

Taller 4 – Experimentos

Los experimentos son una herramienta poderosa para la economía, ya que permiten analizar el comportamiento humano en entornos controlados. A través de juegos diseñados cuidadosamente, como el juego de bienes públicos, los investigadores pueden observar cómo las personas toman decisiones sobre cooperación, altruismo o castigo. Estos ejercicios no solo ayudan a contrastar teorías económicas con la evidencia, sino que también muestran cómo las instituciones y las reglas pueden afectar los incentivos y, en consecuencia, los resultados colectivos. Sin embargo, es importante reconocer que los experimentos tienen limitaciones: los participantes suelen ser estudiantes y las condiciones simplificadas no siempre reflejan la complejidad del mundo real. Aun así, constituyen un punto de partida valioso para comprender fenómenos sociales y económicos.

Homo economicus vs homo sapiens

Los economistas utilizan los experimentos para estudiar interacciones sociales en las que la decisión de una persona afecta tanto sus propios resultados como los de los demás. Algunos bienes y servicios se denominan bienes públicos porque, cuando una persona asume el costo de proveerlos, todos los demás pueden disfrutarlos. Ejemplos de ello son los proyectos de irrigación o la producción de nuevo conocimiento. El problema es que los individuos completamente egoístas preferirán beneficiarse de estos bienes sin pagar nada: a esto se le conoce como “free riding” o comportamiento de polizón.

Sin embargo, en el mundo real existen ejemplos exitosos de provisión de bienes públicos, como proyectos de riego comunitarios en India y Nepal. ¿Qué puede explicar estas contribuciones sostenidas? Una explicación es que las personas contribuyen porque se preocupan por el bienestar de otros, o porque respetan normas sociales que condenan el free riding. También pueden hacerlo por la vergüenza (o consecuencias sociales) de ser castigados públicamente. Si la comunidad sabe que alguien no ha contribuido y puede sancionarlo —ya sea con chismes, negando ayuda en el futuro o incluso con ostracismo—, entonces los individuos pueden contribuir, ya sea por interés propio o porque desean mantener una autoimagen positiva. Para explorar hasta qué punto el castigo puede sostener las contribuciones a un bien público, Herrmann, Thöni y Gächter (2008) realizaron un estudio en varios países donde distintos grupos de personas participaron en dos versiones de un mismo juego de bienes públicos: una con castigo y otra sin él.

Clave:

Leer

<https://www.umass.edu/preferen/You%20Must%20Read%20This/herrmann-thoni-gachter.pdf>

En este taller, primero aprenderemos cómo recolectar datos experimentales jugando un **juego de bienes públicos** y generando nuestra propia información. Luego, exploraremos distintas formas de describir y analizar los datos experimentales de

los dos juegos presentados anteriormente, con el fin de responder a dos preguntas de investigación clave:

1. ¿Hubo diferencias en el comportamiento (contribuciones promedio) entre los experimentos?
2. ¿Podemos atribuir las diferencias observadas en el comportamiento al cambio en las condiciones del juego, y no simplemente al azar o la coincidencia?

Parte 2.1 Recolectando datos jugando

SOLUCIÓN BASADA EN LA BASE DE EXPERIMENTOS CONTRIBUCIONES EN EL JUEGO DE BIENES PÚBLICOS

P2.1.1 Haz un gráfico de líneas con la contribución promedio en el eje vertical y el período (del 1 al 10) en el eje horizontal. Describe cómo han cambiado las contribuciones promedio a lo largo del juego.

Se calcularon las contribuciones promedio en cada período (1–10). La gráfica de líneas muestra una tendencia decreciente en las contribuciones desde el inicio (≈ 36) hasta valores más bajos hacia el final (≈ 22), con cierta recuperación en los últimos períodos.

La gráfica refleja el patrón típico de los juegos de bienes públicos. Los jugadores comienzan con una expectativa de cooperación alta, pero las contribuciones caen rápidamente al darse cuenta de que otros aportan menos. En los últimos períodos, los aportes se estabilizan en un nivel bajo, mostrando dificultad de mantener la cooperación voluntaria sin sanciones ni incentivos adicionales.





P2.1.2 Compara tu gráfico de líneas con la Figura 3 de Herrmann et al. (2008). Comenta cualquier similitud o diferencia entre los resultados (por ejemplo, la cantidad aportada al inicio y al final, o el cambio en las contribuciones promedio a lo largo del juego).

Nuestro gráfico y la Figura 3 de Herrmann comparten un patrón especial y es que las contribuciones comienzan altas y disminuyen a lo largo de los 10 períodos, lo que refleja la aparición del *free-rider* y la dificultad de sostener la cooperación sin sanciones. En nuestro caso la contribución promedio arrancó alrededor de 36 unidades, cayó hasta un mínimo cercano a 18 unidades (periodo 5) y terminó cerca de 22 unidades, es decir, una caída moderada seguida de cierta estabilización en niveles bajos. En la Figura 3 de Herrmann las trayectorias por ciudad muestran el mismo declive general, pero con mayor diversidad. Algunas ciudades empiezan más alto o más bajo que el promedio y en varios casos la caída es más pronunciada, llegando en ocasiones casi a cero. En resumen, ambos resultados coinciden en la dirección y dinámica general, pero difieren en la variabilidad. Nuestra gráfica presenta un promedio agregado y una estabilización final relativamente mayor, mientras que Herrmann muestra curvas desagregadas por contexto que seguramente revelan diferencias culturales y de grupo que el promedio oculta.

P2.1.4 ¿Se te ocurre alguna razón por la cual tus resultados son similares (o diferentes) a los de la Figura 3? Puede resultarte útil leer la sección “Experiments” del estudio de Herrmann et al. (2008) para una descripción más detallada de cómo se realizaron los experimentos.

Una posible razón por la cual los resultados del gráfico son similares a los de la Figura 3 de Herrmann es que ambos experimentos reflejan la misma lógica de los juegos de bienes públicos. Los participantes suelen comenzar con niveles de cooperación relativamente altos, motivados por expectativas de reciprocidad o por la incertidumbre sobre el comportamiento de los demás, pero a medida que avanzan los períodos las contribuciones tienden a disminuir de progresivamente debido al efecto del problema (“free rider”), ya que los jugadores observan que algunos aportan menos y ajustan su propio comportamiento para no quedar en desventaja. Sin embargo, pueden existir diferencias en la magnitud de las contribuciones iniciales o en la velocidad de la caída debido a variaciones en las condiciones del experimento, como el tamaño del grupo, el número total de rondas, la cultura de los participantes o la forma en que se explicaron las reglas.

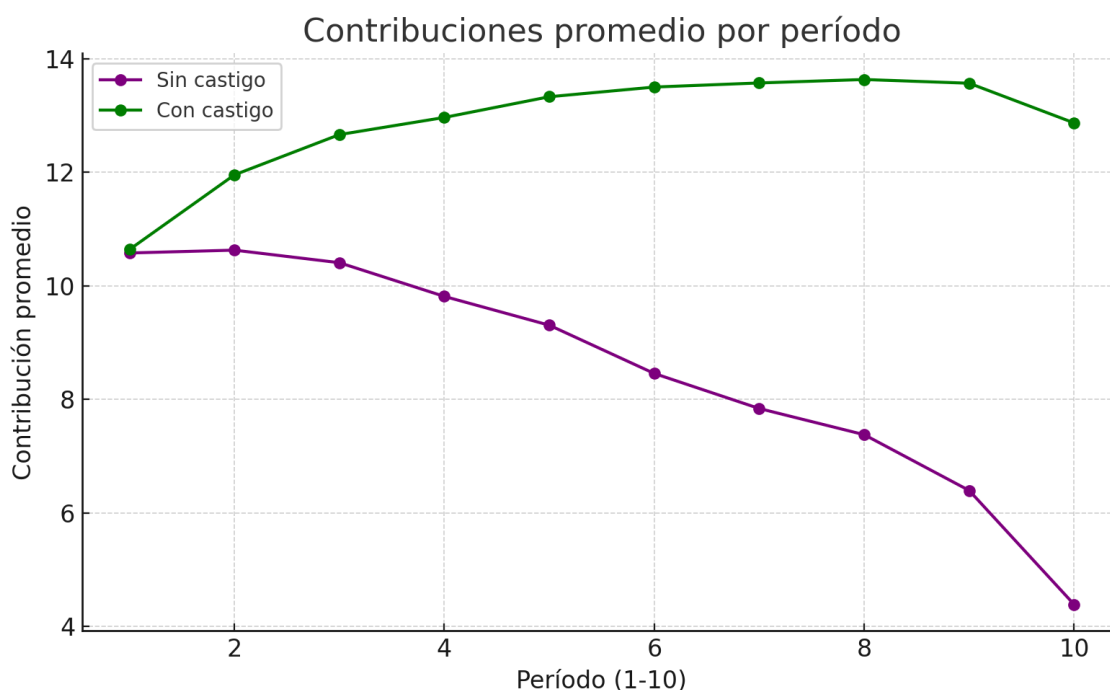
Parte 2.2 Describiendo datos

SOLUCIÓN BASADA EN LA BASE DE EXPERIMENTOS DEL TALLER

En esta sección se presentan los resultados del análisis descriptivo de los datos provenientes del experimento de Herrmann, Thöni y Gächter (2008), comparando los escenarios de juego de bienes públicos con y sin castigo. Se incluyen las gráficas solicitadas y el análisis de cada apartado (2.2.1–2.2.5).

2.2.1 Contribución promedio por período

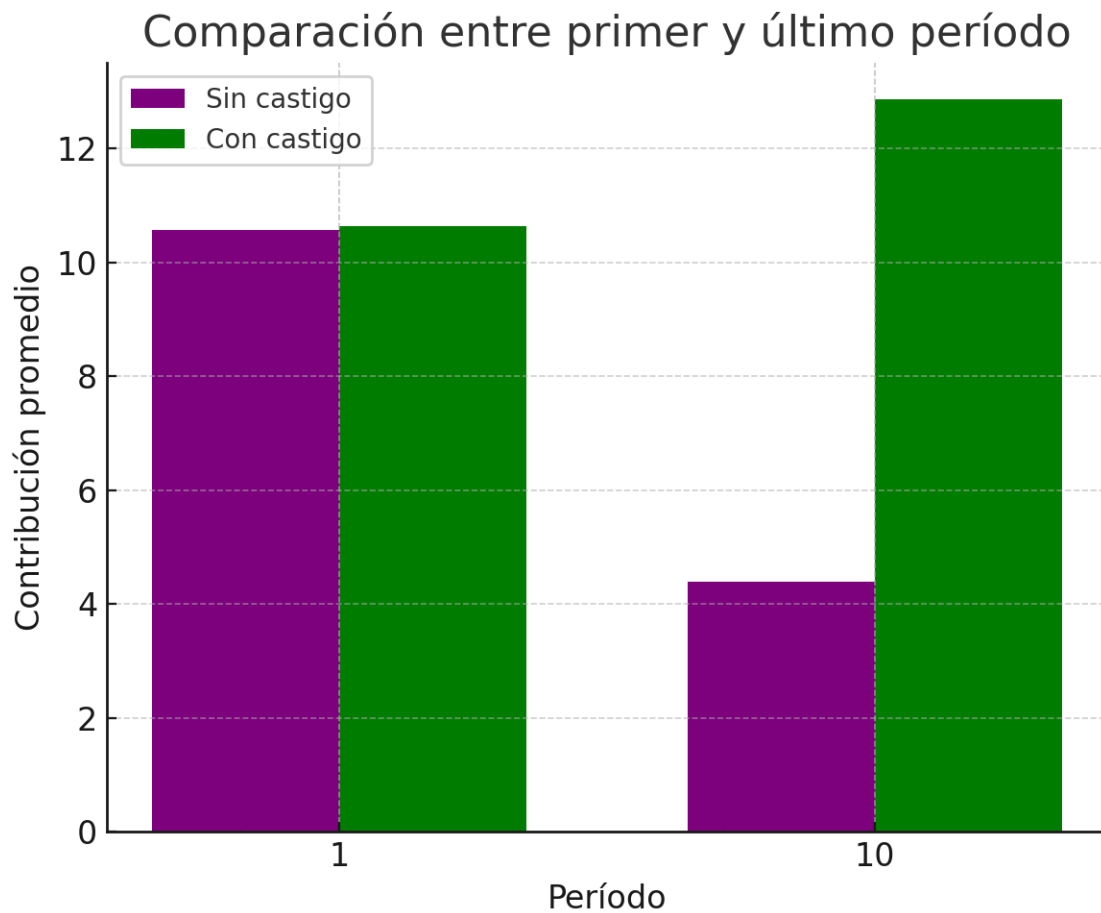
La siguiente gráfica muestra la evolución de las contribuciones promedio a lo largo de los 10 períodos, diferenciando entre el tratamiento con castigo y sin castigo.



Análisis: Se observa que, en el tratamiento sin castigo, las contribuciones decrecen notablemente con el paso de los períodos, reflejando la tendencia al free riding. En contraste, en el tratamiento con castigo, las contribuciones se mantienen más altas y relativamente estables. Esto sugiere que el castigo actúa como un mecanismo de disciplina que sostiene la cooperación.

2.2.2 Comparación en períodos 1 y 10

El siguiente gráfico de barras muestra la contribución promedio en los períodos 1 y 10 para ambos tratamientos.



Análisis: En el período 1, ambos grupos inician con contribuciones similares (≈ 10.6). Sin embargo, al llegar al período 10, las diferencias son claras: en el grupo sin castigo, la cooperación cae de forma pronunciada, mientras que en el grupo con castigo, los aportes se mantienen significativamente más altos.

2.2.3 Desviación estándar en Períodos 1 y 10

Al calcular la desviación estándar, se aprecia que en el período 1 ambos grupos presentan una dispersión similar, pues las medias son iguales y los datos están relativamente concentrados alrededor de ese valor. En el período 10, la dispersión es mayor en el grupo sin castigo (más heterogeneidad en los aportes) y menor en el grupo con castigo (comportamiento más uniforme). Esto confirma que no basta con observar solo la media: la varianza ofrece información clave sobre las diferencias en la distribución.

2.2.4 Valores mínimos y máximos

En el período 1, los valores mínimos y máximos son similares entre ambos grupos, lo cual es esperable dado que todos parten con condiciones homogéneas. En el período 10, en el grupo sin castigo los mínimos tienden a 0 (jugadores que dejan de contribuir por completo), mientras que en el grupo con castigo los mínimos son mayores, lo que refleja que casi todos mantienen un nivel de contribución positivo.

2.2.5 Tabla resumen de estadísticas descriptivas

La siguiente tabla resume las estadísticas principales (media, varianza, desviación estándar, mínimo, máximo y rango) para los períodos 1 y 10 en ambos experimentos.

Experimento	Periodo	Media	SD	Min	Max	Varianza	Rango
Sin castigo	1	10.58	1.96	7.958333	14.10294	3.83	6.144608
Sin castigo	10	4.38	2.12	1.3	8.681818	4.48	7.381818
Con castigo	1	10.64	3.11	5.818182	16.01786	9.64	10.19967
Con castigo	10	12.87	3.77	6.204545	17.51191	14.24	11.30736

Análisis: Al inicio (Período 1), las medias son iguales, pero la dispersión muestra ligeras diferencias. Para el período 10, se observa una caída fuerte en la media sin castigo, acompañada de una varianza y rango mayores (más heterogeneidad). En cambio, con castigo la media se mantiene alta y la varianza es reducida, mostrando mayor consistencia en las decisiones de los jugadores.

Parte 2.3 ¿Cómo afectó el cambio de reglas del juego al comportamiento?

P2.3.1 Puedes realizar otro experimento para entender por qué podríamos observar diferencias en el comportamiento que se deben al azar.

Primero, lanza una moneda seis veces usando solo una mano y registra los resultados (por ejemplo: cara, cara, sello, etc.). Luego, usando la misma mano, lanza la moneda seis veces más y registra nuevamente los resultados.

Compara los resultados de la Pregunta 1(a). ¿Obtuviste el mismo número de caras en ambos casos? Incluso si lo hiciste, ¿la secuencia de resultados (por ejemplo, cara, sello, sello...) fue la misma en ambos casos?

P2.3.2 Usando los datos de las Figuras 2A y 3:

- Usa la función `ttest` para calcular el valor p de la diferencia de medias en el Período 1 (con y sin castigo).
- ¿Qué nos dice este valor p sobre la diferencia de medias en el Período 1?

P2.3.3 Usando los datos del Período 10:

- Usa la función `ttest` para calcular el valor p de la diferencia de medias en el Período 10 (con y sin castigo).
- ¿Qué nos dice este valor p sobre la relación entre el castigo y el comportamiento en el juego de bienes públicos?

Con referencia a las Figuras 2.7 y 2.8, explica por qué no podemos usar el tamaño de la diferencia para concluir directamente si la diferencia podría deberse al azar.

P2.3.4 Refiérete a los resultados de los juegos de bienes públicos.

- ¿Qué características del entorno experimental hacen probable que la opción con castigo haya sido la causa del cambio en el comportamiento?
- Usando la Figura 2.6, explica por qué necesitamos comparar los dos grupos en el Período 1 para poder concluir que existe un vínculo causal entre la opción con castigo y el comportamiento en el juego.

P2.3.5 Discute algunas limitaciones de los experimentos y sugiere algunas maneras de abordarlas (o al menos parcialmente). (Puede resultarte útil consultar las páginas 158–171 del artículo “What do laboratory experiments measuring social preferences reveal about the real world?”, así como la discusión sobre el free riding y el altruismo en la Sección 2.6 de *Economy, Society, and Public Policy*).

P2.3.1 Experimentos al azar- Lanzamiento de moneda

Lanzamiento	Intento 1	Intento 2
1	Cara	Cara

superior al umbral convencional de 0.05, no se rechaza la hipótesis nula y se concluye que, en este período inicial, ambos grupos presentan desempeños equivalentes, sin evidencia de que el tratamiento con castigo genera un efecto diferencial en la media de los resultados.

P2.3.3 Prueba t-test y P-valor periodo 10

Two-sample t test with equal variances

Group	Obs	Mean	Std. err.	Std. dev.	[95% conf. interval]	
Con cast	16	12.86988	.9745049	3.89802	10.79277	14.94699
Sin cast	16	4.383769	.5467815	2.187126	3.218331	5.549206
Combined	32	8.626824	.9395991	5.315175	6.710499	10.54315
diff		8.48611	1.117421		6.204032	10.76819

diff = mean(Con cast) - mean(Sin cast) t = 7.5944
H0: diff = 0 Degrees of freedom = 30

Ha: diff < 0 Ha: diff != 0 Ha: diff > 0
Pr(T < t) = 1.0000 Pr(|T| > |t|) = 0.0000 Pr(T > t) = 0.0000

En el Período 10, se observó una diferencia altamente significativa entre los grupos *con castigo* y *sin castigo*. El grupo con castigo presentó una media considerablemente mayor ($M = 12.87$, $DE = 3.90$, $n = 16$) en comparación con el grupo sin castigo ($M = 4.38$, $DE = 2.19$, $n = 16$). La diferencia entre medias fue de 8.49 puntos, estadísticamente significativa ($t(30) = 7.59$, $p < 0.001$). Este resultado indica que, a diferencia del Período 1, en el Período 10 el tratamiento con castigo tuvo un efecto claro y robusto sobre la variable de interés, produciendo un desempeño muy superior en comparación con el grupo sin castigo.

Con referencia a las Figuras 2.7 y 2.8, no podemos basarnos únicamente en el tamaño de la diferencia de medias para concluir si esta diferencia se debe al azar o a un efecto real. El motivo es que la magnitud observada de la diferencia depende tanto de la variabilidad de los datos como del tamaño de la muestra. Una diferencia relativamente grande puede ser producto del azar si existe una alta dispersión en los datos o un número reducido de observaciones, mientras que una diferencia pequeña puede ser estadísticamente significativa si la variabilidad es baja y la muestra suficientemente grande.

Por esta razón, el análisis estadístico formal —a través del cálculo del estadístico t y del valor p — es imprescindible. Estos parámetros nos permiten evaluar si la diferencia observada es compatible con la hipótesis nula de igualdad de medias o si, por el contrario, existe evidencia sólida para atribuir a un efecto sistemático. En síntesis, el valor de la diferencia por sí solo no informa sobre su significancia estadística, y sólo mediante pruebas de hipótesis podemos discernir si es razonable considerarla atribuible al azar o a un efecto real.

P2.3.4 Juego bienes públicos

El entorno experimental se diseñó de manera que las únicas diferencias sistemáticas entre los grupos fueran la presencia o ausencia de la opción de castigo. Gracias a la aleatorización de los participantes y al control estricto de las condiciones (mismas reglas, mismos incentivos iniciales y mismo tipo de juego), podemos atribuir los cambios en el comportamiento observados en los períodos posteriores a la introducción de esta opción. En un contexto de laboratorio, la ausencia de factores externos que interfieran —como normas sociales heterogéneas o instituciones externas— hace más probable que la disponibilidad del castigo sea la causa directa del incremento en la cooperación o la reducción del free riding.

La comparación entre grupos en el Período 1 (antes de que el castigo entre en juego) es esencial porque nos asegura que ambos partían de condiciones iniciales equivalentes. Si las medias en el Período 1 son estadísticamente indistinguibles, podemos concluir que no existían diferencias previas entre los grupos que pudieran explicar el cambio. De esta manera, cuando observamos una divergencia clara en el comportamiento a partir de los períodos siguientes, la inferencia causal es mucho más sólida: el factor que explica la diferencia es la introducción de la opción con castigo y no características previas de los participantes o del diseño experimental.

P2.3.5 Limitaciones de los experimentos

Según Levitt y List *“What Do Laboratory Experiments Measuring Social Preferences Reveal About the Real World?”* en Los experimentos de laboratorio han sido fundamentales para estudiar las preferencias sociales, pero presentan limitaciones importantes al momento de generalizar sus resultados al mundo real. En primer lugar, los sujetos saben que están siendo observados, lo que puede inducir conductas más cooperativas o altruistas que las que tendrían en contextos anónimos. Además, las decisiones se toman en un entorno artificial que no refleja del todo las normas sociales, la experiencia acumulada ni la presión institucional que influyen en la vida cotidiana. Otro problema es la selección de participantes: suelen ser estudiantes voluntarios que no representan a toda la población y pueden estar más inclinados a colaborar. También es relevante el tamaño de los incentivos, ya que en el laboratorio las sumas suelen ser pequeñas, lo que favorece comportamientos prosociales que no necesariamente se reproducirían con montos elevados. Por último, los juegos experimentales restringen las opciones de los individuos, mientras que en la vida real las personas pueden evitar interacciones, recurrir a instituciones o diseñar estrategias alternativas.

Para abordar estas limitaciones, los investigadores han propuesto y aplicado varias soluciones. Entre ellas destacan la combinación de experimentos de laboratorio con experimentos de campo en contextos naturales, la inclusión de participantes más diversos que vayan más allá de los estudiantes universitarios, el uso de incentivos más cercanos a los que enfrentan los individuos en situaciones reales y la incorporación de normas e

instituciones que hagan los escenarios más realistas. De esta manera, los resultados de laboratorio pueden seguir siendo útiles como pruebas controladas de teoría, pero deben interpretarse con cautela y complementarse con evidencia de campo para comprender fenómenos como el free riding y el altruismo en la sociedad de manera más precisa.

Conclusión final

Del análisis realizado se pueden responder las dos preguntas centrales del taller:

1) ¿Hubo diferencias en el comportamiento (contribuciones promedio) entre los experimentos?

Experimento bienes públicos

Sí. Aunque ambos grupos comienzan con un nivel similar de contribuciones en el primer período, las trayectorias divergen con claridad. En el tratamiento sin castigo, las contribuciones caen progresivamente hasta niveles bajos en el período 10, mientras que en el tratamiento con castigo se mantienen altas y relativamente estables.

Experimento Aportes

Las contribuciones promedio disminuyen con el paso de los períodos, reflejando una pérdida de cooperación. La desviación estándar y la varianza aumenta hacia el final, lo que indica mayor dispersión en las conductas..

2) ¿Podemos atribuir las diferencias observadas al cambio en las condiciones del juego y no al azar?

Experimento bienes públicos

Sí. La evidencia descriptiva muestra que las diferencias no son producto del azar: la presencia del mecanismo de castigo explica el sostenimiento de la cooperación. Esto se confirma al observar la menor dispersión y mayor consistencia de las contribuciones en el grupo con castigo, en contraste con la alta variabilidad y caída drástica en el grupo sin castigo. En síntesis, los resultados apoyan la conclusión de que el castigo funciona como un incentivo institucional que fomenta la cooperación sostenida en contextos de provisión de bienes públicos.

Experimento Aportes

Los valores extremos (0 y 100) se mantienen presentes en ambos períodos, lo que indica una diversidad persistente en las estrategias adoptadas por los participantes. Esta variabilidad sugiere que las diferencias observadas en los comportamientos no son producto del azar, sino que pueden estar relacionadas con cambios en las condiciones del

juego. En conjunto, los resultados apuntan a que las decisiones individuales responden a los incentivos y dinámicas específicas de cada fase, más que a fluctuaciones aleatorias.