es particularmente importante en nuestro caso debido a la inhabilidad del público para aprender de sus opiniones correctas e incorrectas en un mercado de bienes de confianza, y debido a que la literatura de aprendizaje en juegos de señalización, que se utiliza cada vez más como refinamientos, se fundamenta principalmente en procesos de ajuste que dependen de la historia del juego, donde cada jugador puede al menos observar el resultado de sus acciones y ajustar completa o parcialmente sus creencias sobre los tipos utilizando esta información, favoreciendo aquellas estrategias con mayor pago esperado [7]. Por lo tanto, controlar la información disponible sobre los valores de  $P_H$  y  $P_L$  que subyacen en un juego permite analizar la capacidad de aprendizaje de los sujetos para aprender estos parámetros estructurales y ver si ajustan o no su comportamiento en vista de esta información, aunque como Público no observen si su opinión estuvo correcta o no. Esto nos permite chequear la robustez de la teoría y a la vez explorar nuevas posibilidades de aprendizaje para juegos de señalización.

Por lo tanto, como nuestra segunda variable de control incluimos el nivel de información disponible sobre  $P_H$  y  $P_L$  para los jugadores. Establecemos un nivel de información TOTAL, donde las probabilidades  $P_H$  y  $P_L$  que subyacen en la estructura de cada sesión experimental son reveladas al inicio y recordadas a los jugadores para asegurar al menos su conocimiento mutuo. También establecemos un nivel de información NULO, donde ningún tipo de información sobre estas probabilidades es explícitamente dado pero existe en la estructura subyacente de la sesión experimental. El comportamiento observado en las sesiones experimentales bajo un nivel de información TOTAL sirven como un techo natural para lo que (probablemente) ocurra como una situación de largo plazo o steady state en el comportamiento de las sesiones bajo un nivel de información NULO, si algún proceso de ajuste ocurre debido a la experiencia y aprendizaje de las probabilidades estructurales, y también ofrece el mejor ambiente para testear el efecto teórico del signo de  $P_H - P_L$ .

La combinación del primer y segundo control define cuatro posibles tratamientos para nuestro experimento, como se puede observar en la **Tabla 4**.

Finalmente, de acuerdo con Cooper y Kagel [1], el uso de un contexto con significado en lugar de un contexto neutro en los experimentos puede facilitar a los participantes la comprensión de las complejidades estratégicas usualmente envueltas en juegos de señalización, acelerando el proceso de aprendizaje y actuando como un sustituto parcial de la experiencia con el juego (efecto de contexto débil), o también puede afectar la selección de equilibrio de los sujetos (efecto de contexto fuerte). Por otro lado, Cooper et al. [2] señala que un contexto con significado, aplicado a sujetos con experiencia de campo con el fenómeno de señalización bajo estudio, puede gatillar el uso de estrategias familiares a éstos. Debido a que en nuestro experimento utilizamos estudiantes, el uso de un contexto con significado se justifica como una forma de mejorar la comprensión estratégica del juego. Este tipo de contexto puede ser una alternativa útil en experimentos que estudian fenómenos de señalización específicos con implicancias de política debido a la posibilidad de un efecto de contexto fuerte;<sup>7</sup> por consiguiente, nosotros realizamos algunas sesiones en cada tratamiento bajo un contexto con significado y algunas sesiones bajo un contexto neutral, como aparece en la **Tabla 4**.

### **3.2 Procedimientos**

Las sesiones experimentales fueron conducidas a mano en el campus Andrés Bello de la Universidad de Chile. Este campus concentra estudiantes de las carreras de ingeniería comercial, auditoría, arquitectura, geografía y diseño. Cada sesión experimental empleó 12 sujetos sin experiencia reclutados de los primeros tres años de cada carrera. Un total de 16 sesiones fueron conducidas bajo los cuatro tratamientos indicados en la **Tabla 4** entre Octubre del 2003 y Agosto del 2004.

Al inicio de cada sesión, un set común de instrucciones fue leído en voz alta, y cada sujeto tenía una copia escrita del mismo. Todos los participantes llenaron posteriormente un cuestionario corto para verificar su nivel de comprensión del juego y para calcular correctamente sus pagos; las respuestas correctas del cuestionario fueron leídas en voz alta y cualquier duda que quedara era despejada.

En las sesiones de contexto con significado se utilizó en las instrucciones un esquema donde existía una “Organización” cuyos miembros proveían un servicio al “Público”; este servicio podía ser “bueno” o “malo”. La Organización realizaba un control imperfecto sobre sus miembros mediante un nivel activo de vigilancia que podía ser “Alto” o “Bajo”, y la reacción de sus miembros era tal que con un Alto nivel de vigilancia pocos miembros proveían un mal servicio, pero con un Bajo nivel de vigilancia la mayoría de ellos proveían un mal servicio. Estas dos posibilidades definían dos tipos igualmente probables de Organización: una de “Alta Vigilancia – Buen servicio” o H, y otra de “Baja Vigilancia – Mal Servicio” o L. En las sesiones de contexto Neutral se utilizó el esquema Jugador A - Jugador B, donde el mecanismo de búsqueda era sobre un tazón con bolitas rojas y blancas.

---

<sup>7</sup> Por ejemplo, Cooper et al [2] utilizan contextos neutros y con significado para estudiar la existencia del efecto “ratchet” en los administradores de actividades de producción que son planificadas centralmente.

Luego de leer las instrucciones y llenar el cuestionario, los participantes fueron conducidos a otra sala y asignados aleatoriamente a cubículos donde no tenían contacto directo con otros participantes, para mantener la privacidad del ejercicio. Antes de que una sesión empezara los participantes tomaron algunos minutos para leer un resumen corto con la información más importante y las reglas del juego; cualquier duda que quedara fue aclarada en ese momento. Este resumen escrito estaba disponible en cada cubículo y podía ser revisado en cualquier momento. Las instrucciones y la hoja de resumen para el tratamiento TOTAL – NEGATIVO en contexto con significado pueden encontrarse en el **Apéndice Metodológico** de esta tesis.

En las sesiones con un nivel de Información TOTAL se incluyó textual y gráficamente en las instrucciones y resumen la probabilidad de detección de mal servicio para cada tipo. En las sesiones con nivel de información NULO ninguna información probabilística fue incluida; en su lugar se decía a los participantes que ambos tipos de Organización algunas veces podían detectar miembros proveyendo un mal servicio, pero que el número de veces que esto ocurría podía ser igual o distinto para cada tipo.

Cada sesión duró 42 rondas, excepto por dos sesiones que duraron 36 rondas. Evitamos decir a los participantes cuántas rondas tenía cada sesión con el fin de evitar comportamientos de fin de periodo. No se condujeron rondas de prueba, en su lugar las primeras 12 rondas fueron conducidas a un ritmo muy lento.

Al inicio de la sesión a cada participante se le asignó aleatoriamente un rol (público u organización), y a cada organización se le asignó aleatoriamente un tipo, tal que entre los 12 participantes habían seis en cada rol y tres en cada tipo. Estos roles y tipos se mantuvieron fijos durante un set de seis rondas, en el que cada público fue pareado una sola vez con una organización diferente. Cada seis rondas los roles y tipos fueron reasignados aleatoriamente hasta el final de la sesión. No se utilizó el sistema de cambio de roles automático entre un set de seis rondas y otro, como es usual en otros trabajos experimentales con juegos de señalización ([2], [3] y [4], por ejemplo) debido a que esto facilita una coordinación estratégica simple entre los participantes (especialmente en las sesiones con nivel de información NULO). Con un cambio automático de roles, en cada dos sets es posible que aparezca una alta proporción de opiniones “H” entre el público sin importar lo que decida la organización; lo anterior da un pago seguro de  $W$  a la Organización y un pago esperado de  $0.5*U$  al Público, que en conjunto suman un pago esperado que iguala al mayor pago esperado que es posible obtener en equilibrio para cada participante, sin tener que pensar en las complejidades estratégicas, lo que nos haría perder el control del experimento. Debido a que esta coordinación se construye en la amenaza de castigar cualquier desviación grupal (por ejemplo, no opinar siempre H si en el set anterior el otro grupo no lo hizo), una reasignación aleatoria de roles hace más difícil que surja (el grupo que se desvía difícilmente coincidirá exactamente con el grupo que recibe el castigo posterior).<sup>8</sup>

Debido a que es posible que un mismo participante enfrente el mismo rol varias veces, al reasignar los roles se aplica una probabilidad de 0.05 de que cada participante tenga

---

<sup>8</sup> Por otro lado, una asignación aleatoria de roles hace más difícil la construcción de reputación en ambientes de laboratorio cuando se quiere inducir juegos de una sola etapa, lo cual facilita la identificación de efectos de aprendizaje al costo de que los participantes no experimenten ambos roles de manera temprana; debido a esto último, en experimentos futuros debe ser cuidadosamente evaluada la deseabilidad de tener al estructura de set aleatoria utilizada en este trabajo.

el mismo rol por más de dos sets consecutivos, lo que permite mantener la veracidad de decir a los participantes de que el proceso de asignación de roles es aleatorio (necesario para romper el mecanismo de coordinación). También se favoreció las configuraciones de roles en las que los participantes experimentaban ambos tipos de organización de una manera temprana en el experimento, dado el corto número de rondas disponibles.

La asignación de roles y tipos por set, para cada participante y sesión experimental, fue hecha antes de la sesión. También se simuló y estableció por adelantado los descubrimientos de mal servicio de cada organización,<sup>9</sup> y el resultado de la lotería en caso de que el público esté indeciso en cada ronda.

Al inicio de cada ronda, cada organización tenía una tarjeta escrita dividida en tres partes. En la parte izquierda estaba escrito su tipo y si la organización descubrió o no miembros dando un mal servicio en esa ronda. En la parte superior derecha se encontraba un cajón vacío junto a un mensaje que indicaba “Encontré miembros brindando un mal servicio”. En la parte inferior derecha se encontraba tres cajones vacíos junto a tres mensajes: “creo que es ‘H’”, “creo que es ‘L’”, “estoy indeciso”. La tarjeta estaba prepicada de tal manera que la parte izquierda y derecha podían ser separadas con facilidad.

En cada ronda, la interacción entre Organización y Público se daba de la siguiente manera: la organización observaba en la tarjeta si encontró o no miembros dando un mal servicio; solamente si los encontraba la organización podía decidir si anunciaba o no su hallazgo al público (decidía si marcaba o no el anuncio en la parte superior derecha de la tarjeta; cualquier falsa denuncia era explícitamente penalizada); luego, la organización separaba la tarjeta y enviaba la parte derecha al público; observando la tarjeta enviada, el público marcaba uno de los cajones en la parte inferior y se la devolvía a la organización, finalizando la ronda.<sup>10</sup> De esta manera, en cada ronda el público observaba si la organización con la cual estaba pareada en esa ronda denunciaba algo o no, la organización observaba la opinión del público (con lo que podía calcular sus ganancias) pero el público no observaba si su opinión estuvo correcta o incorrecta.

Además del resumen de las instrucciones, cada participante tenía una Hoja de Record en la cual debía registrar los eventos del experimento, como el rol y tipo que tenía en cada set de seis rondas, sus propias decisiones y la de su compañero de grupo en cada ronda.

En los pilotos previos a las sesiones definitivas se observaba una inusual concentración de exposición de fraude al final de cada set de seis rondas cuando se pagaba lo obtenido en todas las rondas, indicando la existencia de algún tipo de efecto riqueza que fue controlado al pagar aleatoriamente una sola ronda de cada set de seis, [8]. En cada ronda se fijó el valor de U y W igual a \$ 2.400 pesos chilenos (USD \$ 3.9) para las sesiones con nivel de información TOTAL, e igual a \$ 2.600 pesos (USD \$ 4.2) en las sesiones con nivel de información NULO. Una sesión completa duraba típicamente dos horas (sólo las 42 rondas duraban 85 minutos).

---

<sup>9</sup> Eligiendo aleatoriamente una de las combinaciones de 2 ú 8 éxitos en 10 intentos, según el tipo de organización, y utilizando los primeros seis resultados de la combinación escogida.

<sup>10</sup> Se entrenó a ocho estudiantes de postgrado como asistentes de laboratorio y se utilizaron tres de ellos en cada sesión para el traslado de las tarjetas.

En los pagos realizados se considera que Chile tiene un costo de oportunidad más bajo entre sus estudiantes universitarios que otros países desarrollados en los que rutinariamente se realizan experimentos; esto puede ser inferido al comparar los índices de paridad de poder compra, salario mínimo legal y el salario promedio esperado de mercado para un profesional graduado del campus en que se realizó el experimento.<sup>11</sup> Adicionalmente, tenemos una excelente medida del costo de oportunidad por hora de nuestros estudiantes en los honorarios que ofrece la Unidad de Encuestas del Departamento de Economía. La Unidad de Encuestas frecuentemente ofrece trabajos temporales a los estudiantes pagándoles un máximo de \$ 2.500 pesos (USD \$ 4.1) la hora de encuesta. En promedio nuestros participantes ganaron, incluyendo el pago por presentación, \$ 10.600 pesos (USD \$ 17.3) por una sesión de dos horas. El 86% de nuestros participantes obtuvieron ganancias entre los \$ 5.000 (USD \$ 8.1) y los \$ 16.800 pesos (USD \$ 27.4). En una de las sesiones, no considerada en los cálculos anteriores, se incrementó significativamente los pagos, los participantes obtuvieron en promedio \$ 16.300 pesos (USD \$ 26.5) y su comportamiento fue indistinguible de otras sesiones bajo el mismo tratamiento.

Al finalizar cada sesión los participantes podían verificar los pagos que obtuvieron utilizando un resumen detallado con las opiniones del público, los resultados de la lotería y los tipos de organización, ordenados por el número de los participantes y grupos formados en cada ronda.

#### **4. RESULTADOS EXPERIMENTALES**

De ahora en adelante nos referiremos a las sesiones experimentales conducidas bajo un nivel de Información TOTAL ó NULO simplemente como sesiones TOTAL ó sesiones NULO. Nos referiremos a las sesiones conducidas bajo  $(P_H - P_L) > 0$  como sesiones POSITIVO y a las sesiones conducidas bajo  $(P_H - P_L) < 0$  como sesiones NEGATIVO. Cuando sea necesario referirnos a algún tratamiento (combinación de controles) primero colocaremos el nivel de información de la sesión y luego el signo de  $(P_H - P_L)$ , por ejemplo la sesión TOTAL – NEGATIVO ó NULO – POSITIVO.

##### **4.1 Validación de la Teoría: SESIONES TOTAL**

###### **4.1.1 Determinación empírica de los equilibrios e interpretaciones.**

Como fue mencionado en la sección 2, las sesiones bajo un nivel de información TOTAL pueden considerarse como una situación de largo plazo para el comportamiento de las sesiones bajo un nivel de información NULO en caso de que algún proceso de ajuste ocurra debido a la experiencia y aprendizaje de la estructura del juego (del signo de  $(P_H - P_L)$ , particularmente). Mucho más importante, sin embargo, es que estas sesiones constituyen el mejor ambiente para identificar el efecto del signo de  $(P_H - P_L)$  sobre el comportamiento de los participantes, y además nos permite determinar cuál equilibrio teórico, si aquel de una BUENA ó una MALA interpretación, encuentra mayor sustento cuando  $(P_H - P_L) > 0$ .

---

<sup>11</sup> El Índice de Paridad de Poder de Compra del Banco Mundial para el año 2000 (2001) fue de 0.56 (0.47) para Chile, 0.70 (0.70) para España, 1.03 (1.00) para EE. UU, 1.02 (1.03) para el Reino Unido y 0.23 (0.23) para China [2]. El salario mínimo legal en Chile para el 2004 fue de \$ 114.000 pesos (USD\$ 186) mensual (48 horas a la semana). El salario medio para un ingeniero comercial que ingresa al mercado laboral es de \$ 500.000 pesos (USD\$ 814) mensuales.

En la **Figura 3** se puede observar el desenvolvimiento del juego a través del tiempo en las sesiones TOTAL. La figura muestra las proporciones jugadas por las organizaciones y público en las sesiones POSITIVO (barras grises) y NEGATIVO (barras negras), separadas también en sesiones conducidas bajo un contexto con significado (barras punteadas) y neutral (barras con cuadros). Las proporciones de cada sesión que se presentan en la figura se obtienen y promedian las decisiones de todos los individuos en el primer set de seis rondas y por cada dos sets consecutivos posteriormente.

En la **Figura 3** se observan cinco gráficos. El *primer gráfico* muestra la proporción de fraude expuesto para cada sesión (en desviaciones de 50%). El *segundo gráfico* muestra la diferencia en las proporciones de las opiniones “creo que es H” y “creo que es L”, que de ahora en adelante denominaremos opiniones H y L respectivamente, cuando la Organización expone mal servicio o fraude; un valor positivo indica que hubo más opiniones H que L, y el máximo de +100% indica que solamente se observó opiniones H cuando las organizaciones exponían mal servicio; un valor negativo indica que hubo más opiniones L que H, y el mínimo de -100% indica que solamente se observó opiniones L cuando las organizaciones exponían mal servicio. El *tercer gráfico* muestra la proporción de opiniones “estoy indeciso” cuando el fraude es expuesto. El *cuarto gráfico* muestra la diferencia en las proporciones de las opiniones H y L cuando el fraude NO es expuesto, y tiene la misma interpretación que el segundo gráfico. Finalmente, el *quinto gráfico* muestra el nivel de respuestas “estoy indeciso” cuando el fraude NO es expuesto. En los gráficos se observa una diferencia fuerte y clara en el comportamiento de los individuos de acuerdo al signo de  $P_H - P_L$ , sin importar mayormente el contexto utilizado.

En la **Tabla 5** se muestran las proporciones de juego en las sesiones Total para cada acción posible (con un intervalo del 95%) que se obtienen al promediar todas las rondas y sesiones de acuerdo al signo de  $(P_H - P_L)$ . Se observa claramente que ninguna de las proporciones de juego esperadas para una interpretación BUENA ó MALA, reproducidas en las últimas dos columnas, se dieron estrictamente. Sin embargo, consistente con la Figura 3, se observa una diferencia marcada en el comportamiento de las organizaciones y público de acuerdo al signo de  $(P_H - P_L)$ , que se corrobora al observar que en cada acción posible los intervalos del 95% no se intersectan, salvo para las opiniones “estoy indeciso”. Lo anterior nos lleva a nuestra primera conclusión:

**CONCLUSION UNO (Tipo de Interpretación según el signo de  $(P_H - P_L)$ ):** *Ninguna de las proporciones de juego de una BUENA ó MALA Interpretación fueron estrictamente observadas; sin embargo, existe una clara diferencia en el comportamiento de los individuos de acuerdo al signo de  $(P_H - P_L)$ . En las sesiones NEGATIVO, una MALA interpretación tiene mayor sustento empírico que los otros equilibrios de no exposición y de una BUENA Interpretación. En las sesiones POSITIVO, una BUENA interpretación tiene mayor sustento empírico. En las sesiones de ambos signos el Público está menos incierto cuando el fraude es expuesto que cuando no lo es.*

Como se puede observar en la Tabla 5, en las sesiones NEGATIVO la proporción de juego observada para cada rol está cercana a la proporción esperada bajo una MALA interpretación: la proporción de exposición de fraude de las organizaciones no fue 0% pero fue baja (30% en promedio), y entre el público hubo substancialmente más

opiniones L que H cuando el fraude fue expuesto (67 puntos más). Bajo una mala interpretación las proporciones de opiniones H y L cuando el fraude no es expuesto no están determinadas, sin embargo se puede observar en la tabla que hubo más opiniones H que L (27 puntos más), lo cual favorece el equilibrio de la Figura 1b.

En la misma Tabla, se observa que en las sesiones POSITIVO la proporción de juego fue cercana a la proporción esperada bajo una BUENA interpretación: la proporción de exposición al fraude no fue 100% pero fue bastante alta (92% en promedio), las opiniones H cuando el fraude fue expuesto fueron substancialmente mayores que las opiniones L (78 puntos más), y las opiniones H cuando el fraude no fue expuesto fueron substancialmente menores a las opiniones L (58 puntos menos).

#### **4.1.2 Análisis de los Patrones de Juego en el Agregado: Sesiones Total**

Una de las preguntas que inmediatamente llega a la mente luego de ver los resultados de la sección anterior es por qué los individuos jugaron como jugaron, es decir, cuáles fueron los determinantes de este comportamiento. Dos variables que intuitivamente deberían afectar el comportamiento de los jugadores son: 1) el signo de  $(P_H - P_L)$ , que se desprende de toda la discusión teórica anterior; y, 2) la capacidad de los jugadores a reaccionar a las acciones que van tomando el resto de los jugadores. Es decir, por un lado los jugadores simplemente pueden utilizar la información del signo de  $(P_H - P_L)$  para determinar su juego o, por otro lado, van ajustando su juego de acuerdo a como observan o infieren que van a jugar el resto de los participantes.

Con respecto a este último punto, hay que notar que en cualquier set de seis rondas un sujeto que tenga el rol de Organización o Público pudo haber experimentado el mismo rol o uno distinto en el set anterior. Luego, si el sujeto tuvo un rol diferente en el set anterior también tuvo la oportunidad de observar las decisiones de otros participantes que tuvieron su rol actual; si en cambio en el set anterior tuvo el mismo rol, también tuvo la oportunidad de observar las decisiones de los otros participantes que tuvieron el rol que nuevamente enfrentará en este set. En ambos casos, los sujetos pueden utilizar esta experiencia para construir conjeturas sobre el comportamiento agregado y ajustar su propio comportamiento. Debido a que la estructura de apareamiento aleatorio utilizada hace incierto para cada participante qué rol y tipo tuvieron y tendrá el resto de los participantes, es más plausible que para tratar de inferir el juego de su oponente en cualquier ronda los participantes piensen en términos del (promedio) agregado.

Para poder separar debidamente estos dos efectos y diferenciarlos de otros comportamientos (como heterogeneidad en el comportamiento debido a factores idiosincrásicos), utilizamos un modelo Probit de datos de panel con efecto aleatorio [13], cuya metodología e interpretación de resultados explicaremos a continuación.

##### **4.1.2.1 Metodología de identificación de motivación de juego en el agregado.**

Estimamos dos modelos, que se diferencian en el set de variables dependientes, con tres estimaciones Probit en cada modelo. En la primera estimación la variable dependiente es la decisión de la Organización de exponer (1) o no exponer (0) fraude cuando este es encontrado; en la segunda estimación, la variable dependiente es la decisión del Público de opinar H (1) ó L (0) cuando el fraude es expuesto; en la tercera estimación, la

variable dependiente es la decisión del Público de opinar H (1) ó L (0) cuando el fraude no es expuesto.<sup>12</sup>

En la **Tabla 6** se presentan los resultados del primer modelo. En este modelo se utiliza solamente la información desde el 2do. set de seis rondas en adelante para hacerlo comparable. Como variables dependientes incluye la interacción de una variable “dummy” para los sets 1-2-3, 4-5 y 6-7 con una variable dummy para los tratamientos Total – Negativo y Total – Positivo. Hay que notar que, debido a que las variables dummy no varían entre individuos y  $\Phi(0) = 0.5$ , el signo de sus coeficientes estimados indican la dirección de juego en cada tratamiento con respecto a un punto medio, que corresponde a una proporción de exposición de fraude del 50% en la primera estimación, y a una igual proporción de opiniones H y L, es decir una diferencia de proporciones de 0%, en la segunda y tercera estimación. Por tanto, un signo positivo en los coeficientes para cualquier tratamiento indica que el comportamiento está por sobre el punto medio, y un signo negativo indica que está por debajo del punto medio. La significancia del coeficiente indica si el comportamiento es estadísticamente distinto al punto medio y su magnitud indica que tan lejano está del mismo.

Como puede apreciarse en la Tabla 6, existe una marcada diferencia (estadísticamente significativa) en el comportamiento de acuerdo al signo de  $(P_H - P_L)$  consistente con la dirección prevista en la conclusión 1, una vez controlada la heterogeneidad. Hay que notar que, de acuerdo a lo que hemos venido señalando, las dummies por set y tratamiento recogen tanto el efecto que la información del signo de  $(P_H - P_L)$  tiene sobre las decisiones de los participantes como cualquier proceso de ajuste que ocurra debido a las reacciones de los jugadores a lo que juega o espera que juegue el resto.

Para separar ambos efectos estimamos el segundo modelo, cuyos resultados se presentan en la **Tabla 7**. Este modelo incluye las mismas variables dependientes del primer modelo y adicionalmente incorpora tres variables adicionales que interactúan con dummies para cada tratamiento: la proporción total de exposición al fraude del set anterior (Expone / encuentra(-1)), la diferencia entre las proporciones totales de las opiniones H y L cuando el fraude es expuesto del set anterior ( $H - L$  / expone (-1)), y la diferencia entre las proporciones totales de las opiniones H y L cuando el fraude no es expuesto del set anterior ( $H - L$  / no expone (-1)). Como se señaló anteriormente, se puede esperar que algunos individuos utilicen su experiencia en sets pasados para ajustar su comportamiento, por lo tanto, se espera que estas **variables promedio agregadas** capturen cualquier ajuste medio del comportamiento en reacción al comportamiento agregado en el pasado. En la medida que el comportamiento esté correlacionado en el tiempo, estas variables agregadas también capturan los ajustes ante conjeturas sobre el posible comportamiento agregado medio actual.

Debido a que el proceso de ajuste puede ser más pronunciado o solo significativo para uno de los tratamientos o cada tratamiento puede tener un proceso de ajuste diferente, el segundo modelo es el resultado de un proceso de dos etapas: en la primera etapa se incluye y estima cada variable promedio agregada por tratamiento y se conserva aquellas que son significativas hasta en un 20% de significancia; en la segunda etapa se reestima el modelo con las variables agregadas significativas de la etapa anterior y se

---

<sup>12</sup> Debido a que las proporciones de opiniones “estoy indeciso” se mantuvieron relativamente iguales a lo largo de todos los sets, como se puede apreciar en la figura 3, tanto cuando la organización expone como cuando no expone fraude, nos enfocamos solamente en las opiniones H y L.



mantiene solo aquellas que son significativas hasta en un 10% de significancia (se incluyen todas las dummies por set y tratamiento en ambas etapas). Este proceso es robusto a otros niveles de corte en significancia y siempre escoge las mismas variables agregadas. De esta manera se logra identificar de una manera más precisa el efecto de las variables agregadas para los tratamientos correctos.

Una vez que controlamos los procesos de ajuste mediante variables agregadas y la heterogeneidad entre los participantes, las dummies por set y tratamiento deben recuperar la mayoría del efecto del conocimiento del signo de  $(P_H - P_L)$  en las decisiones de los participantes.

#### **4.1.2.1 Interpretación de los resultados de motivación del juego a nivel agregado**

Una comparación directa de las Tablas 6 y 7 muestran una clara diferencia en cada estimación, una vez identificado y controlado el efecto de las variables agregadas y heterogeneidad. Las principales diferencias las resumimos en la siguiente conclusión:

##### ***CONCLUSION DOS (Efecto del signo de $(P_H - P_L)$ y comportamiento agregado):***

a) *Sesiones Negativo: Para que la proporción de sujetos que expone fraude sea menor al 50% debe observarse en sets pasados una mayor proporción de sujetos que opinan que una organización que expone tiene mucho fraude (tipo L). El conocimiento del signo negativo de  $(P_H - P_L)$  es suficiente para que la proporción de sujetos que opina que una organización que expone tiene mucho fraude (L) sea mayor a la que opina que tiene poco fraude (H). Para que la proporción de sujetos que opina que una organización que NO expone tiene poco fraude (H) crezca con respecto a la que opina lo contrario (L) también debe observarse en sets pasados que sea mayor, y también ayuda observar exposición de fraude.*

b) *Sesiones Positivo: El conocimiento del signo positivo de  $(P_H - P_L)$  es suficiente para que la proporción de sujetos que expone fraude sea elevada y para que la proporción de sujetos que opina que una organización que NO expone tiene mucho fraude (L) sea mayor a la que opina lo contrario (H). Para que la proporción de sujetos que opinan que una organización que expone tiene poco fraude (H) se mantenga mayor a la que opina que tiene mucho fraude (L) también debe observarse en sets pasados que sea mayor.*

Como se puede apreciar en la Estimación 1 de las tablas 6 y 7, los coeficientes que acompañan a las dummies por set y tratamiento cambian para las sesiones Negativo (pierden su significancia y son menos negativas en la Tabla 7) y casi no varían para las Positivo, y adicionalmente es significativa la variable agregada  $H - L$  / anuncia (-1) para las sesiones Negativo. Esto indica que, parte del 30% de exposición media de fraude en las sesiones Negativo se sostiene cuando el Público prefiere opinar L en más ocasiones que H cuando el fraude es expuesto en sets pasados, aunque el signo negativo de  $(P_H - P_L)$  todavía sigue siendo importante para mantener el nivel de fraude menor o igual al 50%. Asimismo indica que en las sesiones positivo, debido a que ninguna variable agregada resultó significativa, el 92% de exposición de fraude se debe fundamentalmente al conocimiento del signo positivo de  $(P_H - P_L)$ .

Asimismo, en la Estimación 2 de las mismas tablas se puede apreciar que los coeficientes que acompañan a las dummies por set y tratamiento no cambian para las sesiones negativo, pero sí experimentan un cambio para las sesiones positivo (pierden

significancia y son menos positivas en la Tabla 7), y la variable agregada  $H - L / \text{expone} (-1)$  es significativa para las sesiones positivo. Esto indica que en las sesiones negativo la diferencia de 58 puntos porcentuales que favorece a las opiniones L sobre las H ante una exposición de fraude responde principalmente al conocimiento del signo negativo de  $(P_H - P_L)$ .<sup>13</sup> También indica que en las sesiones positivo la diferencia de 78 puntos que favorece a las opiniones H sobre las L ante una exposición de fraude responde en gran medida a un proceso autosostenido en el que los jugadores necesitan observar una diferencia que también favorezca a las opiniones H en sets pasados para replicar el mismo comportamiento; obviamente el signo positivo de  $(P_H - P_L)$  ayuda para iniciar la sesión con este comportamiento, pero su rol se vuelve secundario en rondas posteriores, lo cual se infiere en la no significancia de las dummies por set y tratamiento en la Tabla 7.

Por último, en la Estimación 3 de las Tablas 6 y 7 se aprecia que los coeficientes que acompañan a las dummies por set y tratamiento cambian para las sesiones Negativo (de mostrar un patrón de crecimiento y ser significativas en la Tabla 6 pasan a no mostrar ningún patrón y perder su significancia en la Tabla 7) y también para las sesiones Positivo (no pierden significancia pero se vuelven más negativas en la Tabla 7); adicionalmente las variables agregadas  $H - L / \text{no expone} (-1)$  y  $\text{Expone} / \text{encuentra} (-1)$  son significativas para las sesiones Negativo, y la variable agregada  $H - L / \text{expone} (-1)$  es significativa para las sesiones Positivo. Esto indica que en las sesiones Negativo la diferencia de 27 puntos porcentuales que favorece a las opiniones H sobre las L cuando el fraude NO es expuesto, y su tendencia a crecer observada en la Tabla 6, responde principalmente a un proceso autosostenido en el que los jugadores necesitan observar una diferencia que también favorezca a las opiniones H en sets pasados para replicar el mismo comportamiento, y también a la observancia de exposición del fraude en sets pasados lo que sugiere que los sujetos tratan de inferir de la presencia de una exposición una manera de actuar en el escenario de no exposición;<sup>14</sup> el signo negativo de  $(P_H - P_L)$  casi no cumple rol en este comportamiento debido a la no significancia y bajo coeficiente estimado para las dummies por set y tratamiento. Para las sesiones Positivo lo anterior indica que la diferencia de 58 puntos que favorece a las opiniones L en lugar de las H cuando el fraude NO es expuesto tiene un proceso muy particular: la presencia de un mayor proporción de opiniones H sobre las L cuando el fraude es expuesto en sets pasados aumenta la propensión en el set actual de los sujetos a opinar H en vez de L cuando el fraude NO es expuesto; en ausencia de este efecto agregado el efecto del signo positivo de  $(P_H - P_L)$  es más fuerte y favorece a las opiniones L sobre las H (lo que se puede observar por el drástico aumento en los coeficientes estimados para las dummies por set y tratamiento en la Tabla 7). Creemos que en este caso la variable

---

<sup>13</sup> En la Figura 2 y en el signo de los coeficientes de las Tablas 6 y 7 se observa un elevado nivel inicial de opiniones L sobre las H cuando el fraude es expuesto que luego decrece en sets intermedios para nuevamente aumentar en los sets finales. Esto da lugar a un patrón en forma de U invertida que no respondería a comportamiento agregado pasado alguno. En una encuesta pequeña realizada al final de algunas sesiones, algunos sujetos indicaban que entendían que una exposición era más probable que provenga de una Organización tipo L pero que ellos no entendían por qué una exposición pudo haber ocurrido en primer lugar. Esto sugiere que parte de este patrón puede ser resultado de confusión en sets intermedios debido a la misma presencia de la exposición del fraude en sets intermedios (donde la hipótesis de exposición errónea ya no es muy plausible), pero que finalmente desaparece a medida que la información del signo negativo de  $(P_H - P_L)$  retoma su relevancia en las decisiones del Público.

<sup>14</sup> Una lógica simple y consistente con esto es: “como observo exposiciones y una exposición es más probable que venga de una Organización tipo L, entonces por descarte una no exposición debería ahora ser más probable que provenga de una Organización tipo H”.

agregada está capturando la emergencia de un comportamiento intermedio particular entre los individuos, que será discutido en la siguiente sección.

### 4.1.3 Análisis de los Patrones de Juego a nivel Individual: Sesiones Total

Debido a que ninguna de las proporciones de juego para alguna Interpretación en la Tabla 3 fue observada en las sesiones Total, donde los participantes conocían de antemano el signo de (PH – PL), realizamos un análisis del comportamiento a nivel individual que nos ayude a entender el comportamiento a nivel agregado. Para esto a continuación agruparemos el comportamiento de los individuos y realizaremos el análisis en función de estos grupos.

#### 4.1.3.1 Clasificación del comportamiento individual por rol

En la **Figura 4** se grafica el comportamiento individual observado bajo el nivel de información Total clasificado en sesiones Negativo (barras negras) y Positivo (barras grises). El primer y segundo gráfico muestra el comportamiento individual para las Organizaciones y el Público respectivamente.

Para el rol de **Organización**, cuatro formas de comportamiento individual fueron identificadas: comportamiento de **interpretación negativa** (los sujetos nunca denuncian fraude cuando lo encuentran), comportamiento de **interpretación positiva** (los sujetos siempre denuncian el fraude cuando lo encuentran), comportamiento **condicional** (cuando los sujetos denuncian fraude siempre, nunca o algunas de las veces que lo encuentran condicional en el tipo de Organización), y comportamiento **indeciso** (cuando los sujetos denuncian fraude algunas de las veces que encuentran sin importar el tipo de Organización).

Para el rol de **Público**, cinco formas generales de comportamiento fueron identificados: comportamiento de **interpretación negativa** (los sujetos siempre opinaban L cada vez que el fraude era expuesto), comportamiento de **interpretación positiva** (los sujetos opinaban H cada vez que el fraude era expuesto y L cada vez que NO era expuesto), comportamiento **cuasi-positivo** (los sujetos opinaban H cada vez que el fraude era expuesto y opinaban H, L o estaban indecisos cuando NO era expuesto), comportamiento **indeciso** (los sujetos opinaban H, L o indeciso tanto cuando se exponía como cuando no se exponía fraude), y **otro** comportamiento (los sujetos opinaban H, L o indeciso cuando se exponía fraude y L cuando no).

Para cada rol los sujetos fueron clasificados de acuerdo a las formas anteriores de comportamiento. Muchos individuos consistentemente mantuvieron el mismo comportamiento a lo largo de toda la sesión pero algunos lo cambiaron. Bajo cada forma de comportamiento incluimos aquellos individuos que mantuvieron ese comportamiento por la sesión entera y a aquellos que terminaron la sesión con el mismo. La mayoría de los sujetos no tuvieron una forma consistente de comportamiento en ambos roles (inclusive algunos se comportaron en direcciones opuestas en cada rol), por lo que analizaremos ambos roles por separado.

#### 4.1.3.2 Análisis del comportamiento individual por rol: Sesiones Total.

En la **Tabla 8** se presenta las proporciones de juego individual y se realizan tests estadísticos para verificar la existencia de diferencias en comportamiento entre las sesiones de distinto signo. En la **Figura 4** y **Tabla 8** se evidencia la existencia de grandes diferencias en el comportamiento individual observado de acuerdo al signo de  $(P_H - P_L)$ , que se resumen en la siguiente conclusión:

**CONCLUSION TRES (Comportamiento Individual en sesiones Total):**

*a) Sesiones Negativo: la mayor parte de los sujetos presentan un comportamiento de interpretación negativa ó condicional al tipo en el rol de Organización, y un comportamiento de interpretación negativa en el rol de Público.*

*b) Sesiones Positivo: la mayor parte de los sujetos presentan un comportamiento de interpretación positiva y el resto un comportamiento condicional al tipo en el rol de Organización, y la mayor parte presenta un comportamiento de interpretación positiva ó cuasi-positiva en el rol de Público.*

Cuando el nivel de información es Total se esperaba un gran impacto del conocimiento del signo de  $(P_H - P_L)$  en los sujetos bajo el rol de **Organización** debido a que éstos son capaces también de observar sus propios pagos en cada ronda.

Como se puede observar en la Tabla 8, en las sesiones **Negativo**, bajo el rol de Organización, el 35% de los sujetos tienen o terminan la sesión con una Interpretación Negativa, también hay un 44% de los sujetos con un comportamiento condicional, un 13% indeciso y solo un 8% con comportamiento de interpretación positiva. En cambio para las sesiones **Positivo**, bajo el rol de Organización, el 72% de los sujetos tienen o terminan la sesión con una Interpretación positiva y el restante 28% con un comportamiento condicional. La última columna de la Tabla 8 evidencia una diferencia estadísticamente significativa en las proporciones de comportamiento de las Organizaciones entre las sesiones de ambos signos, excepto en el comportamiento condicional.

Con respecto a las sesiones **Negativo**, lo anterior explica por qué el nivel de exposición del fraude no terminó siendo cero como se esperaba bajo una Interpretación Negativa: 1 de cada 12 sujetos tuvo una interpretación positiva, 1.5 de cada 12 un comportamiento indeciso, y 5 de cada 12 un comportamiento condicional (4.2 de cada 12 sujetos consistentemente mantuvo algún tipo de comportamiento condicional durante toda la sesión). A grosso modo los sujetos indecisos y con interpretación positiva explicarían aproximadamente la mitad del 30% de exposición promedio observada, y los sujetos con comportamiento condicional explicarían la otra mitad.<sup>15</sup> Las dos primeras formas de comportamiento las encontramos tanto en las sesiones con contexto con significado como en las neutrales, por lo que creemos que debe ser el resultado de que algunos sujetos no comprendieron la relevancia de la información sobre  $P_H$  y  $P_L$ , o de racionalidad limitada, o porque algunos sujetos pueden encontrar poco intuitivo que una Organización tipo L encuentre fraude más veces que una tipo H. El comportamiento condicional también fue observado en ambos contextos y de manera consistente en una

---

<sup>15</sup> El número total esperado de descubrimientos de fraude en cada set de seis rondas es de 18. El número esperado de exposición de fraude en cualquier set por el sujeto con una interpretación positiva es de 1.5, y por los 1.5 sujetos indecisos (suponiendo exposición aleatoria) es de 1.125. Por lo tanto, el primer comportamiento explica aproximadamente 8.3 puntos porcentuales, y el segundo explica 6.3 puntos del 30% observado. Además del error muestral hay que considerar que algunos sujetos terminaron la sesión con estas forma de comportamiento, por lo que lo anterior son solamente aproximaciones.

gran proporción de sujetos, por lo que suponemos que no es un comportamiento excepcional sino un potencial candidato a un comportamiento de largo plazo para parte de la población cuando el signo negativo de  $(P_H - P_L)$  es de conocimiento mutuo (tal vez debido a racionalidad limitada).

Con respecto a las sesiones **Positivo**, lo anterior explica por qué el nivel de exposición de fraude fue muy alto pero no llegó al 100% esperado bajo una Interpretación positiva: 3.4 de cada 12 sujetos tuvieron o terminaron la sesión con un comportamiento condicional como Organización. La ausencia total de sujetos con interpretación negativa o indecisos nos permite descartar las dos primeras explicaciones expuestas en las sesiones Negativo para comportamientos equivalentes (interpretación positiva e indecisos) y apoya la hipótesis de que para los sujetos es más intuitivo que una Organización tipo H encuentre más fraude que una tipo L. Por otro lado, a pesar de que la proporción de sujetos con comportamiento condicional es menor (pero no estadísticamente distinta) a la observada en las sesiones Negativo, es notable encontrarla también en las sesiones Positivo y da mayor sustento al supuesto de que corresponde a un potencial comportamiento de largo plazo en parte de la población cuando el signo, positivo o negativo, de  $(P_H - P_L)$  es conocido.

Para el rol de **Público** esperábamos un impacto más moderado del signo de  $(P_H - P_L)$  en el comportamiento de los sujetos debido a su limitación de no observar el tipo de la Organización en cada ronda y decidir solo en base al conocimiento de las probabilidades, por lo que debe razonar en función de pagos esperados sin retroalimentación de ningún tipo.

Como se puede observar en la Tabla 8, en las sesiones **Negativo**, bajo el rol de Público, el 64% de los sujetos tiene o termina con una Interpretación Negativa, un 19% están indecisos y un 17% presentan un comportamiento cuasi-positivo. En las sesiones Positivo, un 44% de los sujetos termina con una Interpretación positiva, un 39% presentan un comportamiento cuasi-positivo, y el restante 17% termina con diferentes formas de comportamiento (1 de 36 sujetos tuvo una interpretación negativa). La última columna de la Tabla 8 evidencia una diferencia estadísticamente significativa en las proporciones de comportamiento del Público entre las sesiones de ambos signos, excepto en el comportamiento indeciso.

Vale la pena destacar que en las sesiones **Negativo**, una mayor proporción de los sujetos presentan una interpretación negativa como Público que como Organización. En las sesiones **Positivo**, 4.7 de cada 12 sujetos presentan un comportamiento cuasi-positivo, lo cual sugiere que es un potencial comportamiento de largo plazo en el Público cuando el signo positivo de  $(P_H - P_L)$  es conocido, es decir, se espera que algunos sujetos posean dificultades para interpretar una no exposición aunque comprendan que una exposición es más probable que provenga de una Organización tipo H. Es el surgimiento de este tipo de comportamiento lo que estaría capturando la variable agregada  $H - L / \expone (-1)$  en la 3ra estimación de la Tabla 7 para el tratamiento Total – Positivo.

También señalamos que la menor proporción de sujetos con un comportamiento de interpretación negativa, indeciso u otro en las sesiones Positivo (2 de cada 12) con respecto a la proporción de sujetos con un comportamiento cuasi-positivo o indeciso en

las sesiones Negativo (4.2 de 12), podría deberse a que los sujetos encuentran la estructura de las sesiones Negativo poco intuitiva.

#### 4.1.4 Análisis de los resultados encontrados en Validación de la Teoría

Los resultados de las sesiones Total-Negativo y Total-Positivo ponen en evidencia la validez de la teoría como un buen predictor cualitativo, es decir, de la dirección del comportamiento observado. Ninguno de los equilibrios esperados bajo una interpretación negativa fueron estrictamente jugados en las sesiones **Negativo**, aunque el comportamiento observado ofrece mayor soporte a una interpretación negativa que a una positiva: hubo un 30% de exposición promedio de fraude y un 75% de opiniones L cuando el fraude fue expuesto. Asimismo, ninguno de los equilibrios esperados bajo una interpretación negativa o positiva fueron estrictamente jugados en las sesiones **Positivo**, aunque el comportamiento observado ofrece mayor soporte a una interpretación positiva: hubo un 92% de exposición del fraude encontrado, un 87% de opiniones H cuando el fraude fue expuesto y un 72% de opiniones L cuando no fue expuesto. Como el comportamiento observado fue similar en ambos contextos (para cada signo), lo anterior ofrece suficiente evidencia para resolver la indeterminación teórica y a la vez dar mayor sustento el argumento de que la Autorregulación puede funcionar por motivos reputacionales en el caso en que es teóricamente factible (cuando en equilibrio es más probable que la OA encuentre fraude al aumentar su vigilancia).

El análisis del comportamiento a nivel agregado permite establecer claramente que para los sujetos no solamente es importante el signo de  $(P_H - P_L)$  en su comportamiento, sino también el cómo va evolucionando el juego del resto de los participantes. En particular, en las sesiones **Negativo**, para mantener la proporción de exposición de fraude bajo el 50% es necesario que los participantes observen en sets pasados que (en promedio) el resto prefiere opinar L en lugar de H cuando se expone fraude (por lo que es una mala idea exponer); asimismo, para observar poco a poco más opiniones H que L cuando el fraude no es expuesto (lo cual sustenta el equilibrio de mala interpretación 1b) es necesario que los participantes observen que en sets pasados los demás empiezan a presentar este mismo comportamiento, aunque también ayuda observar exposición de fraude. Asimismo, en las sesiones **Positivo**, para mantener la proporción de opiniones H por encima de las opiniones L cuando el fraude es expuesto es muy importante que los participantes observen que en sets pasados el resto presenta claramente este mismo comportamiento; por otro lado, observar en sets pasados que los demás prefieren opinar H en vez de L cuando el fraude es expuesto explica que exista en promedio menos gente opinando L en lugar de H cuando el fraude no se expone (este caso captura el surgimiento de un comportamiento individual particular de largo plazo: opinar H cuando exponen fraude y estar indeciso cuando no).

El análisis a nivel individual aclara el comportamiento agregado observado. Por ejemplo, en las sesiones **Negativo**, la presencia de sujetos que de manera consistente exponen siempre, nunca o a veces condicional en el tipo de Organización que les toque, y la presencia de sujetos que encuentran poco intuitivo que una Organización tipo L encuentre más fraude que una Organización tipo H, explican la existencia de un nivel de exposición distinto de cero; asimismo, la poca intuición del signo negativo de  $(P_H - P_L)$  también permite explicar, en parte, por qué el Público no opinó L cada vez que la Organización expuso fraude. En las sesiones **Positivo**, la presencia de sujetos que actuaban consistentemente condicional al tipo de Organización permite explicar por qué

el nivel de exposición no alcanzó el 100%; la presencia de sujetos indecisos permite explicar por qué el Público no opinó H cada vez que el fraude fue expuesto, y la presencia de sujetos a los cuales se les hizo difícil interpretar una no exposición permite explicar por qué el Público no opinó L cada vez que el fraude no fue expuesto.

## **4.2 Robustez de la Teoría: SESIONES NULO**

### **4.2.1 Evolución y diferenciación empírica del comportamiento observado**

Como fue notado en la sección 2, el conocimiento común de  $P_H$  y  $P_L$  no es un punto de partida creíble en cualquier mercado real aunque, sin embargo, puede ser el resultado eventual de un proceso social de ajuste y aprendizaje largo tanto para la OA como para el Público. Este punto es muy importante al momento de analizar mercados de confianza dada la severa limitación en la habilidad del Público para observar fraude (mala calidad, que en nuestro modelo corresponde al tipo de OA) y aprender de esta experiencia. Como resultado de lo anterior, ningún proceso de ajuste de creencias a partir de las experiencias exitosas o fracasos pasados puede ser esperado del Público. Sin embargo, la OA puede observar sus experiencias pasadas (la opinión que tuvo el Público ante sus acciones) para ajustar su comportamiento, y a la vez está consciente de las limitaciones del Público, lo cual puede tratar de usar en su propio beneficio.

Por todo lo anterior las sesiones bajo un nivel de información Nulo son una oportunidad para testear la robustez predictiva de la teoría en un ambiente más real, donde cada individuo parte con su propia idea sobre la estructura del juego (el signo de  $(P_H - P_L)$ ), pero puede experimentar dicha estructura y ajustar en el camino su comportamiento. En resumen, en las sesiones Nulo se puede analizar si: 1) los sujetos son capaces de aprender las probabilidades estructurales  $P_H$  y  $P_L$  de un mercado de confianza; y, 2) si son capaces de reaccionar a dicha estructura y ajustar su comportamiento en la dirección observada en las sesiones bajo un nivel de información Total. Por lo tanto, nuestro análisis se centrará principalmente en comparar el comportamiento observado en los sets de seis rondas iniciales y finales.

En la **Tabla 9** se muestran las proporciones de juego observado en las sesiones Nulo para cada acción posible (con un intervalo del 95%) que se obtienen al promediar las primeras 18 rondas (tres primeros sets) y las últimas 18 rondas (tres últimos sets) de todas las sesiones, de acuerdo al signo de  $(P_H - P_L)$ . Como se puede apreciar y era de esperarse, en los primeros tres sets no existe diferencia estadísticamente significativa en el comportamiento promedio observado de acuerdo al signo de  $(P_H - P_L)$ : los promedios son similares y los intervalos de confianza se intersectan en todos los casos. En cambio, en los últimos tres sets sí existe una diferencia estadísticamente significativa en el comportamiento promedio observado de acuerdo al signo de  $(P_H - P_L)$ : los promedios son distintos y los intervalos de confianza no se intersectan en todos los casos, excepto en la proporción del Público indeciso cuando la Organización expone y no expone fraude. Lo anterior nos lleva a nuestra cuarta conclusión:

**CONCLUSION CUATRO (Evolución y diferenciación del juego según el signo de  $(P_H - P_L)$ ):** *En promedio, al inicio de las sesiones Nulo no existe diferencia en el comportamiento de los individuos de acuerdo al signo de  $(P_H - P_L)$ ; sin embargo, al final de estas sesiones SI existe una clara diferencia en el comportamiento de los individuos de acuerdo al signo de  $(P_H - P_L)$ . En las sesiones Negativo el*

*comportamiento parece dirigirse hacia las proporciones observadas en las sesiones Total – Negativo (hacia una mala interpretación), y en las sesiones Positivo el comportamiento parece dirigirse lentamente hacia las proporciones observadas en las sesiones Total – Positivo (hacia una buena interpretación).*

En las sesiones Negativo la proporción de exposición promedio del fraude baja de un 56% inicial a un 43% final. La proporción de opiniones H cuando el fraude es expuesto baja de un 53% a un 40%, y la de opiniones L cuando el fraude es expuesto sube de un 38% a un 50%. La proporción de opiniones H cuando el fraude no es expuesto sube ligeramente de un 50 a un 54%, y la de opiniones L cuando el fraude no es expuesto baja de un 33% a un 26%. Es decir, el comportamiento tiende en promedio hacia las proporciones observadas en las sesiones Total – Negativo (30%, 17%, 75%, 55% y 28% respectivamente).

Por otro lado, en las sesiones Positivo la proporción de exposición promedio del fraude sube de 58% inicial a un 68% final. La proporción de opiniones H cuando el fraude es expuesto sube de un 52% a un 64%, y la de opiniones L cuando el fraude es expuesto baja ligeramente de un 35% a un 31%. La proporción de opiniones H cuando el fraude no es expuesto baja de un 42% a un 32%, y la de opiniones L cuando el fraude no es expuesto sube de un 35% a un 46%. Es decir, el comportamiento tiende lentamente y en promedio hacia las proporciones observadas en las sesiones Total – Positivo (92%, 87%, 9%, 14% y 72% respectivamente).

Como se señaló anteriormente, el comportamiento en los sets finales es clara y estadísticamente distinto entre las sesiones de ambos signos, es decir, existe una clara diferenciación en el comportamiento de los individuos en la dirección de una MALA interpretación de la exposición del fraude en las sesiones Negativo y en la dirección de una BUENA interpretación en las sesiones Positivo.

Sin embargo, en las sesiones Nulo se observó diferencias de comportamiento entre las sesiones realizadas bajo un contexto con significado y neutral. En la **Tabla 10** se presenta la evolución media de las acciones tomadas por las Organizaciones (proporción de exposición de fraude) y Público (diferencia entre las proporciones de opiniones H y L cuando se expone y no se expone fraude) de acuerdo al signo de  $(P_H - P_L)$  y contexto.

Como se puede observar, en las sesiones Negativo no existen drásticas diferencias en el patrón de comportamiento observado: la proporción de exposición al fraude baja en ambos contextos entre el primer y los dos últimos sets (de 75% a 51% en contexto con significado, y de 59% a 36% con contexto neutral); en ambos contextos los sujetos cambian de opinar más veces H en lugar de L cuando el fraude es expuesto en el primer set, a opinar más veces L en lugar de H en los dos últimos sets (la diferencia de proporciones entre H y L pasa de 3% a -15% en contexto con significado, y de 32% a -33% en contexto neutral); también en ambos contextos los sujetos aumentan su preferencia por opinar más veces H en lugar de L cuando el fraude no es expuesto entre el primero y los dos últimos sets (la diferencia de proporciones entre H y L aumenta de 5% a 25% en contexto con significado y de 0% a 32% en contexto neutral).

Donde se observan diferencias marcadas en el patrón de comportamiento observado al inicio y final es en las sesiones Positivo: la proporción de exposición al fraude aumenta en el contexto con significado entre el primer y los dos últimos sets (pasa de 75% a



82%) y se mantiene casi igual pero menor en el contexto neutral (pasa de 53% a 51%); los sujetos aumentan su preferencia por opinar más veces H en lugar de L cuando el fraude es expuesto entre el primer y los dos últimos sets en el contexto con significado (la diferencia de proporciones entre opiniones H y L aumenta de 12% a 51%), en cambio en el contexto neutral disminuye esta preferencia (diferencia pasa de 11% a 0%); los sujetos muestran un drástico cambio en su preferencia por opinar más veces H en lugar de L cuando el fraude no es expuesto en el primer set a opinar más veces L en lugar de H en los dos últimos sets en el contexto con significado (la diferencia de proporciones entre opiniones H y L pasa de 33% a -20%), y en el contexto neutral se observa el mismo patrón pero de manera mucho menos drástica (diferencia pasa de 2% a -9%). Al observar el desarrollo set por set de las sesiones Positivo con contexto neutral lo que se observa, más que una tendencia hacia una Interpretación Negativa (equilibrio 1d), es una tendencia a mantener el comportamiento cerca del punto de partida. Estas diferencias nos llevan al siguiente corolario:

**COROLARIO UNO (Diferencias entre contextos en sesiones Nulo):** *No existen grandes diferencias en el patrón de comportamiento observado en las sesiones Negativo conducidas bajo un contexto con significado y neutral: en ambos el comportamiento parece tender hacia una Mala Interpretación. En cambio, solamente en las sesiones Positivo conducidas bajo un contexto con significado se observa un comportamiento que tiende hacia una Buena Interpretación; las sesiones Positivo conducidas bajo un contexto neutral no evidencia un patrón claro de comportamiento hacia alguna Interpretación.*

#### **4.2.2 Análisis de los Patrones de Juego en el Agregado: Sesiones Nulo**

Al igual que al analizar el comportamiento agregado en las sesiones Total, en las sesiones Nulo los individuos tienen dos factores que pueden explicar su comportamiento en los sets finales: 1) el **aprendizaje** del signo de  $(P_H - P_L)$  de la estructura subyacente de la sesión; y, 2) la reacción de los jugadores a las decisiones que van tomando el resto. Es decir, por un lado los jugadores pueden aprender el signo de  $(P_H - P_L)$  subyacente en la estructura de la sesión (al experimentar ambos tipos de Organización) y ajustar su comportamiento o, por otro lado, van ajustando su juego de acuerdo a como observan o infieren que van a jugar el resto de los participantes.

Para poder separar debidamente estos dos efectos y diferenciarlos de otros comportamientos (como heterogeneidad en el comportamiento), también utilizamos un modelo Probit de datos de panel con efecto aleatorio que difiere un poco al anterior, y cuya metodología e interpretación de resultados explicamos a continuación.

##### **4.2.2.1 Metodología de identificación de motivación de juego en el agregado.**

Como en las sesiones Total, se estimaron dos modelos que difieren en el set de variables dependientes, con tres estimaciones Probit en cada modelo. Las variables dependientes de cada estimación son las mismas del caso anterior.

En la **Tabla 11** se presentan los resultados del primer modelo. A diferencia de las sesiones Total, en este caso el modelo utiliza la información de todos los sets de seis rondas. Las tres estimaciones que aparecen en el primer modelo son el resultado de un proceso de dos etapas: en la primera etapa se incluye una dummy por el primer set, una

dummy por las sesiones bajo un contexto neutral en el primer set, dummies para los sets 2-3, 4-5, 6-7 interactuadas con dummies para los tratamientos Nulo – Negativo y Nulo – Positivo, y dummies por set y tratamiento bajo un contexto neutral. Las dummy del primer set recupera el comportamiento para las sesiones bajo un contexto con significado en ese set. La dummy del primer set y contexto neutral recupera cualquier desviación (con respecto al contexto con significado) en el comportamiento medio en ese set por el uso de lenguaje neutral. Las dummies por set y tratamiento tienen el mismo uso e interpretación que para las sesiones Total (en especial lo relativo al punto medio) pero para un contexto con significado. Por último, las dummies por set y tratamiento bajo un contexto neutral recuperan cualquier desviación (con respecto al contexto con significado) en los siguientes sets por el uso de lenguaje neutral. En la segunda etapa, se restringe el modelo estimado y se elimina las dummies por set y tratamiento conducidos bajo un contexto neutral que resultaron no significativas en la etapa anterior.<sup>16</sup>

El primer modelo recoge y da un mejor sustento estadístico a dos aspectos, una vez controlada la heterogeneidad: 1) recoge cualquier diferencia estadísticamente significativa al inicio de las sesiones entre un contexto con significado y neutral; y, 2) recoge las diferencias observadas en el Corolario Uno entre contextos (en especial para las sesiones Positivo).

Hay que notar que, de acuerdo a lo que hemos venido señalando, las dummies por set y tratamiento recogen tanto el efecto que el aprendizaje del signo de  $(P_H - P_L)$  tiene sobre las decisiones de los participantes como cualquier proceso de ajuste por la evolución del juego. Para separar ambos efectos estimamos el segundo modelo, cuyos resultados se presentan en la **Tabla 12**. Este modelo incluye las mismas variables dependientes finales del primer modelo y adicionalmente incorpora las mismas tres variables promedio agregadas utilizadas anteriormente, que ahora interactúan con dummies para cada tratamiento y contexto (por lo que se pierde el primer set).

Como en el caso anterior, el segundo modelo es el resultado de un proceso de dos etapas que termina con aquellas variables promedio agregadas que son significativas hasta en un 10% de significancia (se incluyen todas las dummies por set, tratamiento y contexto del primer modelo en ambas etapas). Asimismo, este proceso es robusto a otros niveles de corte en significancia y siempre escoge las mismas variables agregadas. De esta manera se logra identificar de una manera más precisa el efecto de las variables agregadas para los tratamientos y contexto correctos. Una vez que controlamos los procesos de ajuste mediante variables agregadas y la heterogeneidad entre los participantes, las dummies por set y tratamiento deben recuperar la mayoría del efecto del aprendizaje del signo de  $(P_H - P_L)$  en las decisiones de los participantes.

#### **4.2.2.1 Interpretación de los resultados de motivación del juego a nivel agregado**

Al comparar las Tablas 11 y 12 se observan algunas diferencias en cada estimación, una vez identificado y controlado el efecto de las variables agregadas y heterogeneidad. Las principales diferencias las resumimos en la siguiente conclusión:

***CONCLUSION CINCO (Efecto del signo de  $(PH - PL)$  y comportamiento agregado):***

---

<sup>16</sup> Estas restricciones fueron ampliamente aceptadas a niveles convencionales de significancia en tests Z y LRT.

a) *Sesiones Negativo:*

a1. *Con contexto: para que los niveles de exposición se mantengan bajos es importante observar en sets pasados que la proporción de sujetos que opinan que una organización que expone tiene mucho fraude (L) crezca. El efecto del aprendizaje del signo de negativo de  $(P_H - P_L)$  es muy importante para observar un crecimiento en la tasa de sujetos que opina que una organización que expone tiene mucho fraude (L), sin embargo la presencia de exposición de fraude en sets pasados confunde a los sujetos y eleva la proporción de los que opinan lo contrario (H). Para observar un crecimiento en la tasa de sujetos que opina que una organización que NO expone tiene poco fraude (H) es importante el aprendizaje del signo negativo de  $(P_H - P_L)$  y también observar el mismo comportamiento en sets pasados.*

a2. *Neutral: el aprendizaje del signo negativo de  $(P_H - P_L)$  es suficiente para observar un descenso en la tasa de exposición de fraude, y un crecimiento en la tasa de sujetos que opina que una organización que NO expone tiene poco fraude (H). El comportamiento del público cuando la organización expone fraude es igual al del uso de contexto.*

b) *Sesiones Positivo:*

b.1 *Con contexto: el aprendizaje del signo positivo de  $(P_H - P_L)$  es suficiente para observar un crecimiento en la tasa de exposición de fraude, una elevada tasa de sujetos que opinan que una organización que expone tiene poco fraude (H), y un crecimiento en la tasa de sujetos que opina que una organización que NO expone tiene mucho fraude (L).*

b.2 *Neutral: la tasa de exposición de fraude fluctúa alrededor del 50%, aunque decrece débilmente cuando la tasa de sujetos que opina que una organización que expone tiene mucho fraude (L) crece en sets pasados; la tasa de sujetos que opina que una organización que expone tiene poco fraude (H) termina igualando la tasa de los que opinan lo contrario (L); el único efecto que se observa debido al aprendizaje del signo de  $(P_H - P_L)$  es en la opinión del público ante una NO exposición, al crecer la tasa de sujetos que opina que tiene mucho fraude (L).*

Con respecto a las sesiones Negativo conducidas con un contexto con significado vale la pena destacar el efecto de retroalimentación entre exposición del fraude y la opinión del público frente a una exposición. Como se señala en la conclusión, para mantener los niveles de exposición bajos es necesario observar en sets pasados que los sujetos opinan más veces L que H cuando ocurre una exposición, y por otro lado la presencia de exposición de fraude produce un efecto de confusión en los sujetos que produce el efecto contrario y eleva el nivel de opiniones H en vez de L ante una exposición (lo cual se puede inferir de la súbita recuperación de la significancia de las dummies por set y tratamiento en el segunda estimación de la Tabla 12, una vez controlado el efecto de la exposición). El efecto combinado de ambos puede perfectamente resultar en una situación en que se mantienen altos niveles de exposición y de opiniones H ante una exposición, a pesar del aprendizaje del signo negativo de  $(P_H - P_L)$ , como efectivamente parece ocurrir en una de las sesiones.

En las sesiones Negativo conducidas con un contexto Neutral también se observó el efecto de confusión que genera la presencia de exposición en la opinión del Público, aunque en este caso no existe una retroalimentación hacia el nivel de exposición de fraude. Esto indica que el aprendizaje del signo negativo de  $(P_H - P_L)$  tiene una mayor relevancia para los sujetos al momento de decidir si exponen o no fraude, y explica el nivel medio de exposición más bajo observado en las rondas finales de este tipo de

sesión (en la Tabla 10 se puede observar que terminan con un 36% de exposición media de fraude frente a un 51% de las sesiones con contexto con significado).

Con respecto a las sesiones Positivo conducidas bajo un contexto con Significado, vale la pena destacar el papel fundamental que tiene el aprendizaje del signo positivo de ( $P_H - P_L$ ) para determinar la dirección hacia la cual se dirige el comportamiento agregado.

Por último, en las sesiones Positivo conducidas bajo un contexto Neutral se puede observar el mismo comportamiento anotado en el Corolario Uno, ya que solamente el nivel de opiniones cuando NO se expone fraude tiende hacia el comportamiento esperado bajo una Interpretación positiva, y el resto de las acciones presentan más bien un comportamiento errático y al azar.

#### **4.2.3 Análisis de los Patrones de Juego a nivel Individual: Sesiones Nulo**

El comportamiento individual para las sesiones Nulo fue clasificado de la misma forma que en las sesiones Total. Los resultados se pueden observar en la **Tabla 13** donde se presenta las proporciones de juego individual y se realizan tests estadísticos para verificar la existencia de diferencias en comportamiento entre las sesiones de distinto signo y contexto.

##### **4.2.3.1 Análisis del comportamiento individual por rol: Sesiones Nulo.**

En la **Tabla 13** se evidencia la existencia de algunas diferencias en el comportamiento individual observado de acuerdo al signo de ( $P_H - P_L$ ) y contexto. Consistente con el comportamiento observado a nivel agregado, no existe diferencias significativas en el comportamiento individual de las sesiones Negativo de acuerdo al contexto utilizado (estadísticos no se muestran en la Tabla) por lo que las unificamos. En cambio sí se puede apreciar diferencias significativas en el comportamiento individual, tanto del Público como de las Organizaciones, de las sesiones Positivo de acuerdo al contexto utilizado, por lo que ambos contextos fueron separados. Debido a que el contexto Neutral en las sesiones Positivo no muestra un patrón en el agregado, en la siguiente conclusión solamente se considera el contexto con significado de las mismas:

##### ***CONCLUSION SEIS (Diferencias en comportamiento Individual en sesiones Nulo):***

*En el rol de Organización no hay una clara diferencia en el comportamiento individual entre las sesiones Negativo y Positivo con significado, aunque marginalmente hay más sujetos que tienen o terminan con una interpretación positiva en las últimas.*

*En el rol de Público existe una mayor cantidad de sujetos que terminan la sesión con una interpretación negativa en las sesiones Negativo, y una mayor proporción de sujetos que terminan con una interpretación cuasi-positiva o positiva en las sesiones Positivo con significado.*

Como se puede apreciar en la Tabla 13 para el rol de Organización, en las sesiones **Negativo** el 52% de los sujetos tienen o terminan la sesión con una Interpretación condicional al tipo de organización, el resto termina con una interpretación Positiva (19%), Indeciso (15%) ó Negativa (15%). En cambio, en las sesiones **Positivo** (con significado), el 44% de los sujetos tienen una Interpretación condicional al tipo, un 33% tiene una interpretación positiva y el resto tiene una interpretación negativa (11%) o están indecisos (11%). Estadísticamente no existe diferencia en las distribuciones entre

las sesiones Positivo y Negativo, no obstante se observa una mayor proporción de sujetos con una interpretación positiva en las primeras, lo que explica el marcado crecimiento en la tasa de exposición de fraude encontrado en éstas.

En la misma Tabla se puede apreciar que existe una clara diferencia, estadísticamente significativa, en el comportamiento individual observado para el rol de Público. En las sesiones **Negativo**, la gran mayoría de los sujetos tienen o terminan con una interpretación negativa (48%) ó cuasi-positiva (27%). En cambio, en las sesiones **Positivo**, la gran mayoría de los sujetos tienen o terminan con una interpretación cuasi-positiva (47%) ó Indeciso (31%). La diferencia entre ambos tipo de sesión está en la cantidad de sujetos con interpretación negativa, positiva y cuasi-positiva, que se encuentran mayormente en las sesiones negativo, positivo y positivo respectivamente.

Lo anterior sugiere que el impacto conjunto del aprendizaje de la estructura y cualquier proceso de ajuste es más pronunciado para el Público y marginal para la Organización. Esto no se esperaba en un principio, por la limitación informacional del Público y sugiere que justamente la falta de retroalimentación incrementaría su deseo de buscar otras fuentes de información que le permitan tomar una mejor decisión.

#### **4.2.4 Análisis de los resultados encontrados en Robustez de la Teoría**

Las conclusiones que se pueden obtener de los resultados de las sesiones Nulo-Negativo y Nulo-Positivo varían de acuerdo al contexto utilizado. En las sesiones con un contexto Neutral se observa evidencia débil que sustenta a la teoría como un predictor del comportamiento en el largo plazo: solamente las sesiones Nulo-Negativo tienden hacia el equilibrio de una Mala Interpretación observado en las sesiones Total-Negativo, en cambio que las sesiones Nulo-Positivo no presentan ningún patrón claro de comportamiento. En cambio, en las sesiones conducidas con un contexto con significado se observa clara evidencia de la robustez de la teoría: el comportamiento en las sesiones Nulo-Negativo y Nulo-Positivo presentan un patrón claro hacia el final en la dirección observada en las sesiones Total-Negativo y Total-Positivo respectivamente, es decir, hacia una Mala y Buena Interpretación. Esto nos indica que el uso del contexto pudo haber tenido un rol facilitador en la comprensión estratégica del juego por parte de los individuos, según era esperable de acuerdo a la literatura [1].

Los resultados a nivel agregado nos alertan de la presencia de procesos de ajuste ante el comportamiento agregado cuya influencia en el patrón de juego pueden ser tan importantes como el aprendizaje del signo de  $(P_H - P_L)$ , en especial en las sesiones Negativo donde estos procesos pueden llevar a tener patrones en clara oposición al signo negativo de  $(P_H - P_L)$ , como el de un alto nivel de exposición de fraude con la intención de confundir al Público.

La evidencia a nivel individual es consistente con el patrón de comportamiento hacia las diferentes interpretaciones y se constata que es el Público aquel que presenta un patrón más fuerte, debido a su necesidad de suplir su deficiencia informacional mediante el aprendizaje del signo de  $(P_H - P_L)$  subyacente en la estructura de cada sesión.

## **5. RESUMEN Y CONCLUSIONES**

En este trabajo tomamos una estrategia empírica experimental para establecer la validez del motivo reputacional para el correcto funcionamiento de la Autorregulación en mercados de bienes de confianza.

En nuestro modelo teórico, basado en Nuñez [10], dicho incentivo reputacional existe si en el equilibrio bayesiano del juego de señalización en el cual la organización autorregulada (OA) decide exponer o no cualquier evidencia de fraude encontrada, el Público tiene una **Buena Interpretación** de la exposición de fraude, es decir, opina que la organización realizó mucha vigilancia sobre sus miembros y son pocos los que efectúan fraude, en lugar de una **Mala Interpretación**, es decir, opina que la organización realizó poca vigilancia sobre sus miembros y son muchos los que efectúan fraude. En el primer caso, la organización tiene incentivos reputacionales para exponer evidencia de fraude, en cambio que en el segundo caso no los tiene.

Sin embargo, dadas las probabilidades de equilibrio de descubrimiento de fraude por parte de una OA de alta vigilancia y poco fraude ( $P_H$ ), y de baja vigilancia y mucho fraude ( $P_L$ ), el modelo predice que una Buena Interpretación puede ocurrir solo cuando el signo de  $(P_H - P_L)$  es positivo, es decir cuando en equilibrio es más probable encontrar fraude al aumentar el nivel de vigilancia, en cambio que una Mala Interpretación puede ocurrir siempre sin importar el signo de  $(P_H - P_L)$ . Ningún refinamiento para equilibrios en juegos de señalización puede ser aplicado para obtener un resultado plausible para este juego, así como ningún modelo de aprendizaje existente puede ser aplicado debido esto último a que la severa limitación del Público para observar fraude y vigilancia le impide aprender de experiencias pasadas. Por lo tanto, utilizamos un enfoque empírico experimental para salvar las limitaciones de la teoría y encontrar cómo se comportan los sujetos humanos en esta situación.

Realizamos varias sesiones experimentales utilizando dos variables de control: el signo de  $(P_H - P_L)$ , y el nivel de información que tienen los sujetos acerca de estas probabilidades. El segundo control nos permite chequear tanto la validez como la robustez de la teoría, debido a que en un mercado de confianza real el supuesto de conocimiento común de las probabilidades de equilibrio  $P_H$  y  $P_L$  es cuestionable como punto de partida pero puede ser el resultado de un largo proceso de aprendizaje social tanto para la OA como para el Público, cuya convergencia o tendencia hacia alguna interpretación no puede ser garantizadas debido a las limitaciones informacionales del Público. Utilizamos un nivel de información TOTAL donde todos los sujetos conocían de antemano las probabilidades subyacentes en cada sesión, que nos permite testear la validez de la teoría como predictor del comportamiento a largo plazo; y un nivel de información NULO donde ningún sujeto conoce estas probabilidades pero existen en la estructura subyacente del juego, que nos permite testear la robustez de la teoría en una situación más real. Siguiendo a la literatura que sugiere que el uso de un lenguaje con significado en lugar de un lenguaje neutro en las instrucciones puede ayudar a los sujetos a entender la complejidad inherente en juegos de señalización, se realizaron algunas sesiones bajo un contexto con significado y otras bajo un contexto neutro.

Los resultados experimentales sustentan la validez de la teoría como un predictor cualitativo que explica el comportamiento observado en las sesiones TOTAL, y también despejan la incertidumbre teórica acerca de la interpretación más plausible para el caso en que el signo de  $(P_H - P_L)$  es positivo.

Ningún equilibrio fue estrictamente observado cuando el signo de  $(P_H - P_L)$  fue Negativo, sin embargo la tendencia en el comportamiento medio ofrece mayor soporte a una interpretación negativa: en promedio se observó un 30% de exposición fraude encontrado, y un 75% de las veces el público opinó que una exposición provenía de una organización con poca vigilancia y mucho fraude (L). Un análisis a nivel individual permite entender por qué no hubo una mayor convergencia hacia una mala interpretación: varios sujetos encontraron poco intuitivo que una organización con poca vigilancia y mucho fraude (L) encontrara fraude más veces que una organización con mucha vigilancia y poco fraude (H); asimismo, otros sujetos consistentemente creyeron que exponer fraude siempre, a veces o nunca dependiendo del tipo de organización era una estrategia óptima.

Ningún equilibrio fue estrictamente observado cuando el signo de  $(P_H - P_L)$  fue Positivo, sin embargo la tendencia en el comportamiento medio ofrece mayor sustento a una interpretación positiva: en promedio hubo un 92% de exposición de fraude encontrado, un 87% de las opiniones cuando hubo exposición favorecieron al tipo de mucha vigilancia y poco fraude (H), y un 72% de las opiniones cuando NO hubo exposición favorecieron al tipo de poca vigilancia y mucho fraude (L). A nivel individual también se observó varios sujetos con un comportamiento condicional al tipo de organización y algunos sujetos con dificultades para elegir al tipo de poca vigilancia como el mejor candidato ante una NO exposición aún cuando estaban seguros de elegir al tipo de mucha vigilancia ante una exposición de fraude.

Como el comportamiento medio para ambos signos de  $(P_H - P_L)$  fue similar en ambos contextos, los resultados anteriores son suficiente evidencia para resolver la indeterminación teórica para los equilibrios del juego y ofrece una clara oportunidad a la existencia de incentivos reputacionales para que la autorregulación funcione en el caso en que es teóricamente factible, es decir, cuando el signo de  $(P_H - P_L)$  es positivo.

Por otro lado, los resultados experimentales ofrecen resultados mixtos, dependiendo del contexto utilizado, sobre la robustez de la interpretación positiva ante la ausencia del conocimiento común inicial del signo positivo de  $(P_H - P_L)$ . Si tomamos los resultados de las sesiones con contexto Neutral para entender el comportamiento cuando ninguna información es inicialmente provista a los participantes, obtenemos muy malas noticias para la posible emergencia de una Interpretación Positiva como un proceso de aprendizaje social y por tanto para el argumento reputacional para que la autorregulación funcione: cuando el signo de  $(P_H - P_L)$  es positivo no se observa una tendencia clara del comportamiento hacia ninguna interpretación. Sin embargo, si tomamos los resultados de las sesiones con contexto con Significado, encontramos evidencia que soporta el surgimiento de una interpretación positiva cuando el signo de  $(P_H - P_L)$  es positivo. La Interpretación negativa es robusta cuando el signo de  $(P_H - P_L)$  es negativo, sin importar el contexto.

Por lo tanto, en las sesiones Nulo de contexto con significado, encontramos evidencia de que el contexto ayuda a los sujetos a aprender el signo de  $(P_H - P_L)$  y usar esa información para eventualmente ajustar su comportamiento en la dirección observada en las sesiones Total, y encontramos evidencia que ofrece un mayor sustento al surgimiento de una interpretación positiva cuando el signo de  $(P_H - P_L)$  es positivo.

Aunque en las sesiones experimentales el Público tuvo la oportunidad de actuar también como OA y de esa forma aprender el signo de  $(P_H - P_L)$ , esto no implica que en un mercado real el Público tenga que experimentar el papel de una OA, sino que de alguna manera creíble acceda a esta información (en nuestro caso fue experimentarlo durante la sesión). Algunas formas alternativas podrían ser, por ejemplo: una filtración por parte de algunos miembros de la OA a familiares y amigos (y su divulgación), vigilancia paralela externa a la industria como la de una Agencia de Protección al consumidor, experiencias documentadas en otras industrias de confianza, etc.

Estos resultados tienen relevancia al menos en dos áreas. En primer lugar, para el área de organización industrial de mercados de bienes de confianza: si el Público tiene alguna oportunidad de aprender o creíblemente inferir la estructura del mercado, entonces puede usar esta información para ajustar la manera cómo interpreta una exposición de fraude de parte de la OA y el mercado eventualmente terminará cerca de una buena o mala interpretación de equilibrio. En el largo plazo existirán incentivos para que la OA vigile, castigue y exponga el fraude encontrado al Público solo cuando al aumentar la vigilancia se logre encontrar fraude con mayor frecuencia.

Un aspecto relevante que surge de esta conclusión y que tiene implicancias directas para la generalización de los resultados, tiene que ver con las distintas formas en que el Público aprende o infiere la información del signo de  $(P_H - P_L)$  del mercado de confianza. Si el grado de asimetría de información entre Público y expertos es muy grande (es decir, es muy difícil o caro que un consumidor adquiera el conocimiento experto para evaluar el servicio), entonces es más plausible que el Público extrapole experiencias ó conocimientos previos adquiridos en situaciones similares o utilice convenciones sociales para formarse una idea de lo que ocurre, por lo que la estructura propia del mercado puede ser irrelevante al menos hasta que de alguna manera creíble el Público pueda aprenderla. Si por otro lado, el grado de asimetría no es muy grave (por ejemplo, el Público puede tener acceso a una segunda opinión, [11] y [12]), o existe alguna referencia sobre la estructura del mercado disponible, como la de monitoreos paralelos en el mercado (prensa de investigación u organismo de protección al consumidor, por ejemplo), entonces el Público puede aprovechar esta información y reaccionar ante la estructura actual del mercado. En ambos casos la dinámica social hacia alguna interpretación puede ser distinta, y hasta opuesta a la estructura en el primer caso, por lo que sería interesante de investigar en futuros trabajos su relevancia empírica.

En segundo lugar, los resultados encontrados tienen también profundas implicancias para la literatura de aprendizaje en juegos de señalización, debido a que hemos encontrado evidencia de que la ausencia de retroalimentación del resultado de las acciones tomadas por algunos jugadores (el Público) para actualizar sus creencias y acciones futuras puede ser superado mediante el aprendizaje de los parámetros cruciales del juego cuando éstos existen (como era el signo de  $(P_H - P_L)$ ). En las sesiones Nulo se constata que fue justamente el Público quien utiliza en mayor medida el aprendizaje del signo de  $(P_H - P_L)$  para suplir la falta de retroalimentación y ajustar su comportamiento inicial. Estos resultados abren toda una línea futura de investigación en el área de aprendizaje en juegos de información incompleta, ya que hasta el momento esta se ha enfocado en explicar el comportamiento en juegos donde los participantes son capaces de observar posteriormente el resultado de sus acciones [7].



## 6. BIBLIOGRAFIA

- [1] Cooper D., Kagel J. (2003). “The impact of meaningful context in Signalling Games”. *Journal of Economic Behavior and Organization*, 50, pp. 311 – 337.
- [2] Cooper D., Kagel J., Lo W. and Gu Q. (1999). “Games against managers in incentive systems: Experimental results with Chinese students and Chinese managers”. *American Economic Review*, 89, pp. 781 – 804.
- [3] Cooper D., Garvin S. and Kagel J. (1997). “Signalling and Adaptive Learning in an Entry Limit pricing Game”. *RAND Journal of Economics*, 28, pp. 662 – 683.
- [4] Cooper D., Garvin S. and Kagel J. (1997b). “Adaptive Learning vs. Equilibrium Refinements in an Entry Limit Pricing Game”. *The Economic Journal*, 107, pp. 553 – 575.
- [5] Emons W. (1997). “Credence Goods and fraudulent experts”. *RAND Journal of Economics*, Vol 28, No. 1, pp. 107 – 119.
- [6] Fudenberg D. and Tirole J. (1991). “Game Theory”. *MIT Press*.
- [7] Fudenberg D. and Levine D. (1998). “The Theory of Learning in Games”. *MIT Press*.
- [8] Kagel H. and Roth A. editors (1995). “Handbook of Experimental Economics”. *Princeton University Press*.
- [9] Nuñez, J. (1999). “Four Essays on Reputation and Self Regulation”. *Ph.D. in Economics dissertation, University of Oxford*.
- [10] Nuñez J. (2001a). “A Model of Self-Regulation”. *Economic Letters*, 74, pp. 91 – 97.
- [11] Pitchik C and Schotter A. (1994). “Norms and Competition in Markets with Asymmetric Information: an Experimental Study of the Development of Industry Ethics”. *Metroeconomica*, Vol. 45, No. 2, pp. 188 – 207.
- [12] Plott C. and Wilde L. (1982). “Professional diagnosis vs. Self diagnosis: an experimental examination of some special features of markets with uncertainty”. *Research in Experimental Economics*, Vol. 2, Vernon Smith ed., JAI Press.
- [13] Wooldridge J. (2002). “Econometric Analysis of Cross Section and Panel Data”. *MIT Press*.

## 7. ANEXO DE TABLAS Y FIGURAS

**TABLA 1: Estructura de Pagos para el Público y OA**

<b><i>PUBLICICO</i></b>	<b><i>Pago</i></b>
Si opinión coincide con el tipo de OA gana....	U
Si opinión NO coincide con el tipo de OA gana....	0
Si dice estar incierto (lotería)	0.5 U

<b><i>OA</i></b>	<b><i>Pago</i></b>
Si el público opina que es tipo H	W
Si el público opina que es tipo L	0
Si el público está incierto (lotería)	0.5 W

**TABLA 2: Proporción de juego esperada bajo cada Interpretación**

<b>ACCIONES POSIBLES</b>	<b>MALA Interpretación</b>	<b>BUENA Interpretación</b>
<b>OA (ambos tipos)</b>		
Exponer fraude si es encontrado	0%	100%
<b>PUBLICICO</b>		
Si la OA expone fraude		
Opinar "H"	0%	100%
Opinar "L"	100%	0%
Estar Indeciso	0%	0%
Si la OA NO expone fraude		
Opinar "H"	[ 0% , 100% ]	0%
Opinar "L"	[ 0% , 100% ]	100%
Estar Indeciso	[ 0% , 100% ]	0%

**TABLA 3: Probabilidad con la que cada Tipo de Organización logra encontrar fraude**

PH	PL	SIGNO DE (PH-PL)
0.2	0.8	NEGATIVO
0.8	0.2	POSITIVO

**TABLA 4: Controles, Tratamientos y contexto utilizado en las Sesiones Experimentales**

Nivel de Información	Signo de (PH - PL)	Tratamiento	Contexto utilizado	Sesiones
Total	Negativo	Total - Negativo	con Significado Neutral	1 - 2 - 3 4
Total	Positivo	Total - Positivo	con Significado Neutral	5 - 6 7
Nulo	Negativo	Nulo - Negativo	con Significado Neutral	8 - 9 10 - 11
Nulo	Positivo	Nulo - Positivo	con Significado Neutral	12 - 13 - 14 15 - 16

**TABLA 5: Proporción media observada para cada acción posible en las sesiones Total, por signo de (PH – PL)**

Acciones	NEGATIVO		POSITIVO		Interpret. Negativa	Interpret. Positiva
	Prom.	Int. 95% conf.	Prom.	Int. 95% conf.		
Expone fraude / encuentra	30.6%	26.5% - 34.7%	91.6%	88.9% - 94.4%	0.0%	100.0%
Opina H / exposición	16.9%	10.9% - 22.9%	86.6%	83.0% - 90.2%	0.0%	100.0%
Opina L / exposición	75.0%	68.0% - 82.0%	8.5%	5.6% - 11.5%	100.0%	0.0%
Indeciso / exposición	8.1%	3.7% - 12.5%	4.8%	2.6% - 7.1%	0.0%	0.0%
Opina H / NO exposición	55.0%	51.6% - 58.4%	14.1%	10.7% - 17.5%	---	0.0%
Opina L / NO exposición	28.3%	25.2% - 31.4%	72.3%	67.9% - 76.6%	---	100.0%
Indeciso / NO exposición	16.7%	14.2% - 19.3%	13.6%	10.3% - 17.0%	---	0.0%

**TABLA 6: Primer Modelo. Sesiones Total**

Variables Dependientes	Estimación 1	Estimación 2	Estimación 3
Sets 2-3 x Total - Negativo	-0.680 (0.215) <sup>***</sup>	-0.716 (0.223) <sup>***</sup>	0.255 (0.107) <sup>**</sup>
Sets 4-5 x Total - Negativo	-0.594 (0.209) <sup>***</sup>	-0.674 (0.205) <sup>***</sup>	0.464 (0.107) <sup>***</sup>
Sets 6-7 x Total - Negativo	-0.552 (0.241) <sup>**</sup>	-1.534 (0.347) <sup>***</sup>	0.651 (0.118) <sup>***</sup>
Sets 2-3 x Total - Positivo	2.062 (0.384) <sup>***</sup>	1.239 (0.174) <sup>***</sup>	-0.986 (0.172) <sup>***</sup>
Sets 4-5 x Total - Positivo	2.492 (0.443) <sup>***</sup>	1.249 (0.165) <sup>***</sup>	-1.051 (0.173) <sup>***</sup>
Sets 6-7 x Total - Positivo	2.520 (0.417) <sup>***</sup>	1.741 (0.228) <sup>***</sup>	-0.938 (0.163) <sup>***</sup>
$\sigma_c$	1.170 (0.180) <sup>***</sup>	---	0.339 (0.075) <sup>***</sup>
$\rho$	0.578 (0.075) <sup>***</sup>	---	0.103 (0.041) <sup>***</sup>
Log - Likelihood	-267.166	-136.587	-494.447
Número de Sujetos	84	67	82
Sets usados en estimación	2 - 7	2 - 7	2 - 7
Número de Observaciones	739	409	876

Estadísticamente diferente a cero al \*\*\* 1% , \*\* 5% , \* 10% de significancia

Variable Dependiente Estimación 1: Exposición (1) o no (0) de fraude cuando es encontrado

Variable Dependiente Estimación 2: Opinión H (1) ó L (0) cuando fraude es expuesto

Variable Dependiente Estimación 3: Opinión H (1) ó L (0) cuando fraude NO es expuesto

**TABLA 7: Segundo Modelo. Sesiones Total**

Variables Dependientes	Estimación 1	Estimación 2	Estimación 3
Sets 2-3 x Total - Negativo	-0.430 (0.275)	-0.716 (0.223)***	-0.157 (0.136)
Sets 4-5 x Total - Negativo	-0.413 (0.244)	-0.674 (0.205)***	0.068 (0.154)
Sets 6-7 x Total - Negativo	-0.126 (0.345)	-1.534 (0.347)***	0.046 (0.162)
Sets 2-3 x Total - Positivo	2.030 (0.367)***	0.344 (0.430)	-1.833 (0.399)***
Sets 4-5 x Total - Positivo	2.462 (0.430)***	0.234 (0.469)	-2.066 (0.482)***
Sets 6-7 x Total - Positivo	2.480 (0.398)***	0.604 (0.533)	-2.073 (0.509)***
H-L / expone(-1) x Total - Neg	0.006 (0.003)**	---	---
H-L / expone(-1) x Total - Pos	---	0.014 (0.006)**	0.013 (0.006)**
H-L / no expone(-1) x Total - Neg	---	---	0.010 (0.004)**
Expone / encuentra(-1) x Total - Neg	---	---	0.008 (0.002)***
$\sigma_c$	1.143 (0.168)***	---	---
$\rho$	0.566 (0.072)***	---	---
Log - Likelihood	-265.161	-134.088	-483.226
Número de Sujetos	84	67	82
Sets usados en estimación	2 - 7	2 - 7	2 - 7
Número de Observaciones	739	409	876

Estadísticamente diferente a cero al \*\*\* 1%, \*\* 5%, \* 10% de significancia

Variable Dependiente Estimación 1: Exposición (1) o no (0) de fraude cuando es encontrado

Variable Dependiente Estimación 2: Opinión H (1) ó L (0) cuando fraude es expuesto

Variable Dependiente Estimación 3: Opinión H (1) ó L (0) cuando fraude NO es expuesto

**TABLA 8: Proporción observada del comportamiento individual en las Sesiones Total**

<i>Comportamiento individual por rol</i>	<i>Total - Negativo (48 casos)</i>	<i>Total - Positivo (36 casos)</i>	<i>Ho: igualdad de proporciones (estadístico <math>\chi^2</math>)</i>
<b>Organización: Si encuentra fraude ....</b>			
<i>Interpretación Negativa : Nunca lo denuncia</i>	35.4%	0.0%	15.98 ***
<i>Interpretación Positiva : Siempre lo denuncia</i>	8.3%	72.2%	36.57 ***
<i>Condicional : Lo denuncia algunas veces, nunca o siempre dependiendo del tipo.</i>	43.8%	27.8%	2.25
<i>Indeciso : Lo denuncia algunas veces sin depender del tipo.</i>	12.5%	0.0%	4.84 **
<i>Ho: La distribución de ambos tratamientos es igual</i>			42.18 ***
<b>Público: Si la Organización denuncia fraude .... / Si la Organización NO denuncia fraude ....</b>			
<i>Interpretación Negativa : Opina L / Opina cualq.</i>	64.6%	2.8%	33.32 ***
<i>Interpretación Positiva : Opina H / Opina L</i>	0.0%	44.4%	26.35 ***
<i>Cuasi-positivo : Opina H / Opina cualquiera</i>	18.8%	38.9%	4.19 **
<i>Indeciso : Opina cualquiera / Opina cualquiera</i>	16.7%	5.6%	2.42
<i>Otro : Opina L / Opina cualquiera</i>	0.0%	8.3%	4.15 **
<i>Ho: La distribución de ambos tratamientos es igual</i>			51.14 ***
<b>Como Organización y Público el mismo individuo.....</b>			
<i>Siempre actúa con Interpretación Negativa</i>	18.8%	0.0%	
<i>Termina la sesión con Int. Neg</i>	6.3%	0.0%	
<i>Siempre actúa con Interpretación Positiva</i>	0.0%	8.3%	
<i>Termina la sesión con Int. Pos</i>	0.0%	16.7%	

La Ho de igualdad es **rechazada** al \* 10% of significancia \*\* 5% of significancia \*\*\* 1% of significancia

**TABLA 9: Proporción media observada para cada acción posible al inicio y final de las Sesiones Nulo, por signo de (PH – PL)**

**TRES PRIMEROS SETS**

Acciones	NEGATIVO		POSITIVO		Interpret. Negativa	Interpret. Positiva
	Prom.	Int. 95% conf.	Prom.	Int. 95% conf.		
Expone fraude / encuentra	56.1%	49.7% - 62.6%	57.5%	51.6% - 63.4%	0.0%	100.0%
Opina H / exposición	53.1%	44.5% - 61.8%	52.2%	44.4% - 60.0%	0.0%	100.0%
Opina L / exposición	38.3%	29.9% - 46.7%	35.0%	27.6% - 42.5%	100.0%	0.0%
Indeciso / exposición	8.6%	3.7% - 13.4%	12.7%	7.5% - 18.0%	0.0%	0.0%
Opina H / NO exposición	49.5%	43.8% - 55.2%	42.1%	37.1% - 47.1%	---	0.0%
Opina L / NO exposición	33.4%	28.1% - 38.8%	34.5%	29.7% - 39.3%	---	100.0%
Indeciso / NO exposición	17.1%	12.8% - 21.3%	23.4%	19.2% - 27.7%	---	0.0%

**TRES ULTIMOS SETS**

Acciones	NEGATIVO		POSITIVO		Interpret. Negativa	Interpret. Positiva
	Prom.	Int. 95% conf.	Prom.	Int. 95% conf.		
Expone fraude / encuentra	42.6%	35.5% - 49.8%	67.5%	61.9% - 73.2%	0.0%	100.0%
Opina H / exposición	39.7%	28.9% - 50.6%	64.2%	57.2% - 71.3%	0.0%	100.0%
Opina L / exposición	50.0%	38.9% - 61.1%	30.7%	24.0% - 37.5%	100.0%	0.0%
Indeciso / exposición	10.3%	3.5% - 17.0%	5.0%	1.8% - 8.2%	0.0%	0.0%
Opina H / NO exposición	54.3%	48.4% - 60.1%	32.3%	27.4% - 37.2%	---	0.0%
Opina L / NO exposición	25.9%	20.8% - 31.0%	45.8%	40.6% - 51.0%	---	100.0%
Indeciso / NO exposición	19.9%	15.2% - 24.5%	21.9%	17.6% - 26.2%	---	0.0%

**TABLA 10: Proporción media observada en las Sesiones Nulo por signo de (PH – PL) y contexto**

<b>1. Proporción de exposición de Fraude</b>				
Sets	NEGATIVO		POSITIVO	
	Significado	Neutral	Significado	Neutral
1	75%	69%	75%	53%
2 - 3	39%	57%	61%	41%
4 - 5	44%	63%	75%	47%
6 - 7	51%	36%	82%	51%
<b>2. Diferencia prop. H vs. L cuando hay exposición</b>				
Sets	NEGATIVO		POSITIVO	
	Significado	Neutral	Significado	Neutral
1	3%	32%	12%	11%
2 - 3	17%	12%	42%	-23%
4 - 5	-25%	9%	38%	-14%
6 - 7	-15%	-33%	51%	0%
<b>3. Diferencia prop. H vs. L cuando NO hay exposición</b>				
Sets	NEGATIVO		POSITIVO	
	Significado	Neutral	Significado	Neutral
1	5%	0%	33%	2%
2 - 3	19%	26%	7%	-3%
4 - 5	35%	29%	-14%	-7%
6 - 7	25%	32%	-20%	-9%

**TABLA 11: Primer Modelo. Sesiones Nulo.**

Variables Dependientes	Estimación 1	Estimación 2	Estimación 3
Set 1	0.772 (0.205)***	0.114 (0.154)	0.330 (0.136)**
Set 1 x Neutral	-0.738 (0.294)**	0.204 (0.254)	-0.313 (0.199)
Sets 2-3 x Nulo - Negativo	0.109 (0.181)	0.191 (0.155)	0.338 (0.097)***
Sets 4-5 x Nulo - Negativo	0.003 (0.173)	-0.068 (0.146)	0.538 (0.103)***
Sets 6-7 x Nulo - Negativo	0.220 (0.235)	-0.243 (0.175)	0.498 (0.106)***
Sets 6-7 x Nulo - Negativo x Neut.	-0.781 (0.304)***	----	----
Sets 2-3 x Nulo - Positivo	0.616 (0.243)**	0.660 (0.183)***	0.043 (0.088)
Sets 4-5 x Nulo - Positivo	0.856 (0.223)***	0.549 (0.156)***	-0.180 (0.091)**
Sets 6-7 x Nulo - Positivo	1.086 (0.236)***	0.746 (0.156)***	-0.245 (0.094)***
Sets 2-3 x Nulo - Positivo x Neut.	-0.842 (0.337)***	-0.991 (0.307)***	----
Sets 4-5 x Nulo - Positivo x Neut.	-0.691 (0.317)**	-0.729 (0.264)***	----
Sets 6-7 x Nulo - Positivo x Neut.	-1.011 (0.334)***	-0.746 (0.261)***	----
$\sigma_c$	0.890 (0.095)***	---	---
$\rho$	0.442 (0.052)***	---	---
Log - Likelihood	-642.308	-386.896	-810.167
Número de Sujetos	108	102	108
Sets usados en estimación	1 - 7	1 - 7	1 - 7
Número de Observaciones	1135	602	1227

Estadísticamente diferente a cero al \*\*\* 1%, \*\* 5%, \* 10% de significancia

Variable Dependiente Estimación 1: Exposición (1) o no (0) de fraude cuando es encontrado

Variable Dependiente Estimación 2: Opinión H (1) ó L (0) cuando fraude es expuesto

Variable Dependiente Estimación 3: Opinión H (1) ó L (0) cuando fraude NO es expuesto



**TABLA 12: Segundo Modelo. Sesiones Nulo.**

Variables Dependientes	Estimación 1	Estimación 2	Estimación 3
Sets 2-3 x Nulo - Negativo	0.153 (0.201)	-0.676 (0.388)*	0.285 (0.099)***
Sets 4-5 x Nulo - Negativo	0.031 (0.191)	-0.789 (0.332)**	0.397 (0.113)***
Sets 6-7 x Nulo - Negativo	0.301 (0.271)	-0.841 (0.304)***	0.455 (0.107)***
Sets 6-7 x Nulo - Negativo x Neut.	-0.729 (0.322)**	----	----
Sets 2-3 x Nulo - Positivo	0.753 (0.226)***	0.660 (0.183)***	0.043 (0.088)
Sets 4-5 x Nulo - Positivo	0.951 (0.228)***	0.549 (0.156)***	-0.180 (0.092)**
Sets 6-7 x Nulo - Positivo	1.220 (0.234)***	0.746 (0.156)***	-0.245 (0.094)***
Sets 2-3 x Nulo - Positivo x Neut.	-0.812 (0.354)**	-0.991 (0.306)***	----
Sets 4-5 x Nulo - Positivo x Neut.	-0.672 (0.339)**	-0.729 (0.264)***	----
Sets 6-7 x Nulo - Positivo x Neut.	-1.154 (0.342)***	-0.746 (0.261)***	----
HLa_1 x None Neg x Mean	0.008 (0.003)***	----	----
HLa_1 x None Pos x Neu	0.009 (0.004)*	----	----
Ann_1 x None Neg	----	0.014 (0.006)**	----
HLn_1 x None Neg x Mean	----	----	0.008 (0.003)***
$\sigma_c$	0.949 (0.106)***	---	---
$\rho$	0.474 (0.056)***	---	---
Log - Likelihood	-545.525	-311.936	-696.263
Número de Sujetos	108	100	107
Sets usados en estimación	2 - 7	2 - 7	2 - 7
Número de Observaciones	966	496	1065

Estadísticamente diferente a cero al \*\*\* 1% , \*\* 5% , \* 10% de significancia

Variable Dependiente Estimación 1: Exposición (1) o no (0) de fraude cuando es encontrado

Variable Dependiente Estimación 2: Opinión H (1) ó L (0) cuando fraude es expuesto

Variable Dependiente Estimación 3: Opinión H (1) ó L (0) cuando fraude NO es expuesto

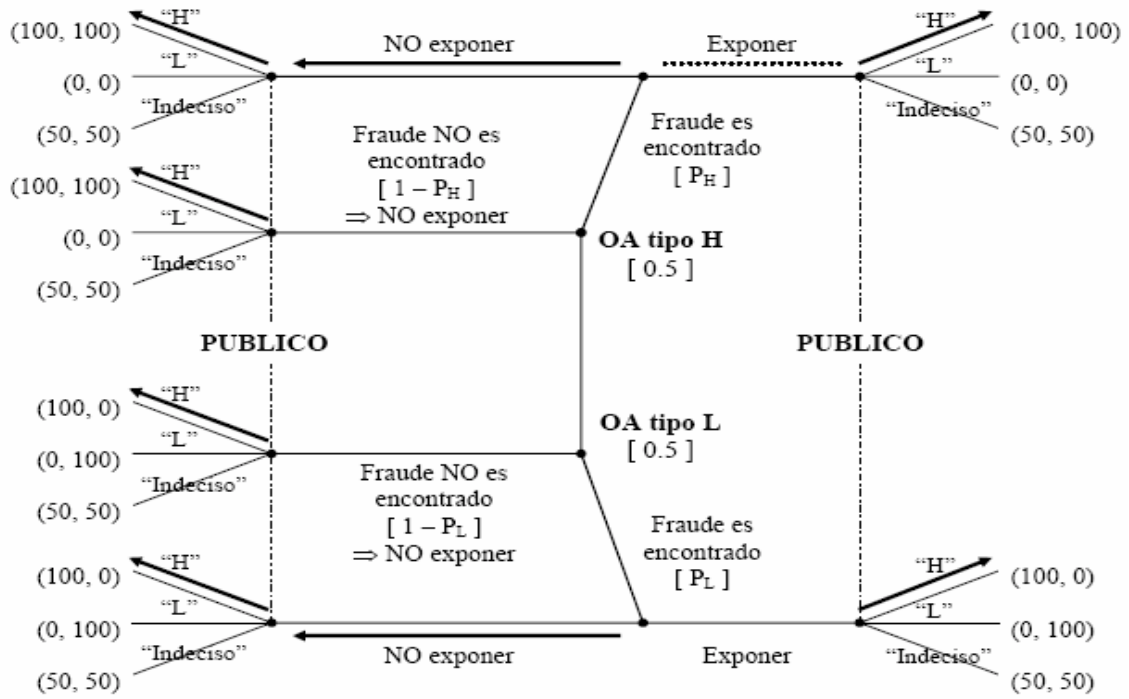
**TABLA 13: Proporción observada del comportamiento individual en las Sesiones Nulo**

<i>Comportamiento individual por rol</i>	<i>Nulo - Neg (48 casos)</i>	<i>Nulo - Pos - Sig (36 casos)</i>	<i>Nulo - Pos - Neu (24 casos)</i>	<i>Nulo Pos Sig vs Nulo Pos Neu (estadístico <math>\chi^2</math>)</i>	<i>Nulo Neg vs Nulo Pos Sig (estadístico <math>\chi^2</math>)</i>
<b>Organización: Si encuentra fraude ....</b>					
<i>Interpretación Negativa: Nunca lo denuncia</i>	14.6%	11.1%	4.2%	0.91	0.22
<i>Interpretación Positiva: Siempre lo denuncia</i>	18.8%	33.3%	8.3%	5.03 **	2.33
<i>Condicional: Lo denuncia algunas veces, nunca o siempre dependiendo del tipo.</i>	52.1%	44.4%	79.2%	7.14 ***	0.48
<i>Indeciso: Lo denuncia algunas veces sin depender del tipo.</i>	14.6%	11.1%	8.3%	0.12	0.22
<i>Ho: La distribución de ambos tratamientos es igual</i>				7.78 *	2.37
<b>Público: Si la Organización denuncia fraude .... / Si la Organización NO denuncia fraude ....</b>					
<i>Interpretación Negativa: Opina L / Opina cualq.</i>	47.9%	8.3%	37.5%	7.65 ***	15.08 ***
<i>Interpretación Positiva: Opina H / Opina L</i>	0.0%	13.9%	8.3%	0.43	7.08 ***
<i>Cuasi-positivo: Opina H / Opina cualquiera</i>	27.1%	47.2%	33.3%	1.14	5.47 **
<i>Indeciso: Opina cualquiera / Opina cualquiera</i>	22.9%	30.6%	20.8%	0.70	0.12
<i>Otro: Opina L / Opina cualquiera</i>	2.1%	0.0%	0.0%	n.a.	0.76
<i>Ho: La distribución de ambos tratamientos es igual</i>				7.68 *	21.56 ***
<b>Como Organización y Público el mismo individuo.....</b>					
<i>Siempre actúa con Interpretación Negativa</i>	2.1%	0.0%	0.0%		
<i>Termina la sesión con Int. Neg</i>	8.3%	5.6%	4.2%		
<i>Siempre actúa con Interpretación Positiva</i>	0.0%	2.8%	0.0%		
<i>Termina la sesión con Int. Pos</i>	0.0%	5.6%	0.0%		

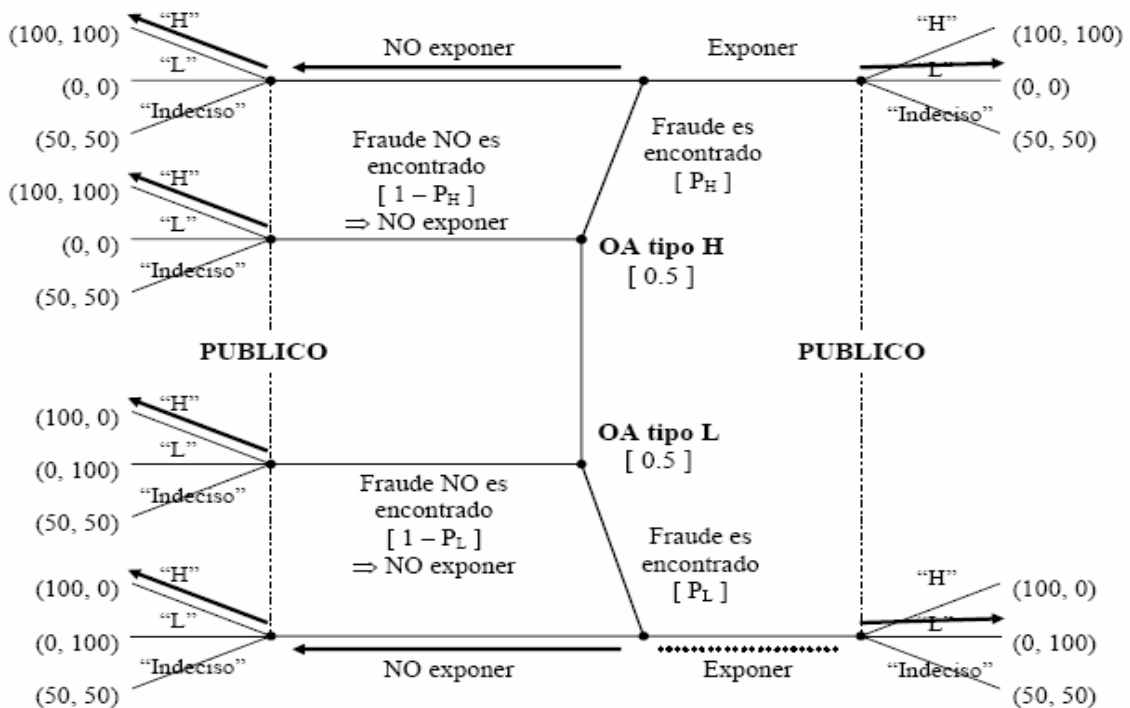
La Ho de igualdad es **rechazada** al \* 10% of significancia \*\* 5% of significancia \*\*\* 1% of significancia

**FIGURA 1 (a y b): Equilibrios de No exposición**

Entre corchetes están las probabilidades de la naturaleza [0.5] y las probabilidades de Hallazgo de fraude para cada tipo [ $P_i$ ,  $i = H, L$ ]. Las flechas gruesas indican las acciones de equilibrio para ambos jugadores. Las flechas punteadas indican las creencias fuera del equilibrio que sustentan la acción del Público. Solo el equilibrio en la figura 1b es consistente con una MALA Interpretación.



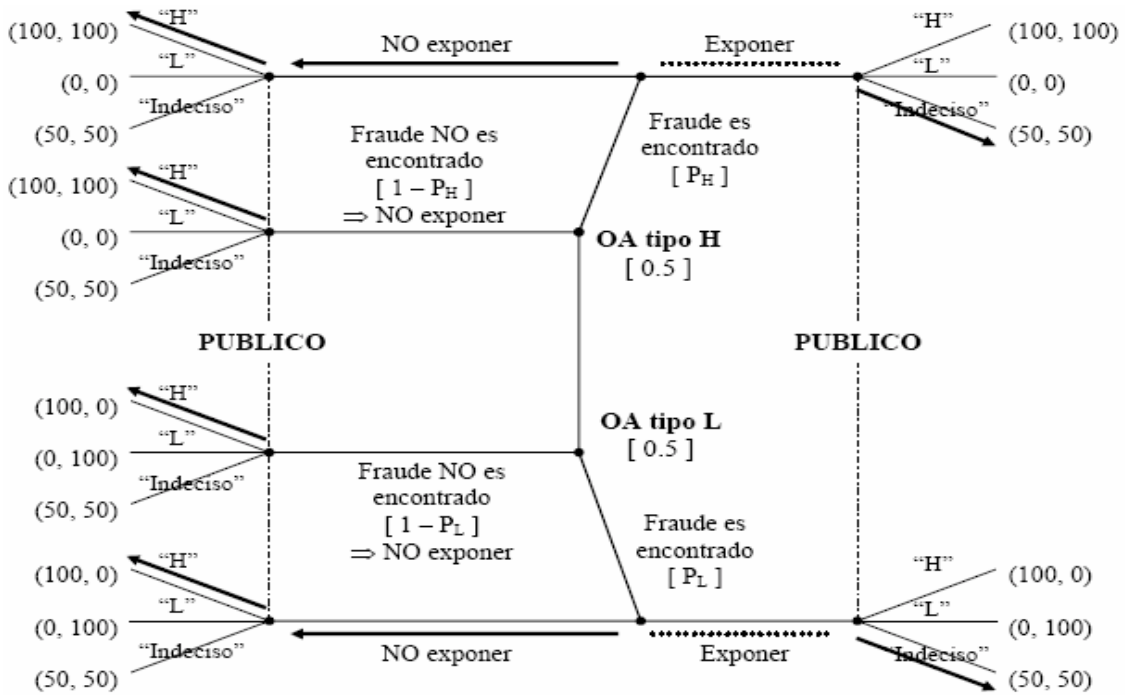
(a)



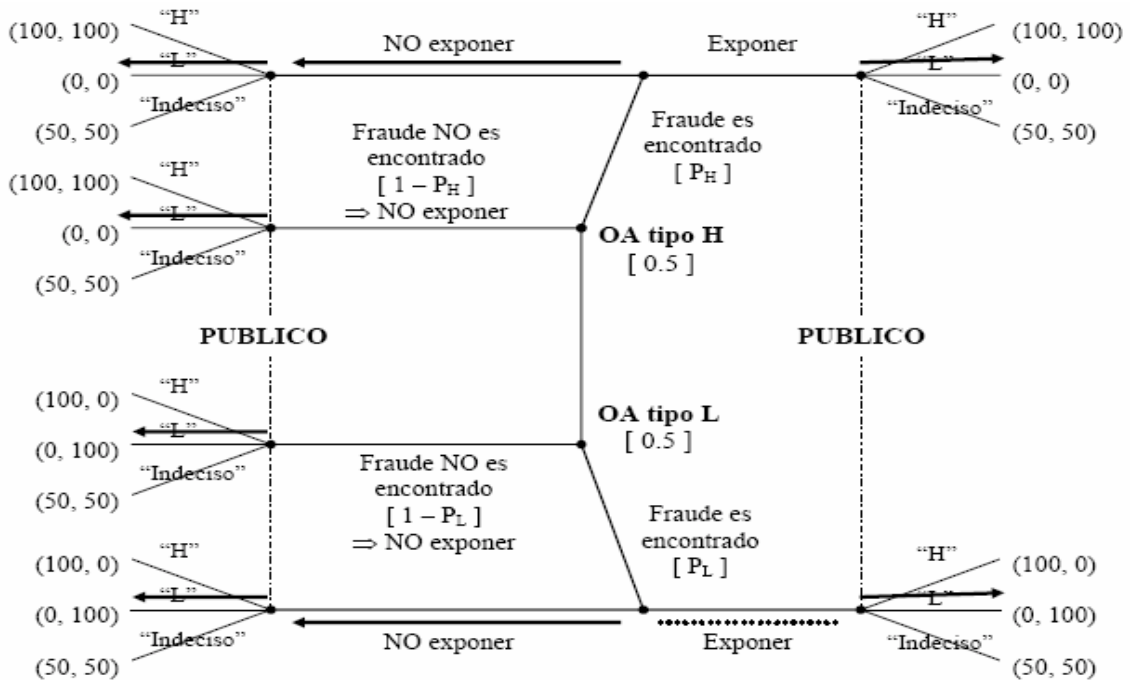
(b)

**FIGURA 1 (c y d): Equilibrios de No exposición**

Entre corchetes están las probabilidades de la naturaleza [0.5] y las probabilidades de Hallazgo de fraude para cada tipo [ $P_i$ ,  $i = H, L$ ]. Las flechas gruesas indican las acciones de equilibrio para ambos jugadores. Las flechas punteadas indican las creencias fuera del equilibrio que sustentan la acción del Público. Solo el equilibrio en la figura 1d es consistente con una MALA Interpretación.



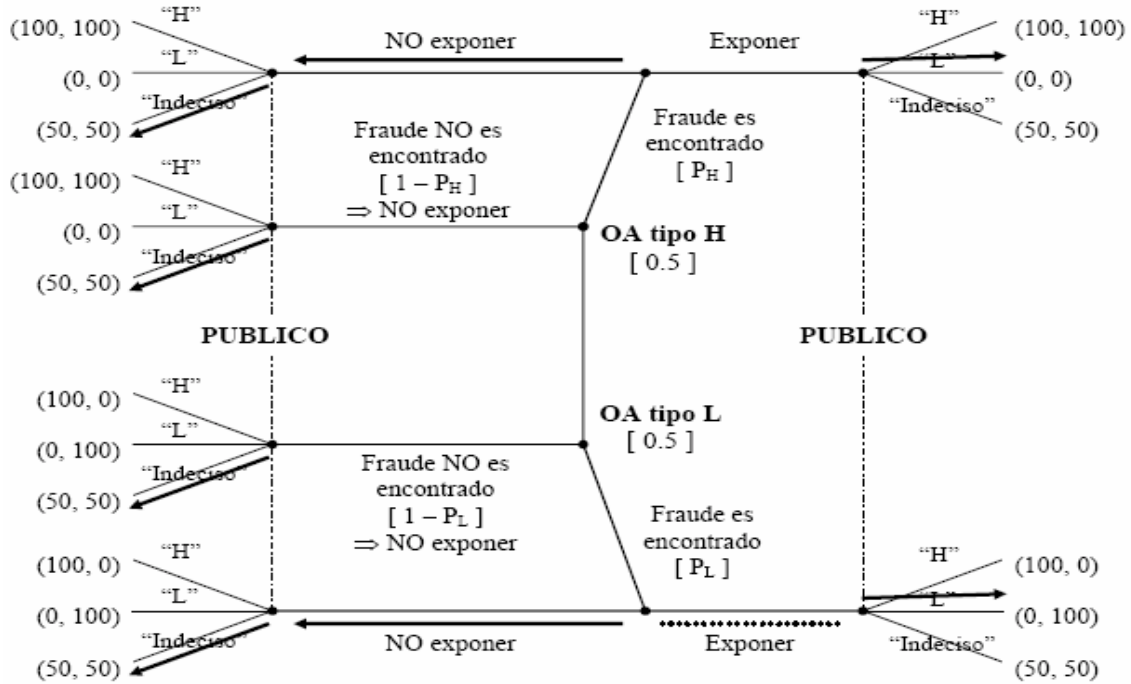
(c)



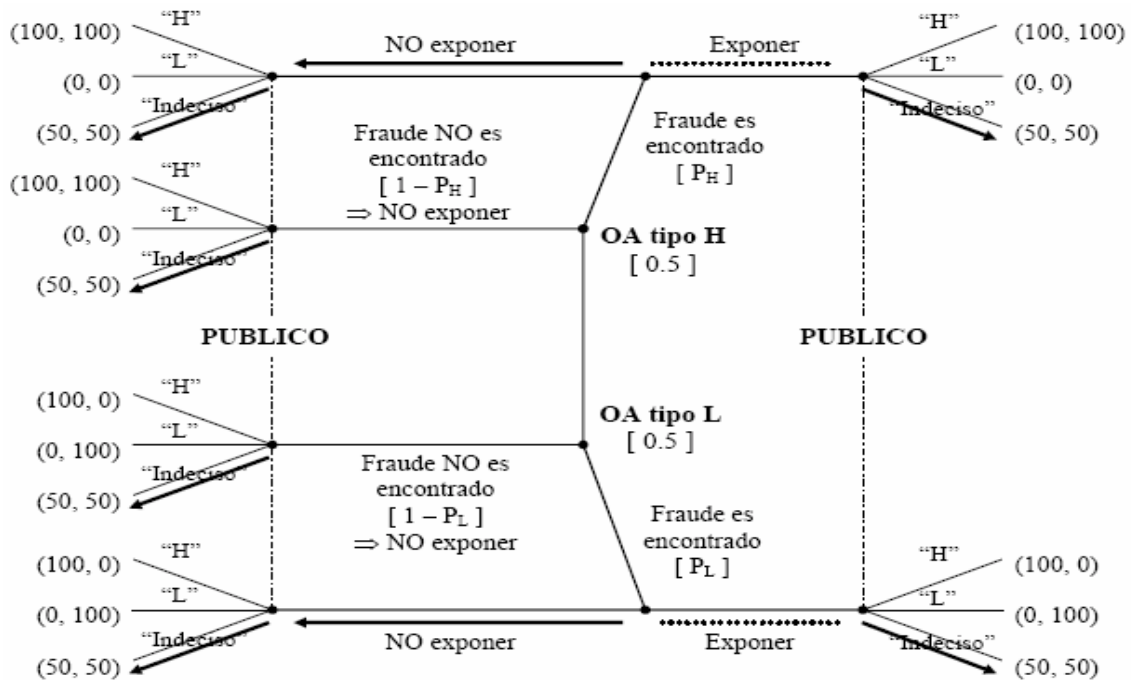
(d)

**FIGURA 1 (e y f): Equilibrios de No exposición**

Entre corchetes están las probabilidades de la naturaleza [0.5] y las probabilidades de Hallazgo de fraude para cada tipo [ $P_i$ ,  $i = H, L$ ]. Las flechas gruesas indican las acciones de equilibrio para ambos jugadores. Las flechas punteadas indican las creencias fuera del equilibrio que sustentan la acción del Público. Solo el equilibrio en la figuras 1e es consistente con una MALA Interpretación.



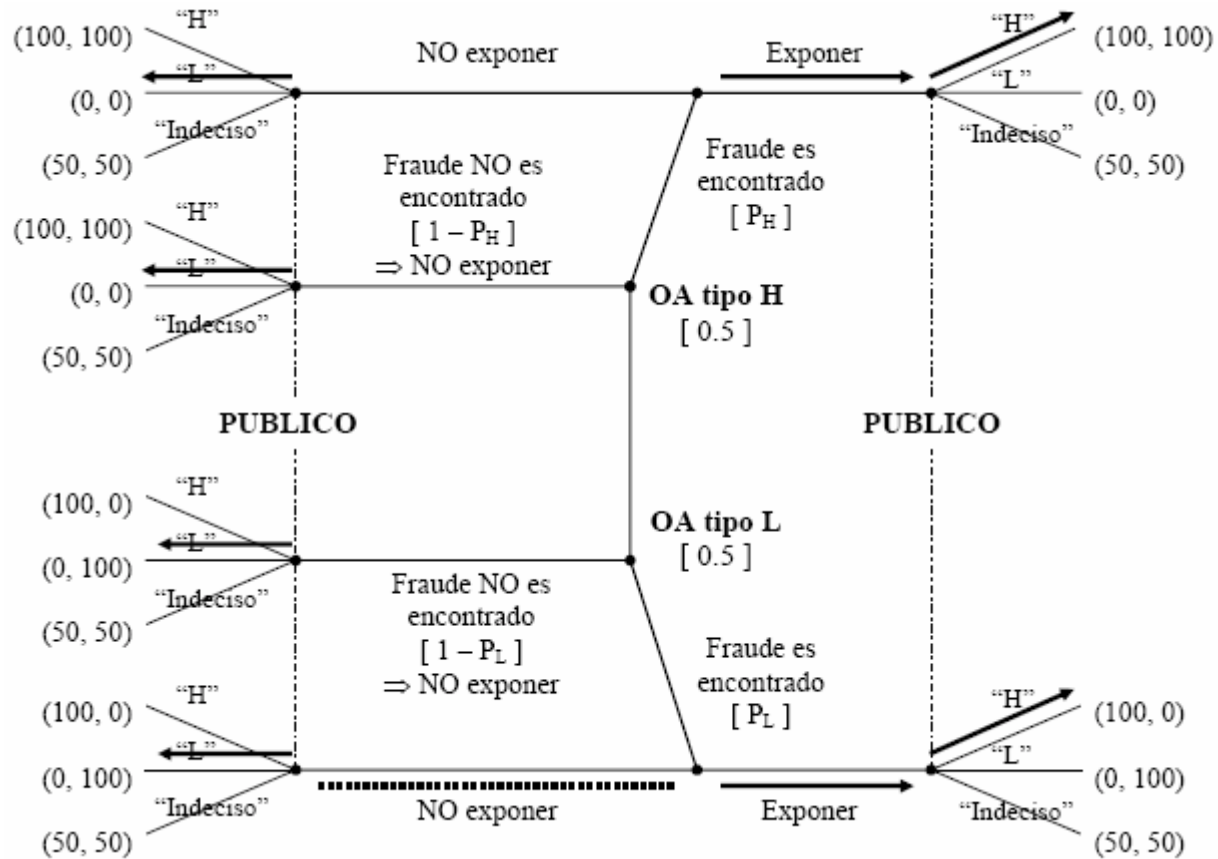
(e)



(f)

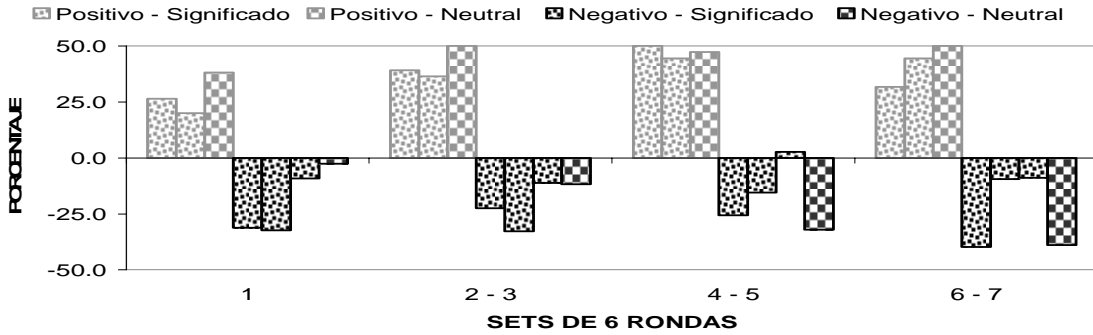
**FIGURA 2: Equilibrio de Exposición**

Entre corchetes están las probabilidades de la naturaleza [0.5] y las probabilidades de hallazgo de fraude para cada tipo [ $P_i$ ,  $i = H, L$ ]. Las flechas gruesas indican las acciones de equilibrio para ambos jugadores. Las flechas punteadas indican las creencias fuera del equilibrio que sustentan la acción del Público.

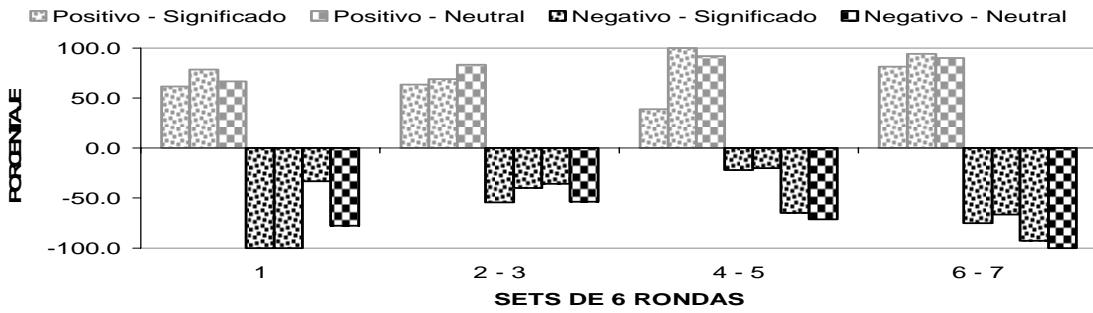


### FIGURA 3: JUEGO PROMEDIO OBSERVADO EN SESIONES TOTAL

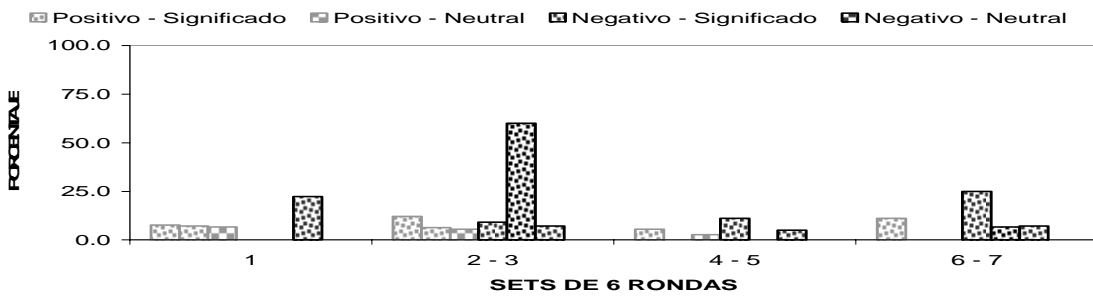
OA: PROPORCION DE EXPOSICIÓN DE FRAUDE MENOS 50 %



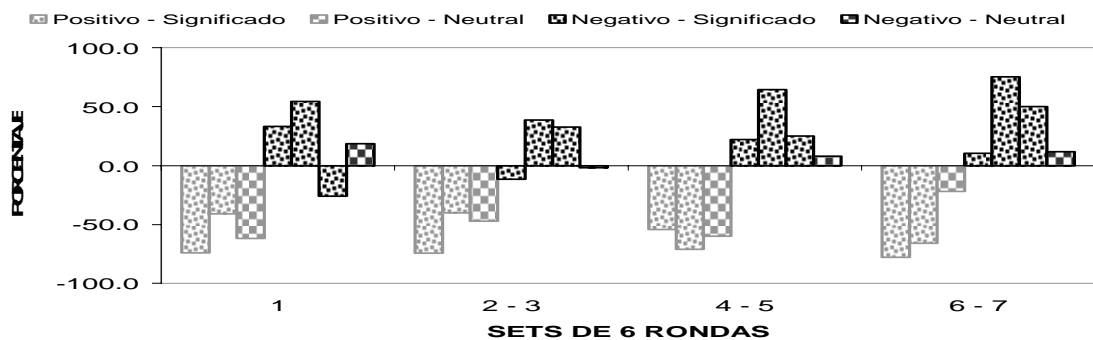
PUBLICO: PROP. DE OPINIONES "H" MENOS PROP. DE OPINIONES "L" CUANDO OA EXPONE FRAUDE



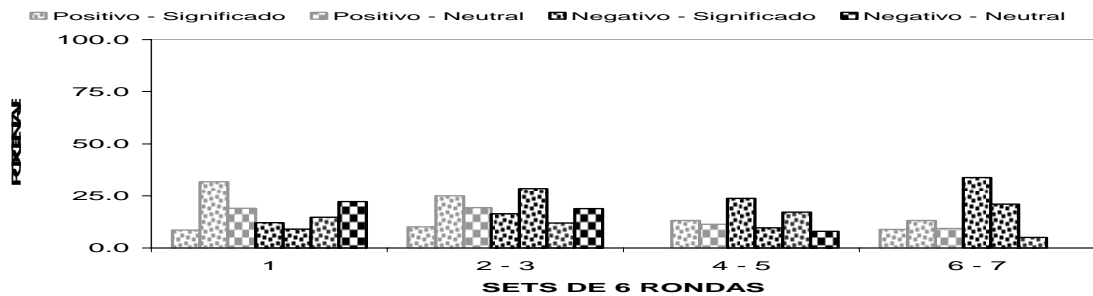
PUBLICO: PROP. INDECISOS CUANDO OA EXPONE FRAUDE



PUBLICO: PROP. DE OPINIONES "H" MENOS PROP. DE OPINIONES "L" CUANDO LA OA NO EXPONE FRAUDE



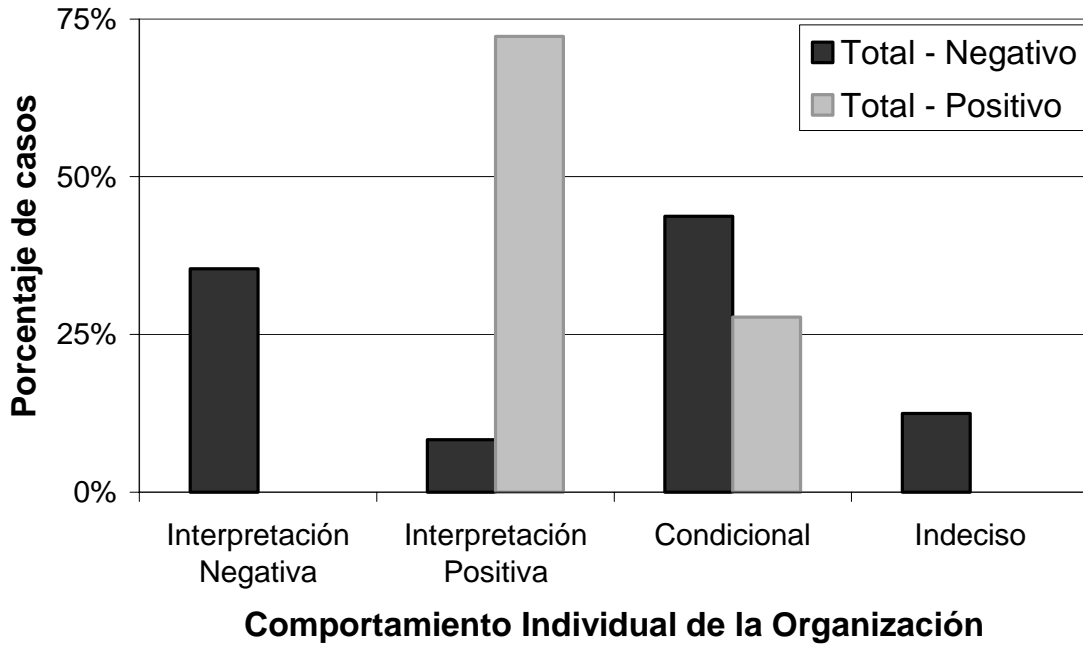
PUBLICO: PROP. INDECISOS CUANDO OA NO EXPONE FRAUDE



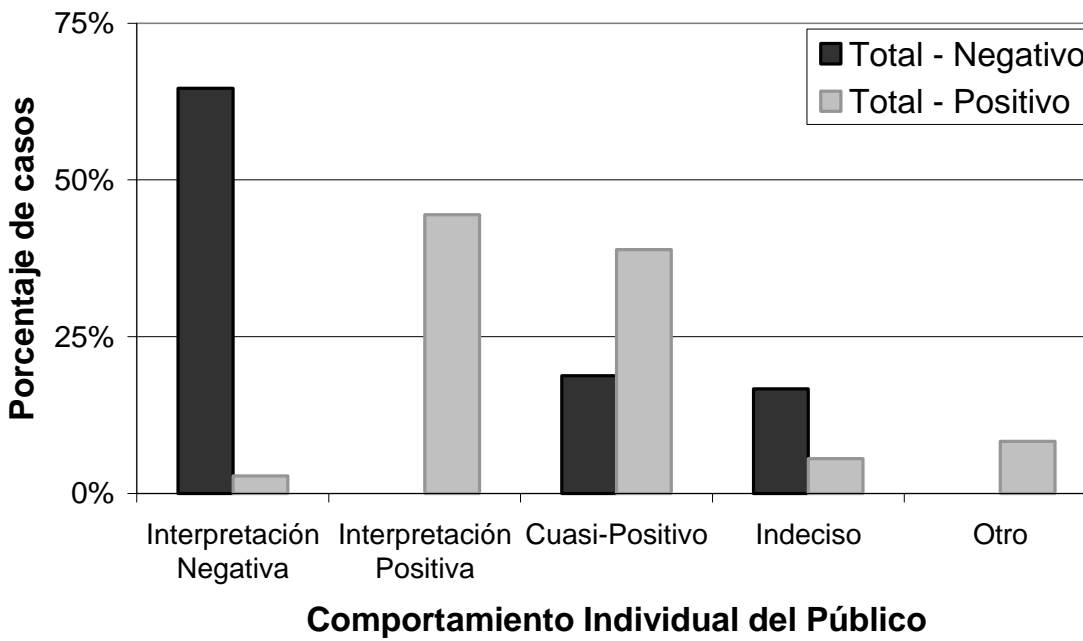
**FIGURA 3: JUEGO PROMEDIO OBSERVADO EN SESIONES TOTAL**

**FIGURA 4:**

**Comportamiento de las Organizaciones  
en Sesiones TOTAL**



**Comportamiento del Público en  
Sesiones TOTAL**





**FIGURA 3: JUEGO PROMEDIO OBSERVADO EN SESIONES TOTAL**

**FIGURA 5:**

