



UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO

FACULTAD DE PSICOLOGÍA

TRANSICIONES EN JUEGOS 2 X 2

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
LICENCIADO EN PSICOLOGÍA

P R E S E N T A:

Jesus Manuel Villarreal Ulloa

DIRECTOR: Dr. Arturo Bouzas Riaño

REVISOR: Dr. German Palafox Palafox

SINODALES: Dr. Florente Lopez Rodriguez

SINODALES: Dr. Oscar Zamora Arevalo

SINODALES: M en C. José Salvador Zamora Muñoz

APOYADO POR PROYECTO PAPIIT: IN307214



2016

Ciudad Universitaria, CDMX



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Resumen

En situaciones de conflicto puro, modeladas a través de la Teoría de Juegos, se ha encontrado que el comportamiento de participantes humanos se aproxima a la solución óptima conforme adquieren experiencia en un juego en el cual las consecuencias del comportamiento no varían a través del tiempo. En el presente experimento, se aborda una situación más realista en la cual se presentan cambios en los pagos que alteran la solución óptima de un juego a lo largo del tiempo, sin transformar la estructura general de la situación en la que se encuentran los jugadores. Para abordar este problema, se utilizaron juegos asimétricos de conflicto puro con dos alternativas para cada jugador (2×2) con 400 repeticiones, introduciendo un cambio en la estructura de los pagos para cada jugador después de cada 100 ensayos. Los resultados muestran que en equilibrio, el modelo de “Equilibrio en respuesta cuantil” (McKelvey y Palfrey, 1995) es el que mejor aproxima la proporción de ocasiones en que los sujetos eligen una de sus alternativas. La dinámica del comportamiento de los sujetos ante los cambios en la función puede resumirse en 3 clases: el comportamiento no cambia a través del tiempo, el comportamiento cambia en la dirección del equilibrio, y sujetos en los cuales el comportamiento cambia en la dirección contraria. Sin embargo, estos tres tipos de comportamiento pueden ser descritos por el “Modelo de Aprendizaje de Atracciones Ponderadas por la Experiencia” (Ho, Camerer y Chong, 2007).

Índice general

| | |
|---|-----------|
| Resumen | 3 |
| Introducción | 11 |
| 1. Marco Teórico | 17 |
| 1.1. Teoría de Juegos | 17 |
| 1.2. Cambio y Adaptación | 23 |
| 1.3. Modelos Conductuales de Equilibrio | 30 |
| 1.4. Aprendizaje en Juegos | 40 |
| 1.4.1. Aprendizaje por Reforzamiento | 42 |
| 1.4.2. Aprendizaje por Igualación de Impulsos | 46 |
| 1.4.3. Juego Ficticio Estocástico | 48 |
| 1.4.4. St-EWA | 52 |
| 2. Método | 65 |
| 2.1. Procedimiento | 67 |
| 3. Resultados | 71 |
| 3.1. Análisis General | 72 |

| | |
|--|------------|
| 3.1.1. Proporción de Elecciones | 72 |
| 3.1.2. Proporción de Elecciones a Través del tiempo | 77 |
| 3.2. Equilibrio | 81 |
| 3.2.1. Distribuciones Posteriores y Posterior Predictiva | 81 |
| 3.2.2. Comparación de Modelos | 93 |
| 3.3. Aprendizaje | 106 |
| 3.3.1. Distribuciones Posteriores y Posterior Predictiva | 107 |
| 3.3.2. Comparación de Modelos | 121 |
| 4. Discusión | 133 |
| 5. Conclusión | 143 |
| Referencias | 147 |

Índice de figuras

| | |
|--|-----|
| 1.1. Equilibrio de Nash y Equilibrio en Respuestas Cuantales | 31 |
| 3.1. Proporción de Respuestas por Juego | 73 |
| 3.2. Proporción de Elecciones a Través del Tiempo | 78 |
| 3.3. Modelo Gráfico: PSE | 82 |
| 3.4. Probabilidad Posterior del Parámetro η | 84 |
| 3.5. Posterior Predictiva PSE | 86 |
| 3.6. Modelo Gráfico: QRE | 87 |
| 3.7. Distribución Posterior QRE | 88 |
| 3.8. Posterior Predictiva QRE | 92 |
| 3.9. Modelo Gráfico para Comparación de Modelos de Equilibrio . | 94 |
| 3.10. Distribuciones Posteriores Predictivas por Modelo | 95 |
| 3.11. Distancia Cuadrática Media por Modelo | 98 |
| 3.12. Probabilidad Media Posterior por Modelo | 105 |
| 3.13. Simulación del Modelo IM | 108 |
| 3.14. Modelo Gráfico: CR | 111 |
| 3.15. Distribución Posterior y Posterior Predictiva: REL | 112 |
| 3.16. Modelo Gráfico: SFP | 114 |

| | |
|--|-----|
| 3.17. Distribución Posterior y Posterior Predictiva: SFP | 116 |
| 3.18. Modelo Gráfico: St-EWA | 117 |
| 3.19. Distribución Posterior y Posterior Predictiva: STE | 119 |
| 3.20. Puntuación Cuadrática Media | 122 |
| 3.21. Modelo Gráfico Para la Comparación de Modelos de Aprendizaje | 124 |
| 3.22. Media de la Probabilidad Posterior Por Modelo | 129 |
| 3.23. Modelo Gráfico Para la Comparación de Modelos de Aprendizaje | 130 |
| 3.24. Media de la Probabilidad Posterior Por Modelo y Tipo de Jugador | 131 |

Índice de cuadros

| | |
|--|-----|
| 1.1. “Mathicng Pennies” | 19 |
| 2.1. Juegos Experimentales (Método) | 66 |
| 2.2. Juego Fase de Entrenamiento | 66 |
| 2.3. Orden de Presentación | 67 |
| 3.1. Juego “A1” | 74 |
| 3.2. Juego “B1” | 74 |
| 3.3. Juego “A2” | 75 |
| 3.4. Juego “B2” | 76 |
| 3.5. Puntos de Equilibrio modelo: IBE | 96 |
| 3.6. Probabilidad Posterior Por Pareja Juego A1 | 99 |
| 3.7. Probabilidad Posterior Por Pareja Juego B1 | 101 |
| 3.8. Probabilidad Posterior Por Pareja Juego A2 | 102 |
| 3.9. Probabilidad Posterior Por Pareja Juego B2 | 104 |
| 3.10. Probabilidad Posterior por Modelo J. Fila | 125 |
| 3.11. Probabilidad Posterior por Modelo J. Columna | 126 |
| 3.12. Clasificación por Modelo | 127 |

Introducción

Uno de los supuestos con mayor arraigo en la Psicología, es que el comportamiento de los organismos se encuentra controlado por sus consecuencias. El estudio de la interacción social no es la excepción, sin embargo, en este caso las consecuencias no dependen únicamente de las acciones de un organismo, sino de las relaciones que existen entre las acciones de todos los organismos que participan de una situación particular.

Esta relación de dependencia entre las consecuencias del comportamiento de los organismos y el comportamiento de otros es una de las características más relevantes de una situación de interacción. La Teoría de Juegos (TJ) permite abstraer este tipo de relaciones a partir del uso de reglas, especificando cuál es la consecuencia de un organismo cuando lleva a cabo una acción tomando en cuenta las acciones de los demás, en este caso, las consecuencias de cada organismo pueden variar dependiendo de las acciones de otros. Otras de las características relevantes que la TJ permite modelar mediante el uso de reglas son, si los organismos eligen de manera simultánea y sin conocer las acciones de los demás o eligen de manera secuencial; o cuántas acciones tienen disponibles quienes participan en la situación. Estas especificaciones definen las reglas de un juego y mediante su uso es posible modelar situaciones de

competencia, cooperación o coordinación.

Dentro de las situaciones que pueden ser modeladas a través del uso de la TJ, la instancia más simple son las situaciones de conflicto puro en las cuales los intereses de los organismos son completamente opuestos, es decir, que las ganancias de un organismo representan las pérdidas del otro, además de que sólo dos organismos participan de la situación cada uno con dos alternativas disponibles. Este tipo de situaciones se conocen como Juegos de suma cero 2×2 . Debido a que son la instancia más simple han sido empleadas ampliamente en la experimentación. En este caso, uno de los hallazgos más importantes indica que los jugadores pueden distribuir su comportamiento de una manera cercana a lo que se considera una solución óptima (O'Neill, 1987), denominada Equilibrio de Nash (EN; Nash, 1950).

El EN se refiere al punto en el cual ninguno de los organismos que forman parte del juego tiene incentivos para cambiar su comportamiento, es decir, tomando en cuenta el comportamiento de los demás, un cambio en el comportamiento del organismo no le generará mayores ganancias.

Estudios posteriores mostraron que el comportamiento de participantes humanos no siempre se ajusta a la solución óptima de un juego, sino que se presentan algunas discrepancias. Estos resultados llevaron al planteamiento de modelos de equilibrio que fueran capaces de dar cuenta del comportamiento de los sujetos en situaciones experimentales. Este tipo de modelos se denominan modelos conductuales de equilibrio y se tratarán con mayor detalle en la sección 1.3.

Uno de los principales supuestos del EN como solución óptima, implica que los sujetos deben tener información completa acerca de la situación en

la que se encuentran; en otras palabras, para que el EN pueda sea considerado una solución óptima, los jugadores deben conocer sus propios pagos, los pagos (preferencias) del oponente, las acciones disponibles de ambos, etc. Este supuesto, fue puesto en duda desde la formalización de la teoría (Von Neumann y Morgenstern, 1945). El argumento en contra radica en que es poco plausible que los jugadores cuenten con la información completa sobre la situación y las preferencias de su oponente desde el inicio de la interacción, en cambio, es posible que los jugadores adquieran dicha información en encuentros sucesivos con oponentes similares bajo las mismas condiciones (reglas).

Los mismos autores presentaron su argumento como la necesidad de una teoría dinámica de juegos, la cual fuese capaz de dar cuenta del proceso mediante el cual los jugadores adquirirían información, así como el comportamiento que se esperaría de los sujetos una vez que contaran con la información completa. Dentro de la TJ, los modelos de aprendizaje surgieron como una vía para formalizar una teoría dinámica. El estudio del aprendizaje en juegos ha derivado en dos ramas importantes, la primera es el estudio de las propiedades a largo plazo de los diferentes modelos que han sido propuestos; la segunda, estudia a nivel empírico, qué modelos son capaces de replicar el proceso de adaptación de los organismos cuando éstos se encuentran en una situación de interacción. Los modelos de aprendizaje y los resultados de su comparación se tratarán con mayor detalle en la sección 1.4.

Uno de los problemas que han surgido al comparar modelos de aprendizaje en situaciones de interacción, es la dificultad que existe para discriminar entre dos o más modelos. Debido a que los modelos de aprendizaje en juegos son

similares en términos de sus supuestos a los a los modelos de aprendizaje en elección individual, podría resultar de utilidad el incorporar la metodología que ha sido utilizada en este último caso para poder discriminar entre las predicciones de diferentes modelos de aprendizaje.

La metodología que ha permitido contrastar de mejor manera los modelos de aprendizaje en situaciones individuales, es el estudio del aprendizaje en ambientes que cambian con el tiempo. Esto se debe principalmente a que cada modelo responde de manera diferente a los cambios en la estructura de los pagos, por lo que el comportamiento de los organismos en estas situaciones se vuelve más informativo en términos de las predicciones de cada modelo. Algunos autores han probado el ajuste de los organismos a cambios en la estructura de los pagos en una situación experimental, sin embargo, existen algunos problemas metodológicos los cuales se trataran con mayor detalle en la sección 1.2.

Los resultados del presente experimento indican que el comportamiento de participantes humanos cambia cuando cambia la estructura de los pagos de un juego sin alterar la estructura general (reglas) del mismo. El cambio en el comportamiento de los participantes no siempre presenta la misma dirección, no obstante, es posible dar cuenta del comportamiento de los participantes, tanto a un nivel “estacionario” como a nivel dinámico, utilizando modelos conductuales de equilibrio y aprendizaje respectivamente.

El modelo de equilibrio que es capaz de dar cuenta del comportamiento de los participantes es áquel que supone que la utilidad de los sujetos se ve afectada por perturbaciones que tienen una distribución de probabilidad determinada, a este modelo se le conoce como *Equilibrio en respuesta cuantal*

(“*Quantal Response Equilibrium*”; QRE, por sus siglas en inglés).

Por otro lado, el modelo de aprendizaje que es capaz de dar cuenta de la dirección y magnitud del cambio en el comportamiento de los participantes, es el de *Modelo de aprendizaje auto-regulable de atracciones ponderadas por la experiencia* (“*Self-tuning Experience Weighted Attraction Learning*”; St-EWA, por sus siglas en inglés; Ho, Camerer y Chong, 2007).

Las siguientes secciones se dividen de la siguiente manera, en la sección 1.1 se presenta con mayor detalle la Teoría de Juegos. En la sección 1.2 se abordara el tema del ajuste a cambios en situaciones de elección individual y estudios previos que han buscado incorporar la noción de ajuste a cambios en situaciones de interacción. En la sección 1.3 se presentan los modelos conductuales de equilibrio y, en la 1.4, los modelos de aprendizaje.

Marco Teórico

1.1. Teoría de Juegos

Como se mencionó en la introducción, utilizando las reglas de un juego es posible modelar diferentes situaciones de interacción que van desde el conflicto puro hasta situaciones de cooperación. Formalmente las reglas de un juego deben establecer lo siguiente:

- i) El número de jugadores que se encuentran involucrados $N = (1, 2, \dots, n)$
- ii) Un conjunto de estrategias $s_i = \{1, 2, \dots, k\}$ para cada jugador $i = \{1, 2, \dots, n\}$.
- iii) Una función de pago $h_i(s_i, s_{-i})$ para cada jugador $i = 1, 2, \dots, n$, en donde s_{-i} representa las estrategias de los oponentes del i -ésimo jugador. Si $h_i(s_i, s_{-i}) > 0$ el jugador i recibe un pago, si $h_i(s_i, s_{-i}) < 0$ el jugador debe realizar un pago y si $h_i(s_i, s_{-i}) = 0$ ambos jugadores quedan igual.

Además de estas reglas, en un juego se debe especificar el orden en el que los jugadores elegirán entre sus estrategias, es decir, si todos los jugadores elegirán al mismo tiempo y sin conocimiento de las elecciones de los otros o si

los jugadores elegirán por turnos. A estos juegos se les denomina simultáneos y secuenciales respectivamente (Von Neumann y Morgenstern, 1945).

Utilizando estas reglas, los juegos utilizados en el presente experimento son una instancia de un juego simultáneo de dos jugadores con dos estrategias disponibles para cada jugador; en los cuales la función de pago cuenta con dos restricciones:

$$\sum_{i=1}^2 h_i(s_i, s_{-i}) = 0 \quad (1.1)$$

$$h_1(s_i = s_{-i}, s_{-i}) > h_1(s_i \neq s_{-i}, s_{-i}) \quad (1.2)$$

$$h_2(s_i = s_{-i}, s_{-i}) < h_2(s_i \neq s_{-i}, s_{-i})$$

La primera implica que, sin importar las estrategias que utilice cada jugador, la suma de los pagos de ambos siempre es igual a cero. Cuando la función de pagos cuenta con esta restricción se dice que el juego es de suma cero. Este tipo de pagos se ha utilizado para modelar situaciones de conflicto y se consideran la instancia más simple debido a que no existen incentivos que puedan llevar a los jugadores a cooperar.

La segunda restricción en la función de pagos indica que el pago para uno de los jugadores (jugador fila o jugador 1) siempre será mayor si utiliza la misma estrategia que su oponente, mientras que el pago para el otro jugador será mayor cuando las estrategias de ambos sean diferentes (jugador 2), mantener esta restricción hace que, aún cuando se cambian los valores específicos de la función de pago, el juego siga siendo el mismo (las reglas se mantienen iguales); a los juegos que mantienen esta segunda restricción se les conoce como “Matching Pennies Generalizado”. En el cuadro 1.1 se presenta la matriz de pagos del juego de “Matching Pennies” en su forma

más simple. Los juegos utilizados en el experimento son generalizaciones de éste, es decir, en los juegos utilizados solo se alteraron los pagos dejando las mismas restricciones.

| | | |
|--------|--------|-----|
| | Águila | Sol |
| Águila | 1 | -1 |
| Sol | -1 | 1 |

Cuadro 1.1: Juego de “Matching Pennies”, este juego fue utilizado durante la fase de entrenamiento de los participantes.

En el caso del juego en el cuadro 1.1 las filas representan las estrategias de uno de los jugadores (jugador Fila), mientras que las columnas representan las del otro. Las intersecciones de la matriz muestran la función de pagos para el jugador 1 o jugador Fila, los pagos del jugador columna se han omitido debido a que se obtienen multiplicando cada pago por -1 . Por ejemplo, si ambos jugadores eligieran “Águila” o “Sol”, el jugador Fila ganaría una unidad y su oponente (jugador 2 o Columna) perdería la misma cantidad; en cambio, si las elecciones fueran diferentes “Águila” para un jugador y “Sol” para el oponente, el jugador Fila perdería una unidad y su oponente ganaría la misma cantidad. Los juegos utilizados en este experimento mantuvieron las reglas que se mencionaron anteriormente, pero alterando las magnitudes de los pagos.

En la situación modelada por el juego de “Matching Pennies” el jugador Fila ganará el juego cuando sea capaz de predecir la estrategia que elegirá su oponente, mientras que el jugador columna ganará el juego solo cuando las estrategias de ambos sean diferentes. Von Neumann y Morgenstern (1945) afirmaron que esta era una de las características más relevantes de una si-

tuación de conflicto, es decir, en una situación de conflicto, los jugadores siempre tienen incentivos para predecir el comportamiento de su oponente y al mismo tiempo, tienen incentivos para ser impredecibles.

Dadas las reglas de este juego ¿Cuál es la mejor estrategia que puede llevar a cabo un jugador cuando no conoce la elección de su oponente? Para poder dar una respuesta a la pregunta anterior, es necesario utilizar el supuesto de que los jugadores buscan maximizar sus ganancias (o minimizar sus pérdidas). El equilibrio de Nash es la respuesta más general a esta pregunta e indica qué estrategias por parte de cada jugador llevarán a que nadie tenga incentivos para cambiar su comportamiento. Es decir, el equilibrio de Nash es el punto en el cual ninguno de los jugadores incrementaría sus ganancias cambiando su comportamiento, cuando el comportamiento de los oponentes permanece constante (Young, 2004).

Como se mencionó anteriormente, en situaciones de conflicto los jugadores por lo general tienen incentivos para predecir el comportamiento de sus oponentes y al mismo tiempo ser impredecibles (Von Neumann y Morgenstern, 1945). Por lo tanto, la mejor estrategia de cada jugador sería seleccionar de manera aleatoria entre sus dos opciones, a este tipo de estrategias se les denomina estrategias mixtas. Una estrategia mixta implica que los jugadores ya no elegirán una estrategia particular de las que tienen disponibles, sino que elegirán entre diferentes distribuciones de probabilidad sobre sus estrategias (Nash, 1950). La solución óptima en estrategias mixtas se encuentra buscando para cada jugador, cuál es la distribución de probabilidad sobre sus opciones (*e.g.* “Águila” y “Sol”) que llevaría a que el oponente sea indiferente entre sus alternativas.

En el caso del ejemplo en el cuadro 1.1 esta distribución es elegir cada una de las opciones con una probabilidad de 0.5, de esa manera, los pagos esperados para cada alternativa del oponente serán iguales entre sí. El mismo argumento se aplica para el segundo jugador, es decir, el jugador Columna también debe de elegir una distribución de probabilidad sobre sus propias opciones que lleve al jugador Fila a ser indiferente entre sus alternativas. Al punto formado por las dos distribuciones (una por jugador) se le conoce como equilibrio en estrategias mixtas.

Estos equilibrios cumplen con la noción de que si alguno de los jugadores cambia su comportamiento, el oponente podría aprovecharse, por ejemplo, si el jugador Fila comenzará a utilizar con mayor probabilidad su estrategia “Sol”, su oponente, tendría una ganancia esperada mayor si comienza a usar su estrategia “Águila” en cada ocasión posible.

No obstante, para que el EN sea considerado una solución óptima en un juego, tienen que cumplirse ciertos supuestos acerca de la información con la que cuenta cada uno de los jugadores, es decir, es necesario que los jugadores sepan cuántas estrategias tienen disponibles los oponentes y cuáles son sus pagos. Estos supuestos, aunados al de maximización de las ganancias, convierten al EN en una solución óptima. En otras palabras, el EN implica un conocimiento pleno por parte de cada jugador sobre la situación en la que se encuentra.

Los equilibrios representan el comportamiento óptimo de los sujetos cuando éstos ya no tienen nada que aprender, es decir, pueden entenderse como el límite de algún proceso de ajuste. Los modelos de aprendizaje en juegos buscan dar cuenta de este proceso de ajuste mediante el uso de supuestos

acerca de la forma en la que los organismos adquieren información sobre la situación en la que se encuentran a través del tiempo.

A diferencia de los modelos de equilibrio, los modelos de aprendizaje permiten representar las trayectorias que podría seguir el comportamiento de los sujetos a través del tiempo en una situación de interacción, lo cual a su vez permite predecir en qué situaciones se esperaría que los sujetos actuarán de acuerdo con el concepto de equilibrio y en qué tipo de situaciones se esperaría que el comportamiento de los sujetos fuese diferente.

Si bien los modelos de aprendizaje tienen ventajas con respecto a los modelos de equilibrio, vale la pena resaltar que a diferencia de éstos, los modelos de aprendizaje requieren de una gran cantidad de información para poder llevar a cabo predicciones, es decir, estos modelos requieren que se cuente con información ensayo a ensayo mientras que los de equilibrio requieren únicamente de las características generales de un juego para generar predicciones. Debido a esto no es posible dejar de lado el estudio de los modelos de equilibrio.

Sin embargo, antes de profundizar en los modelos de aprendizaje y equilibrio que han sido presentados en la literatura en situaciones de interacción, es necesario presentar el concepto de ambientes dinámicos (ambientes que cambian a través del tiempo), así como las relaciones entre este tipo de ambientes y las nociones de equilibrio y aprendizaje en situaciones de elección individual. En la siguiente sección se define este concepto y se discutirán dos experimentos previos que han buscado introducir la noción de cambio en situaciones de interacción.

1.2. Cambio y Adaptación

Una de las propiedades de los ambientes naturales, es que las relaciones entre las acciones de los organismos y las consecuencias de su comportamiento no son estáticas sino que cambian a lo largo del tiempo. Un ejemplo de esto serían los cambios en la disponibilidad de alimento asociados con diferentes estaciones del año en los que es posible predecir el siguiente estado del sistema (ambiente) a partir del estado actual; en este caso, el cambio tiene una estructura determinada, es decir, se sabe que después de la primavera sigue el verano. A estos cambios se les denomina “cambios predecibles” (O’Reilly, 2013), y tienen la propiedad de que el estado al que llegarán conforme pasa el tiempo puede ser predicho a partir del estado actual.

Por otro lado, existen cambios que pueden llevar a un estado desconocido. Por ejemplo, un día por la mañana puede ser soleado, sin embargo, existe la posibilidad de que en la tarde se encuentre nublado, que siga soleado o incluso que llueva, en ese caso, los tres estados posibles son (sol, nublado o lluvia), no obstante, aunque se conocen los posibles estados, no se puede predecir cuál de los tres ocurrirá. Este tipo de cambios se denominan “cambios impredecibles”, y aunque son una función del tiempo, el nuevo estado del sistema no depende, o no puede ser predicho, a partir del estado anterior (O’Reilly, 2013).

El problema que enfrentan los organismos al encontrarse con ambientes que cambian en función del tiempo, es el de detectar cuándo ha ocurrido un cambio, debido a que la ocurrencia de un cambio en el ambiente determina si la información que los organismos pudieron haber acumulado sigue siendo relevante. En ambientes de este tipo, la tarea para los organismos es integrar

la información extraída de las consecuencias recibidas en el pasado con la información extraída de las consecuencias más recientes.

El problema de la integración de información es un problema fundamental en el aprendizaje, sin embargo, cuando existe la posibilidad de la ocurrencia de un cambio, éste adquiere una nueva dimensión, ya que la estrategia óptima para un organismo dependerá de si el ambiente ha sufrido un cambio. Por ejemplo, cuando la variabilidad en el ambiente permanece estable, los organismos deberían otorgarle mayor peso a la información acumulada, por otra parte, cuando el ambiente en el que se encuentran los organismos cambia frecuentemente, éstos deberían otorgarle mayor peso a la información más reciente (O'Reilly, 2013).

Cuando un organismo se enfrenta con ambientes que cambian de forma impredecible a través del tiempo, surgen dos preguntas, la primera: ¿son los organismos capaces de adaptarse a este tipo de ambientes? y de ser así, ¿cuál es el curso que sigue la adaptación a estos cambios?. Este problema ha sido abordado experimentalmente en el ámbito de la elección individual.

Uno de los hallazgos más importantes en la literatura sobre elección individual es que los organismos son capaces de adaptarse de manera rápida cuando las consecuencias de su comportamiento cambian de forma impredecible aun cuando el cambio no es señalado (*e.g.* Bailey y Mazur, 1990; Mazur, 1992; Mark y Gallistel, 1994). Además, la velocidad a la que se adaptan los organismos puede variar dependiendo de la frecuencia con la que ha enfrentado cambios en el pasado, siendo la velocidad de adaptación mayor cuando los cambios se presentan de manera frecuente en contraste a cuando éstos son esporádicos (*e.g.* Mazur, 1997).

Este tipo de resultados representan un reto para los modelos de aprendizaje “clásicos”, ya que uno de sus supuestos básicos es que la información acumulada y la información de los ensayos más recientes se ponderan de manera constante a través del tiempo. Por ejemplo, la instancia más simple de los modelos de aprendizaje conocidos como integradores (precursores del aprendizaje por reforzamiento), utilizan el supuesto de que la información acerca de las consecuencias que un organismo ha encontrado en el pasado y las consecuencias más recientes se combinan de manera lineal (*ver ecuación 1.3*; Staddon, 2001).

$$V(t) = \alpha V(t - 1) + (1 - \alpha)X(t - 1) \quad (1.3)$$

La ecuación 1.3 muestra el modelo de integrador, en este modelo V representa el valor asignado a una alternativa en el ensayo t , X representa la consecuencia en ese ensayo y α es un parametro que descuenta la información acumulada hasta el ensayo actual mientras que la última consecuencia recibida es descontada por $(1 - \alpha)$. En este caso al parametro α se le denomina “ventana temporal” ya que otorga un peso diferente a cada tipo de información.

El problema que enfrentan los modelos que suponen que el factor de ponderación α es constante, en entornos en los cuales se presentan cambios aleatorios en las consecuencias del comportamiento, es que si le otorgan mayor peso a las consecuencias pasadas los organismos serían menos susceptibles a los cambios en el ambiente; por otra parte, si se le otorga mayor peso a las observaciones más recientes, existe la posibilidad de que los organismos detecten como un cambio en las consecuencias lo que es sólo la variabilidad

del entorno.

Para poder dar cuenta de un proceso de adaptación en ambientes dinámicos, esta clase de modelos requieren del supuesto de que la proporción a la que se descuentan las experiencias previas (α) y cómo se acumula la información sobre los últimos ensayos varía a través del tiempo. Modelos con estas características ya han sido utilizados (*e.g.* O'Reilly, 2013; Behrens, *et al.*, 2007) pero no serán abordados debido a que no se han propuesto modelos similares en juegos.

En el caso de las situaciones de interacción (juegos) el problema se magnifica debido a que no solo las consecuencias de las elecciones de los organismos pueden cambiar a través del tiempo, sino que también el comportamiento de los oponentes puede cambiar. Por ejemplo, cuando dos o más sujetos se encuentran interactuando, uno podría aprender acerca de las elecciones del otro, de manera que sea capaz de distribuir su comportamiento de manera más eficiente. Sin embargo, los demás también podrían aprender acerca del primero y en respuesta, modificar su comportamiento. En otras palabras, en el caso de las situaciones de interacción, *“aprender modifica aquello que debe ser aprendido”* (Young, 2007).

Sin embargo, las reglas que definen la estructura del juego (principalmente la función de pagos) restringen las posibles combinaciones en el comportamiento de ambos jugadores. La tarea de los modelos de aprendizaje en situaciones de interacción, es dar cuenta de cómo se adaptan los organismos a estas diferentes estructuras de pagos.

Cuando la función de pagos cambia a lo largo del tiempo, el problema para los modelos de aprendizaje en juegos es similar al que enfrentan en

situaciones de elección individual cuando el entorno no es estacionario.

Al revisar la literatura empírica, se encontraron sólo dos experimentos publicados que han estudiado el comportamiento en juegos cuando ocurren cambios en las relaciones entre las estrategias de un jugador y los pagos que éste recibe. El primero fue llevado a cabo por Ochs en 1995 con participantes humanos, el segundo fue llevado a cabo por Sanabria y Thrailkill en 2009 con palomas.

Ochs (1995) utilizó tres generalizaciones del juego de “Matching Pennies” con el objetivo de estudiar si después de un cambio en la función de pagos, el comportamiento de los jugadores se ajustaba al nuevo equilibrio de Nash. El primer juego utilizaba la función de pagos que se presentó en el cuadro 1.1, con la excepción de que no se utilizaban pagos negativos, es decir, si alguno de los participantes “perdía”, la consecuencia era no recibir puntos.

La manipulación importante consistió en cambiar a los sujetos a un nuevo juego en el cual se modificaba el pago sólo para el jugador Fila, de tal forma que si ambos jugadores utilizaban su estrategia “Águila”, el jugador Fila ahora recibía un pago ya sea de 4 o 9 dependiendo de la condición; mientras que todos los otros pagos en la matriz permanecían constantes.

El cambio en la matriz de pagos introducido por Ochs, lleva a que el equilibrio de Nash cambie para el sujeto cuyos pagos permanecían constantes, sin embargo, el equilibrio sigue siendo una estrategia mixta.

Los resultados de Ochs (1995) mostraron que, en general, los jugadores no se ajustaban al nuevo equilibrio del juego, mostrando incluso en algunos casos un cambio en el comportamiento en la dirección contraria a la predicción del equilibrio de Nash.

No obstante, existen algunas características del método que podrían explicar estos resultados. En primer lugar, en los juegos utilizados por Ochs al cambiar la matriz de pagos, el equilibrio de Nash se modifica únicamente para uno de los jugadores (jugador Columna), mientras que el del oponente se mantiene constante. En este caso, si durante el primero de los juegos el jugador Fila se encontrara utilizando sus estrategias de manera consistente con la solución óptima del juego, al momento de cambiar, el oponente no tendría ningún incentivo para cambiar su comportamiento para aproximarse al nuevo equilibrio.

Un segundo problema es que cada jugador elegía para cada 10 repeticiones del juego la combinación de águilas y soles que utilizaría. El problema con esta metodología es que altera la distribución de probabilidad sobre el número de veces en que un jugador puede ganar cuando el juego se repite en una cantidad determinada de ocasiones. Por lo tanto, en el experimento de Ochs los jugadores estarían recibiendo una retroalimentación diferente a la del juego original, lo cual podría alterar el comportamiento de los participantes (Shachat, 2002).

Sanabria y Thrailkill (2009), utilizaron palomas a las cuales se les otorgaba como “pago” un tiempo determinado de acceso a un reforzador dependiendo de la combinación de elecciones de dos palomas. Las funciones de pagos utilizadas son similares a las que se presentaron en el cuadro 1.1, sin embargo, los pagos negativos fueron reemplazados por “*Black outs*” de la misma duración que las ganancias, por lo que podrían ser interpretados como pagos negativos. Al igual que en el experimento de Ochs, estos juegos también son una generalización del juego de MP.

El objetivo del experimento de Sanabria y Thraillkill (2009), fue el de observar si las elecciones de las palomas se aproximaban a las predicciones del equilibrio de Nash después de un cambio en la matriz de pagos. Al igual que en el caso de Ochs, el cambio consistía en la modificación de uno de los pagos del juego. En este caso, los autores reportaron que la proporción de veces que cada paloma elegía cada una de sus estrategias, se aproximaba con el paso de las sesiones a la solución óptima del juego. Además, cuando cambiaba la función de pagos, el comportamiento de las palomas se ajustaba a la nueva solución óptima.

Los resultados de ambos experimentos son contrastantes, no obstante, vale la pena destacar que en el experimento de Sanabria y Thraillkill se utilizaron únicamente dos parejas de palomas. Además existe el problema de que, de acuerdo con los autores, no es claro si las respuestas de los animales se estabilizaron con el paso del tiempo, aunque la dirección del cambio en el comportamiento presentara la dirección predicha por el equilibrio de Nash.

Si bien el ajuste a cambios ha sido ampliamente estudiado en ambientes de un solo agente, éste no ha sido el caso para los ambientes interactivos. En comparación con la adaptación a cambios en situaciones de elección individual, la adaptación en situaciones de interacción podría ser más complicada ya que, como se mencionó anteriormente, en este tipo de situaciones las consecuencias que recibe un organismo por su comportamiento dependen de la interacción entre la estructura general del juego (reglas) y la distribución del comportamiento del oponente. El problema radica en que, para estas situaciones, el cambio podría darse en cualquiera de estos dos niveles.

Los modelos de aprendizaje que han sido utilizados para dar cuenta del

proceso de adaptación en juegos repetidos, toman en cuenta el comportamiento de los oponentes como parte del proceso de adaptación. Estos mismos modelos podrían tener problemas cuando las reglas (función de pagos) cambian a través del tiempo, ya que utilizan los mismos supuestos que los modelos en aprendizaje individual. Sin embargo, el problema es que no se cuenta con datos experimentales con los cuales estos modelos puedan ser comparados en este tipo de situaciones. En las siguientes secciones se presentan los modelos de equilibrio y de aprendizaje que fueron comparados en el presente experimento.

1.3. Modelos Conductuales de Equilibrio

Los modelos de equilibrio y los modelos de aprendizaje surgieron en la TJ como una alternativa al equilibrio de Nash con miras a dar cuenta del comportamiento de sujetos reales en situaciones de interacción. La necesidad de estos modelos surgió a partir de los resultados de diversos experimentos (*e.g.* Brown y Rosenthal, 1990; Eliaz y Rubinstein, 2011; Goeree, Holt y Palfrey, 2003; Mckelvey y Palfrey, 1995, etc.) que mostraron que, si bien el comportamiento de los participantes se aproximaba al equilibrio de Nash, existen diferencias entre el comportamiento agregado de los sujetos y dicho equilibrio.

La figura 1.1 muestra de manera gráfica este problema¹. Cada punto en la gráfica representa el comportamiento agregado de los sujetos y el equilibrio de Nash para experimentos en los cuales se utilizaron juegos de dos jugadores

¹Gráfica tomada de *Behavioral Game Theory: Experiments in Strategic Interaction* de Collin F. Camerer (2003).

con dos estrategias para cada jugador con solución en estrategias mixtas. La diagonal que aparece en la gráfica representa una correlación perfecta, en este caso, si los jugadores se comportaran de acuerdo con la predicción teórica se esperaría que los puntos se aproximaran a esta recta, sin embargo, esto no ocurre en todos los casos, ni siquiera entre juegos que comparten una solución.

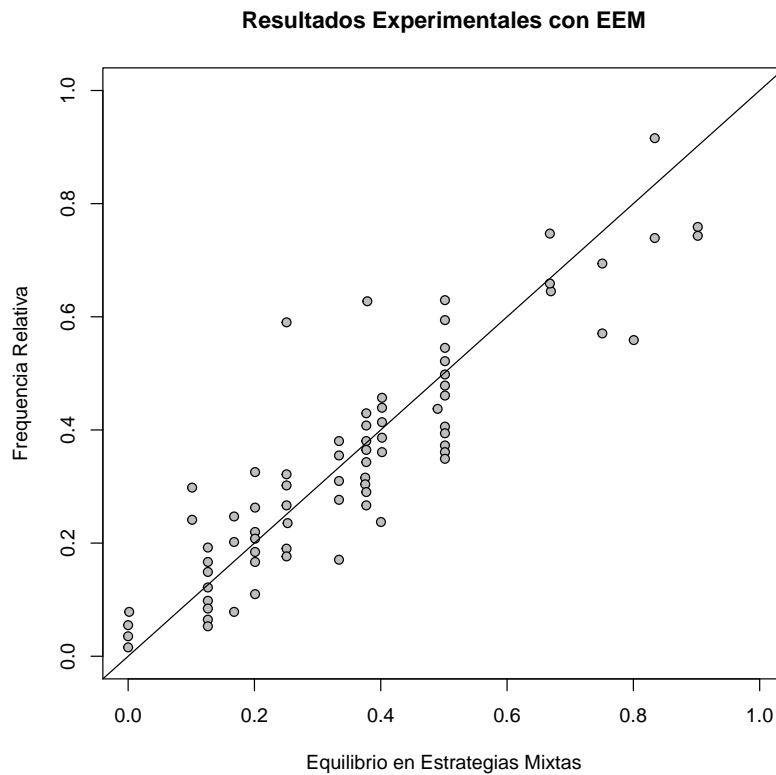


Figura 1.1: Gráfica tomada de Camerer (2003). Cada uno de los puntos en la gráfica representa un experimento con el equilibrio de Nash en un experimento en las abscisas y los resultados en términos de frecuencia relativa en las ordenadas.

Aunado a esto, se ha argumentado que los supuestos del equilibrio de Nash son poco plausibles a nivel psicológico ya que, como se menciono an-

teriormente, requiere que los sujetos tengan información completa acerca de la situación en la que se encuentran y que sean capaces de maximizar sus ganancias. Los modelos conductuales de equilibrio surgieron a apartir de la modificación de estos supuestos y su principal objetivo ha sido el de dar cuenta del comportamiento en situaciones reales.

Los modelos de equilibrio que se compararon en el presente experimento son:

- i) Equilibrio de Nash (“*Nash Equilibrium*”; NE, por sus siglas en inglés; Nash, 1950).
- ii) Equilibrio en Respuestas Cuantales (“*Quantal Response Equilibrium*”; QRE, por sus siglas en inglés; McKelvey y Palfrey, 1995).
- iii) Equilibrio por Muestreo de Pagos (“*Payoff Sampling Equilibrium*”; PSE, por sus siglas en inglés; Osborne y Rubinstein, 1998).
- iv) Equilibrio por Balance de Impulsos (“*Impulse Balance Equilibrium*”; IBE, por sus siglas en inglés; Selten, Abbink y Cox, 2005)

De estos modelos, solamente el de QRE puede ser considerado como una generalización del equilibrio de Nash, ya que solo “relaja” algunos de sus supuestos; en cambio, los modelos de PSE e IBE utilizan un conjunto de supuestos diferentes de los cuales sólo algunos pueden relacionarse con el dicho concepto de equilibrio.

El modelo de QRE fue presentado por McKelvey y Palfrey como una alternativa paramétrica al modelo de equilibrio de Nash. A diferencia de este

último, el modelo de QRE “debilita” el supuesto de maximización de las ganancias sustituyéndolo por una forma “suavizada” de maximización.

El principal supuesto detrás de este modelo, es que la utilidad² de cada alternativa tiene asociado un error aleatorio, es decir, se asume que la utilidad es una variable aleatoria. En este caso, la probabilidad de que un jugador elija una de sus estrategias depende de los pagos asociados a cada alternativa, la probabilidad con la que el oponente elige entre sus alternativas y la variabilidad asociada a la utilidad de las alternativas del jugador.

Al igual que en el equilibrio de Nash, el punto de equilibrio se alcanza cuando ninguno de los jugadores tiene incentivos para cambiar su comportamiento, manteniendo el comportamiento de los oponentes constante. La probabilidad de que un jugador elija una de sus estrategias disponibles se expresa de la siguiente manera:

$$p_i(s_i = j) = \frac{e^{\frac{\bar{u}_{ij}}{\mu}}}{\sum_{k \in s_i} e^{\frac{\bar{u}_{ik}}{\mu}}} \quad (1.4)$$

En donde \bar{u}_{ij} representa el pago esperado para el i -ésimo jugador tomando en cuenta la distribución de probabilidad que el oponente asigna sobre sus estrategias.

La función 1.4 liga las utilidades asociadas a cada estrategia con una probabilidad de elección por parte de los jugadores; son muchas las formas que puede tomar esta función, pero ésta dependerá del supuesto que se realice acerca del comportamiento que sigan los errores asociados a las utilidades.

²Cantidad que esperaría ganar un jugador por utilizar una de sus estrategias tomando en cuenta el comportamiento del oponente.

Por ejemplo, la función 1.4 surge de suponer que los errores tienen una distribución de probabilidad conocida como función generalizada de valor extremo, la cual debe ser idéntica para todas las alternativas. La diferencia de las utilidades entre dos alternativas cuyos errores se comportan de acuerdo con la distribución anterior, sigue una distribución de probabilidad denominada distribución logística. Esta es una de las funciones más utilizadas en la literatura experimental (*i.e.* Camerer y Ho, 1998; Fudenberg y Levine, 2009).

Una de las propiedades de esta función es que cuando el parámetro μ se aproxima a cero, la predicción del modelo se aproxima al equilibrio de Nash, mientras que, cuando el parámetro se aproxima a ∞ , la predicción del modelo es que los jugadores elegirán con igual probabilidad entre todas sus estrategias disponibles sin importar la función de pagos del juego.

El modelo de QRE ha sido uno de los más utilizados en la literatura experimental (*e.g.* Goeree, Holt y Palfrey, 2003; Selten y Chmura, 2012) debido a sus propiedades y a que su principal supuesto (la utilidad es una variable aleatoria) ha sido aplicado con éxito en el área de la elección individual.

Como se mencionó anteriormente, el modelo de QRE mantiene el supuesto de que los sujetos tienen información completa, pero sustituye el supuesto de maximización por una forma suavizada. En cambio, el modelo de PSE no utiliza el supuesto de información completa, sin embargo, mantiene el supuesto de que los jugadores buscan maximizar sus ganancias.

El modelo de PSE asume que para cada una de sus estrategias, el jugador selecciona una muestra de tamaño n de posibles juegos, después considera todos los posibles escenarios (elecciones del oponente) y cuenta, para cuántos de esos escenarios cada estrategia hubiera sido la mejor alternativa. La

probabilidad de elegir una estrategia será proporcional al número de veces en que cada estrategia hubiera sido la mejor en los n posibles juegos que el jugador consideró.

El proceso de muestreo que forma parte del modelo, no requiere que los jugadores tengan experiencia previa, pues supone que es un proceso hipotético en el cual solo se toma en cuenta, dadas las posibles combinaciones de estrategias que podría usar el oponente en una muestra de n juegos, qué proporción de veces sería mejor elegir la estrategia “a” sobre la estrategia “b”.

Formalmente, el modelo de PSE se expresa de la siguiente manera. Para un juego con dos estrategias, “Águila” y “Sol”, y dos jugadores; sean k_A^i , k_S^i el número de “Águilas” utilizadas por el oponente en una muestra de tamaño n para las estrategias “Águila” y “Sol” del jugador i (respectivamente). La suma de los pagos en la muestra para el jugador i se define como:

$$H_A^i = k_A^i h_i(A, A) + (n - k_A^i) h_i(A, S) \quad (1.5)$$

$$H_S^i = k_S^i h_i(S, A) + (n - k_S^i) h_i(S, S) \quad (1.6)$$

En donde $h_i(s_i, s_{-i})$ es el pago para el jugador i .

Con las sumas de los pagos es posible definir la probabilidad de que el

jugador i elija su estrategia “Águila” de la siguiente manera:

$$\beta_i(k_A^i, k_S^i) = \begin{cases} 1 & \text{si } H_A^i > H_S^i \\ \frac{1}{2} & \text{si } H_A^i = H_S^i \\ 0 & \text{de otra forma} \end{cases}$$

Debido a que k_A^i , k_S^i se distriubuyen de manera Binomial, es posible encontrar la proporción de ocaciones en las cuales la sumatoria de los pagos será mayor para cada estrategia dado un tamaño de la muestra n . El punto de equilibrio del modelo de PSE se obtiene resolviendo el siguiente sistema de ecuaciones:

$$p_A = \sum_{k_A^f=0}^n \sum_{k_S^f=0}^n \binom{n}{k_A^f} \binom{n}{k_S^f} q_A^{k_A^f+k_S^f} (1 - q_A)^{2n-k_A^f-k_S^f} \beta_f(k_A^f, k_S^f) \quad (1.7)$$

$$q_A = \sum_{k_A^c=0}^n \sum_{k_S^c=0}^n \binom{n}{k_A^c} \binom{n}{k_S^c} (1 - p_A)^{k_A^c+k_S^c} p_A^{2n-k_A^c-k_S^c} \beta_c(k_A^c, k_S^c) \quad (1.8)$$

En donde p_A y q_A representan la probabilidad de que el jugador Fila y Columna (respectivamente) elijan su estrategia “Águila”.

De acuerdo con el artículo original (Osborne y Rubinstein, 1998) existe una relación entre el parámetro n del modelo de PSE (*ver ecuaciones 1.7 y 1.8*) y el equilibrio de Nash. En juegos que tienen solución única en estrategias

mixtas, el equilibrio por muestreo de pagos converge con el NE cuando el tamaño de la muestra “ n ” tiende al infinito.

El modelo de IBE asume que el punto de equilibrio en un juego surge de un proceso dinámico, en el cual cada estrategia tiene asociada un impulso que es el resultado de la diferencia entre lo que un jugador obtuvo y lo que pudo haber obtenido, si su comportamiento hubiese sido diferente. Utilizando el ejemplo que emplean los autores (Selten y Chmura, 2008), si una persona intenta disparar a un blanco con un arco y la flecha golpea a la izquierda del centro, el modelo sugiere que en el siguiente ensayo el sujeto buscara corregir ese error apuntando un poco más a la derecha.

Del ejemplo anterior, se puede ver que el modelo predice cambios en la dirección de aquellas estrategias que pudieron resultar en una mayor ganancia para el sujeto. El equilibrio en el modelo se refiere al punto en el cual los impulsos para cada alternativa se igualan entre sí, tomando en cuenta el comportamiento del oponente. Este supuesto es la base de modelos que se revisarán más adelante concidos como modelos de arrepentimiento.

Selten y Chmura (2008) formalizaron el modelo y lo presentaron como una alternativa al NE para juegos 2x2. De manera formal, los impulsos para el jugador i desde su estrategia j hacia su estrategia k se definen de la siguiente forma:

$$r_k^i = \text{máx}(0, h_i(s_i^k, s_{-i}) - h_i(s_i^j, s_{-i})) \quad (1.9)$$

Debido a la estructura de los juegos utilizados en este experimento, estos impulsos sólo tendrán valores positivos para el caso de las combinaciones $\{(A, S), (S, A)\}$ para el jugador Fila y $\{(A, A), (S, S)\}$ para el jugador Co-

lumna. Estas combinaciones definen los impulsos de “A” a “S” y de “S” a “A”, respectivamente para cada jugador.

Una vez calculados los impulsos para cada estrategia, es posible encontrar el punto de equilibrio predicho por el modelo mediante la solución al siguiente sistema de ecuaciones, utilizando como ejemplo las estrategias en el presente experimento:

$$p_A(1 - q_A)r_S^f = (1 - p_A)q_Ar_A^f$$

$$q_Ap_Ar_S^c = (1 - q_A)(1 - p_A)r_A^c$$

Asumiendo que los impulsos son estrictamente positivos, el equilibrio se encuentra cuando se igualan los impulsos esperados.

Es posible definir la probabilidad de elegir la estrategia “Águila” de cada uno de los jugadores de la siguiente manera:

$$p_A = \frac{q_Ar_A^f}{q_Ar_A^f + (1 - q_A)r_S^f} \quad (1.10)$$

$$(1.11)$$

$$q_A = \frac{(1 - p_A)r_A^c}{(1 - p_A)r_A^c + p_Ar_S^c} \quad (1.12)$$

A diferencia de los dos modelos presentados anteriormente, el de IBE no utiliza parámetros libres.

Evaluación Empírica de Los Modelos de Equilibrio.

Estos cuatro modelos de equilibrio han sido comparados de manera expe-

rimental sólo en una ocasión (Selten y Chmura, 2008). Dicho experimento consistió de 12 diferentes juegos con 864 participantes divididos en 12 grupos (por juego) de 8 participantes cada uno para los primeros 6 juegos, y 6 grupos de 8 participantes cada uno para los juegos restantes. Para cada participante el juego se repitió en 200 ocasiones, en cada una de las repeticiones el jugador se enfrentaba contra un oponente de su grupo seleccionado al azar.

Los resultados del experimento de Selten y Chamura (2008), mostraron que de los 4 modelos, aquel que mejor se aproximaba al comportamiento de los participantes era el modelo de IBE, seguido por el modelo de PSE y después QRE y el equilibrio de Nash, en este orden. Sin embargo, Brunner, Camerer y Goeree, J. K. (2011) al realizar un nuevo análisis de los datos del experimento de Selten y Chmura, encontraron algunos errores en la estimación de los puntos de equilibrio y parámetros del modelo de QRE; con esto, se observó que, a diferencia de la conclusión de Selten y Chamura, en juegos de suma constante³ el modelo de QRE proporcionaba una mejor aproximación al comportamiento de los sujetos.

Los juegos utilizados en el presente experimento son una instancia de los juegos de suma constante en los cuales el modelo de QRE parece aproximarse mejor al comportamiento de los sujetos, sin embargo, los juegos de suma cero (que son una instancia de los juegos de suma constante) no fueron utilizados por Selten y Chamura (2008) por lo que es importante comparar estos modelos de equilibrio bajo estas condiciones de manera que sea posible evaluar el desempeño de los mismos en situaciones de conflicto puro.

Los modelos de equilibrio, en comparación con los modelos de aprendizaje,

³Juegos en los que la suma de cada una de las entradas de la matriz es igual a k .

tienen la ventaja de que no es necesario contar con la información respecto a los resultados de cada repetición de una situación de interacción, únicamente se requiere de la estructura del juego de manera que sea posible generar un conjunto de predicciones.

Para dar cuenta de las “trayectorias” que sigue el comportamiento de los sujetos cuando interactúan de manera repetida, es necesario realizar supuestos acerca de los procesos de aprendizaje, en particular, cómo los organismos adquieren información acerca de la situación y cómo dicha información se “traduce” en una elección.

1.4. Aprendizaje en Juegos

Los modelos de aprendizaje son modelos dinámicos que buscan dar cuenta del comportamiento de sujetos reales a través del tiempo cuando éstos se enfrentan en una situación de interacción. Para esto, los modelos hacen uso de supuestos acerca del tipo de información que adquieren los sujetos conforme experimentan con una situación determinada, así como el tipo de variable que se modifica con la experiencia.

Al clasificar los modelos en función de las variables que se modifican a través de la experiencia, existen dos grandes grupos; los primeros, son aquellos que suponen que lo que cambia son las preferencias (atracciones o fuerza) de los jugadores con respecto a sus estrategias, este tipo de modelos se denominan modelos de refuerzo. El segundo tipo son aquellos en los que se asume que los sujetos forman un “modelo” acerca del comportamiento de los oponentes y eligen con base en dicho modelo, en este caso, la experiencia

no modifica la preferencia, en cambio, es el comportamiento del oponente el que guía las elecciones de los participantes, estos modelos se conocen como modelos bayesianos.

Dentro del primer grupo, existen variantes las cuales pueden asumir que lo que altera las preferencias de los sujetos son los pagos recibidos o en otros casos lo que cambia las preferencias son las ganancias que los jugadores pudieron haber recibido, los cuales se denominan modelos de “arrepentimiento”, sin embargo, forman parte de la familia de modelos de reforzamiento.

De los modelos evaluados en el presente experimento, dos corresponden a esta categoría: Reforzamiento Acumulado (*Cummulative Reinforcement*; CR, por sus siglas en inglés, Erev y Roth, 1998) y el Aprendizaje por Igualación de Impulsos (*Impulse Matching*; IM, por sus siglas en inglés; Chmura, Goerg y Selten, 2012). Este último pertenece a la segunda variante de los modelos de reforzamiento ya que utiliza los pagos que los jugadores pudieron haber recibido.

El segundo tipo de modelos denominados bayesianos, asumen que el comportamiento de los oponentes se refleja en una creencia por parte de los jugadores acerca de la distribución de probabilidad que los oponentes utilizan para llevar a cabo sus elecciones, estas creencias o distribución de probabilidad son a las que se les denomina “modelo” (Young, 2007).

De los modelos utilizados en el presente experimento, solamente el de Juego Ficticio Estocástico (*Stochastic Fictitious Play*; SFP, por sus siglas en inglés; Fudenberg y Kreps, 1993) corresponde a esta familia. En 1993 Fudenberg y Kreps demostraron que el modelo de SFP converge en el equilibrio de Nash en juegos 2x2, y además, que el modelo formado a través de la regla de

aprendizaje converge en el comportamiento real del oponente, sin embargo, esto sólo ocurre cuando los oponentes utilizan una distribución de probabilidad estacionaria, es decir, si el comportamiento de los oponentes no cambia a través del tiempo.

Por último, el modelo Auto-Regulable de Atracciones Ponderadas por la Experiencia (*Self-tuning Experience Weighted Attraction Learning*; St-EWA, por sus siglas en inglés; Ho, Camerer y Chong, 2007) puede ser considerado como un híbrido entre los modelos de reforzamiento y los modelos bayesianos, ya que utiliza un modelo (creencias) acerca del comportamiento de los oponentes además de los pagos asociados a cada estrategia para modificar las preferencias de los jugadores.

A continuación, se revisarán los modelos mencionados junto con algunos de los resultados experimentales que apoyan a cada uno, así como los resultados del contraste entre los mismos.

1.4.1. Aprendizaje por Reforzamiento

Los modelos de reforzamiento fueron uno de los primeros modelos de aprendizaje utilizados en la Economía (Young, 2004) debido a que, en su forma más general, asumen que los sujetos solo requieren información acerca del desempeño pasado de sus acciones para tomar una decisión, lo cual los convierte en uno de los modelos más simples.

De la familia de modelos de Reforzamiento en juegos el que ganó mayor atención y cuyas propiedades a largo plazo han sido más estudiadas, es el propuesto por Erev y Roth (1998) conocido como Reforzamiento Acumulado

(CR). Este modelo se basa en dos principios generales derivados de la investigación en aprendizaje individual que son: la ley de poder de la práctica y una extensión a la ley del efecto que supone que la elección es probabilística (Erev y Roth, 1998).

La ley del efecto, de acuerdo con el modelo de CR, implica que las acciones o estrategias que han resultado en una ganancia en el pasado se usarán con mayor probabilidad en el futuro. Mientras que la ley de poder de la práctica, hace referencia a una reducción en el impacto que tienen los pagos a medida que transcurren los ensayos, es decir, el impacto que tienen las ganancias sobre la probabilidad de elegir alguna de las estrategias, se reduce conforme los jugadores adquieren experiencia.

Por último, el supuesto de que la elección es probabilística se refleja en el uso de una función de respuesta probabilística, no obstante, cuando existen pagos negativos en un juego ésta puede quedar indeterminada. Por lo tanto, es necesario transformar las matrices de manera que todos los pagos sean mayores o iguales a cero, sin alterar el equilibrio de Nash del juego.

Como se mencionó al inicio del capítulo, el principal supuesto en los modelos de reforzamiento es que la experiencia modifica las preferencias que los jugadores tienen respecto a cada una de sus estrategias. En la literatura en aprendizaje en juegos, las preferencias han sido denominadas atracciones, sin embargo, vale la pena destacar que este tipo de modelos han sido utilizados en otras áreas en las cuales pueden tomar otro tipo de nombres (*e.g.* fuerza, valor, etc).

De acuerdo con el modelo de CR, las atracciones de cada jugador con respecto a cada una de sus estrategias se modifican de acuerdo con la siguiente

función:

$$A_i^j(t) = A_i^j(t-1) + \begin{cases} h_i(s_i^j, s_{-i}) & \text{si } s_i^j = s_i(t) \\ 0 & \text{en otro caso} \end{cases}$$

En donde $A_i^j(t)$ representa la atracción o preferencia del jugador i por su estrategia j y $h_i(s_i^j, s_{-i})$ representa el pago para el jugador en el ensayo por haber usado su estrategia j dado el comportamiento del oponente. Esta función refleja el supuesto de la ley del efecto, es decir, que las consecuencias de las elecciones son las que modifican la atracción (fuerza) de cada una de las alternativas.

De manera que sea posible asignarle una probabilidad a cada una de las estrategias del jugador en el primer ensayo, el modelo requiere de un valor inicial para las atracciones. Este valor debe cumplir con dos restricciones:

- i) La atracción inicial de un jugador por cada una de sus estrategias es la misma, es decir: $A_i^j(0) = A_i^k(0)$.
- ii) Toda atracción inicial debe ser mayor a 0: $A_i^j(0) > 0; \forall j$

La probabilidad de que el jugador elija su estrategia j se define como:

$$p_i^j(t) = \frac{A_i^j(t)}{\sum_{k \in s_i} A_i^k(t)} \quad (1.13)$$

Por lo tanto, la probabilidad de utilizar la estrategia j en el ensayo t es una función de la tasa relativa de las atracciones. Si la sumatoria de los pagos recibidos por un jugador a lo largo del tiempo tuviera un valor de 0, esta ecuación podría quedar indeterminada, debido a esto, es necesario

transformar los pagos de los juegos del presente experimento modificando la función de reforzamiento de la siguiente manera:

$$R_i^j(s_i^j, s_{-i}) = h_i(s_i^j, s_{-i}) + |\text{mín } h(\bullet, \bullet)| \quad (1.14)$$

En donde $\text{mín } h(\bullet, \bullet)$ representa el pago mínimo en el juego tomando en cuenta a todos los jugadores; de esta forma R_i^j reemplaza a $h_i(s_i^j(t), s_{-i}(t))$ en la ecuación para la actualización de las atracciones:

$$A_i^j(t) = A_i^j(t-1) + \begin{cases} R_i^j(s_i^j, s_{-i}) & \text{si } s_i^j = s_i(t) \\ 0 & \text{en otro caso} \end{cases}$$

La función 1.13 es la que refleja el supuesto asociado a la ley de poder de la práctica, en este caso, conforme transcurren los ensayos el valor de la suma de las atracciones de los jugadores por sus estrategias tiende a acumularse, por lo tanto, el efecto que las consecuencias tienen sobre la probabilidad de elegir cada estrategia decrementa conforme pasa el tiempo.

El modelo de CR fue de los primeros modelos de reforzamiento utilizados en la TJ (Erev y Roth, 1998); sin embargo, existen varias extensiones del mismo, las cuales incluyen supuestos acerca de que las preferencias de los jugadores pueden cambiar cuando una estrategia no ha sido utilizada o que los pagos que los jugadores no recibieron pueden afectar la preferencia. No obstante, este último supuesto pertenece a la variante de los modelos de “arrepentimiento”.

A continuación, se presentara el modelo de IM (Chmura, Goerg y Selten, 2012), el cual pertenece a la variante de los modelos de “arrepentimiento”, de los cuatro modelos utilizados, éste es el único que no utiliza parámetros libres.

1.4.2. Aprendizaje por Igualación de Impulsos

EL modelo de IM fue propuesto por Chmura, Goerg y Selten (2012) y, al igual que el de CR, forma parte de la familia de los modelos de reforzamiento. Este modelo es la formalización del proceso de aprendizaje que subyace al *Equilibrio por Balance de Impulsos* (IBE) propuesto por uno de los autores en 2008 (Selten, Abbkin y Cox, 2008).

En este modelo de aprendizaje, se asume que lo que determina las preferencias de los jugadores por sus estrategias son los pagos que pudieron haber recibido si su comportamiento hubiera sido diferente; es decir, las preferencias cambian en favor de aquellas estrategias que hubieran resultado mejores dado el comportamiento del oponente. Debido a esto, el modelo pertenece a la variante de los modelos de reforzamiento denominada modelos de “arrepentimiento”.

En este modelo, las preferencias se definen a través de lo que los autores denominan impulsos (Chmura, Goerg y Selten, 2012). La diferencia con las atracciones utilizadas en el modelo anterior, es que los impulsos dependen únicamente de los pagos que los jugadores pudieron haber recibido (cuando dichos pagos son positivos) y no de los pagos recibidos en sí.

El impulso para el jugador i desde su estrategia k hacia su estrategia j

se define de la siguiente manera:

$$r_i^j(t) = \text{máx}(0, \pi_i(s_i^j, s_{-i}(t)) - \pi_i(s_i^k, s_{-i}(t))) \quad (1.15)$$

En donde $\pi_i(s_i^j, s_{-i}(t))$ representa el pago transformado para el jugador i por utilizar su estrategia j dado el comportamiento del oponente en el ensayo t . Al igual que en el modelo de CR, una de las restricciones sobre los impulsos es que éstos deben ser siempre positivos.

De acuerdo con la ecuación anterior, los impulsos o preferencias cambian en favor de una de las estrategias, únicamente cuando el pago por haber usado esa estrategia, es mayor al pago por haber usado la otra, tomando en cuenta la elección del oponente en el ensayo t .

Usando como ejemplo el juego “A1” en el presente experimento, si el jugador Fila eligiera su estrategia “Águila” y el oponente su estrategia “Sol”, la preferencia del jugador Fila por su estrategia “Sol” se incrementaría debido a que si hubiera usado esa estrategia podría haber ganado 2 puntos en lugar de perderlos. En cambio, en el mismo ensayo los impulsos para el jugador Columna serían iguales a 0 ya que ese jugador no habría podido ganar más si su comportamiento fuera diferente.

De manera similar al modelo de CR, el de IM asume que los jugadores acumulan la información acerca de los impulsos en cada ensayo hacia cada una de sus estrategia. Los impulsos acumulados hasta el ensayo t , se calculan de la siguiente manera:

$$R_i^j(t) = R_i^j(t - 1) + r_i^j(t) \quad (1.16)$$

Debido a que los impulsos son siempre positivos, este modelo también presenta un efecto similar a la ley de poder de la práctica, es decir, el impacto de cada nuevo impulso sobre la probabilidad de elección de los jugadores se reduce conforme pasan los ensayos.

De acuerdo con el modelo, la probabilidad de que cada jugador elija una de sus estrategias es una función de la tasa relativa de los impulsos acumulados:

$$p_i^j(t) = \frac{R_i^j(t)}{\sum_{k \in S_i} R_i^k(t)} \quad (1.17)$$

Esta es la misma función que utiliza el modelo de CR, sin embargo, la diferencia radica en el tipo de información que utiliza el modelo; mientras que en el primero se utilizan los pagos recibidos, en éste se utilizan los pagos que no se obtuvieron y que hubieran resultado mejores (mayores a 0).

Debido a que este modelo no utiliza parámetros libres, es considerado como uno de los modelos de aprendizaje más simples.

Juntos, el modelo de CR y el de IM, son los únicos modelos pertenecientes a la familia de los modelos de reforzamiento que se compararon en el presente experimento.

De los modelos que se presentan a continuación, uno pertenece a la familia de modelos bayesianos mientras que el segundo es un híbrido entre esta familia de modelos y los de reforzamiento.

1.4.3. Juego Ficticio Estocástico

El Juego Ficticio Estocástico (SFP) pertenece a la familia de modelos denominados modelos bayesianos, en el cual se asume que las preferencias de los

jugadores no cambian a través del tiempo, sino que los jugadores forman un “modelo” acerca del comportamiento de sus oponentes el cual utilizan para elegir entre sus alternativas.

En el modelo de SFP, se asume que los jugadores utilizan las elecciones de sus oponentes hasta el último ensayo para predecir con qué probabilidad elegirán sus estrategias en ensayos subsecuentes (Fudenberg y Kreps, 1993); para después calcular un valor esperado para cada estrategia que se basa en dicha probabilidad. En otras palabras, de acuerdo con el modelo de SFP, los jugadores asumen que sus oponentes utilizan una distribución de probabilidad sobre sus estrategias, la cual puede inferirse a partir de la frecuencia relativa con la que han utilizado cada estrategia hasta el ensayo actual, además de una creencia previa acerca de la probabilidad con la que sus oponentes elegirán en el primer ensayo (Fudenberg y Levine, 2009).

Una de las características de este modelo es que supone que los jugadores no esperan que sus oponentes cambien su comportamiento en los siguientes ensayos, es decir, asumen que la distribución de probabilidad que utilizan los oponentes para llevar a cabo sus elecciones es estacionaria⁴ (Fudenberg y Levine, 2003).

Para construir una creencia o modelo acerca del comportamiento del oponente, el modelo de SFP asume que los jugadores realizan conteos acerca de las veces en las que el oponente ha usado cada una de sus estrategias. En este caso, el conteo que el jugador i realiza sobre la estrategia j de su oponente

⁴No cambia a través del tiempo.

hasta el ensayo t es:

$$k_i^j(t) = k_i^j(t-1) + \begin{cases} 1 & \text{si } s_{-i}(t-1) = j \\ 0 & \text{en cualquier otro caso} \end{cases}$$

En donde $k_i^j(t)$ representa el conteo que el jugador i le asigna a la estrategia j del oponente hasta el ensayo t .

Mediante estos conteos el jugador i infiere la probabilidad de que su oponente elija su estrategia j en el ensayo t de acuerdo con la siguiente ecuación:

$$\gamma_i^j(t) = \frac{k_i^j(t)}{\sum_{k \in S_{-i}} k_i^k(t)} \quad (1.18)$$

El “modelo” (o creencias) que los jugadores tienen acerca del comportamiento del oponente se conforma de un valor de γ_i^j para cada una de las estrategias que dicho oponente tiene disponibles y puede cambiar a través del tiempo.

Utilizando las creencias formadas a partir de las dos ecuaciones anteriores, se puede definir la utilidad o valor esperado para cada una de las estrategias del jugador. La utilidad esperada de la estrategia r del jugador i en el ensayo t se define como:

$$\bar{u}(s_i = r | \gamma_i(t)) = \sum_{j \in S_{-i}} \gamma_i^j(t) * h_i(s_i = r, s_{-i} = j) \quad (1.19)$$

En donde, $\gamma_i^j(t)$ representa la probabilidad estimada por el jugador i de que el oponente elija su j -ésima estrategia y $h_i(s_i = r, s_{-i} = j)$ representa el pago para el jugador i por usar su estrategia r tomando en cuenta que el

oponente elige su estrategia j .

Al igual que el modelo de QRE, en este modelo la probabilidad de cada estrategia depende de la utilidad esperada, sin embargo, en el modelo de SFP, ésta depende del modelo que el jugador ha construido acerca del comportamiento de su oponente.

El modelo de SFP utiliza la función logística que se presentó anteriormente para ligar la utilidad esperada con una probabilidad de elección por parte de los jugadores (Fudenberg y Levine, 2009); la probabilidad de que el jugador i elija su estrategia k en cualquier ensayo se define como:

$$p_i^k(t) = \frac{e^{\bar{u}(s_i=k|\gamma_i(t))/\mu}}{\sum_{r \in S_i} e^{\bar{u}(s_i=r|\gamma_i(t))/\mu}} \quad (1.20)$$

En donde, $\bar{u}(s_i = k | \gamma_i(t))$ representa la utilidad esperada del jugador i por utilizar su estrategia j condicional al modelo acerca del comportamiento del oponente $\gamma_i(t)$ generado a través de la función 1.18 .

A diferencia de los tres modelos descritos para los cuales es posible definir a qué familia pertenece cada uno, el modelo de St-EWA (Ho, Camerer y Chong, 2007) incorpora supuestos que pertenecen a ambas familias de modelos, es decir, toma supuestos de los modelos de reforzamiento como los cambios en las preferencias, pero también toma en cuenta el efecto que tienen las elecciones de los oponentes sobre el comportamiento de los jugadores. Por lo tanto, este modelo podría ser considerado como un “híbrido” entre los modelos bayesianos y los modelos de reforzamiento (Camerer y Ho, 1999).

1.4.4. St-EWA

El modelo de St-EWA propuesto por Ho, Camerer y Chong (2007), deriva de un modelo presentado anteriormente por Ho y Camerer (1999), denominado “Atracciones Ponderadas por la Experiencia” (*Experience Weighted Attraction Learning*, EWA; por sus siglas en inglés). Con el modelo original, los autores mostraron que una instancia de la familia de modelos bayesianos y los modelos de reforzamiento podían entenderse como instancias de un modelo más general.

Sin embargo, el modelo original recibió críticas debido a que contaba con una gran cantidad de parámetros (7 para un juego de dos jugadores, en un juego 2 x 2). Debido a esto, Ho, Camerer y Chong (2007) presentaron el modelo de St-EWA, el cual reemplaza la mayoría de los parámetros con los que contaba el modelo original con funciones deterministas, basadas en supuestos respecto a cómo los sujetos utilizan la información que adquieren ensayo a ensayo para definir su preferencia con respecto a las estrategias disponibles. Estos supuestos son:

- i) La preferencia de los jugadores es la misma para todas las alternativas al inicio del juego.
- ii) Al inicio del juego los jugadores no cuentan con experiencia previa.
- iii) Las preferencias de los jugadores cambian en función de las consecuencias de cada elección.
- iv) Las preferencias de los jugadores cambian en función de las ganancias que pudieron haber obtenido si su elección hubiera sido diferente.

- v) Los sujetos ponderan los resultados previos de cada estrategia en función de la variabilidad en el comportamiento del oponente a lo largo del tiempo.

Estos 5 supuestos, son una combinación de los modelos que han sido presentados anteriormente, sin embargo, algunos pertenecen a la familia de modelos de reforzamiento, mientras que otros pertenecen a los modelos bayesianos.

El primer supuesto implica que las atracciones iniciales (preferencia del sujeto al inicio del juego) deben tener un valor constante k , el cual es el mismo para todos los sujetos y estrategias; sea A_i^j la atracción del sujeto i hacia su estrategia j , las atracciones iniciales se definen como:

$$A_i^j(t = 0) = k ; \forall j \in S_i$$

En otros experimentos se ha asignado un valor de 1 para k , esto debido a que es posible asumir que los jugadores no tienen una preferencia definida al iniciar el juego cuando no conocen los pagos.

Por otra parte, la experiencia de los jugadores se define como el número de ocasiones en las cuales un jugador se ha enfrentado a un juego en particular. Esta experiencia modifica el impacto que tienen otras variables sobre la probabilidad de elección de los jugadores. En general, tiene un efecto similar a la ley de poder de la práctica que se mencionó anteriormente (*ver sección 1.4.1*), sin embargo, en el caso de este modelo en particular, la reducción en el impacto que tiene cada nuevo ensayo depende también de la variabilidad en el comportamiento del oponente.

La experiencia de un jugador en el ensayo t se define como:

$$N_i(t) = \begin{cases} 1 & \text{si } t = 0 \\ \phi_i(t)N_i(t-1) + 1 & \text{en otro caso} \end{cases}$$

En donde, $N_i(t)$ representa la experiencia del jugador i hasta el ensayo t y $\phi_i(t)$ es un factor de descuento que depende del comportamiento del oponente. Conforme más “estables” son las elecciones del oponente, la experiencia de los jugadores se descuenta menos y, por lo tanto, menor es el impacto de un nuevo ensayo sobre las probabilidades de elección de los jugadores.

De las funciones que forman parte del modelo de St-EWA, las que adquieren mayor relevancia en este experimento podrían ser las que definen a la variable $\phi_i(t)$; el valor de esta variable se conforma de 4 ecuaciones las cuales dependen del comportamiento del oponente.

De acuerdo con el modelo, el efecto acumulado de los resultados de cada estrategia sobre las atracciones o preferencias de cada jugador, decae a lo largo del tiempo, no obstante, la velocidad a la que esto sucede depende de la variable $\phi_i(t)$, la cual está asociada a la historia de juego del oponente. Por ejemplo, conforme el comportamiento del oponente se hace más variable, menor impacto tienen los resultados anteriores sobre las preferencias de un jugador; en cambio, mientras más estable se hace el comportamiento del oponente mayor impacto tienen los resultados anteriores.

Las funciones que definen la variable $\phi_i(t)$ son 4 y en un juego de dos jugadores 2x2 con equilibrio en estrategias mixtas⁵, se definen de la siguiente

⁵La función 1.21 es diferente cuando los juegos tienen una solución en estrategias puras.

manera:

$$r_k^{s^{-i}}(t) = \frac{\sum_{\tau=t-w+1}^t I(s_k^{-i}, s^{-i}(\tau))}{2} \quad (1.21)$$

$$h_k^{s^{-i}}(t) = \frac{\sum_{\tau=1}^t I(s_k^{-i}, s^{-i}(\tau))}{t} \quad (1.22)$$

$$S_i(t) = \sum_{k=1}^2 (h_k^{s^{-i}} - r_k^{s^{-i}})^2 \quad (1.23)$$

$$\phi_i(t) = 1 - \frac{1}{2} S_i(t) \quad (1.24)$$

Estas 4 ecuaciones determinan el valor de $\phi_i(t)$ para cada jugador en cada ensayo. La variable $S_i(t)$ que aparece en las ecuaciones 1.24 y 1.23 se denomina índice de “*sorpresa*” y funciona como una medida de la diferencia entre el número total de veces en que el oponente ha utilizado su estrategia k a lo largo del tiempo, y el número de veces en las que el oponente ha utilizado la misma estrategia en los últimos ensayos $r_k^{-i}(t)$. Por lo tanto, esta variable funciona como una medida de la variabilidad en el comportamiento del oponente (Chmura, Goerg y Selten, 2012).

La función 1.22 es la que utiliza el modelo de St-EWA para generar un modelo acerca del comportamiento del oponente a lo largo del tiempo, esta función es la misma que utiliza el modelo de SFP, por lo tanto, incluye el mismo supuesto, es decir, que la probabilidad con la que el oponente elige entre sus estrategias es estacionaria. En cambio, la función 1.21 tiene una forma similar, sin embargo, ésta se utiliza para inferir la probabilidad con la que el oponente ha elegido entre sus estrategias en los últimos w ensayos⁶.

⁶La variable w en la ecuación 1.21, representa el número de estrategias que reciben una

En la función 1.24, el índice de sorpresa aparece ponderado por un medio, debido a que esta variable puede tomar valores entre cero y dos, lo cual podría llevar a ϕ a valores negativos.

En conjunto, estas funciones han sido denominadas “*detector de cambio*” a causa de su comportamiento. Por ejemplo, cuando los oponentes han jugado una sola estrategia durante los t ensayos anteriores, el valor de la variable ϕ es de uno, lo cual implica que las atracciones anteriores tienen el mismo peso que el pago del último ensayo. Esta propiedad de la función es importante debido a que en ambientes que cambian poco, los resultados de ensayos anteriores (resumidos en las atracciones) brindan información acerca de los resultados que pueden esperarse en el futuro y, por lo tanto, es deseable que determinen en mayor medida la probabilidad de utilizar una estrategia (Ho, Camerer y Chong, 2007).

Por otra parte, cuando el comportamiento de los oponentes cambia repentinamente, los resultados de ensayos anteriores dejan de ser un buen indicador de lo que es posible esperar del oponente, por lo tanto, las atracciones que contienen información acerca de los ensayos anteriores, deberían de decrementar su impacto sobre la elección, otorgando un mayor peso a los pagos recibidos y los pagos que pudieron resultar mejores dado el comportamiento del oponente en los últimos ensayos (Ho, Camerer y Chong, 2007). Esto es lo que sucede cuando se utilizan las funciones anteriores para determinar el decaimiento de las atracciones. Por ejemplo, si el oponente ha utilizado una estrategia pura en los últimos $t - 2$ ensayos la cual cambia en los últimos dos ($t - 1$ y t), el valor de $\phi_i(t)$ se aproximará a cero, descontando por completo

probabilidad mayor a cero en equilibrio

las atracciones acumuladas.

En su conjunto, las ecuaciones 1.21, 1.22, 1.23 y 1.24, son las que relacionan al modelo de St-EWA con la familia de modelos Bayesianos, ya que incluyen la construcción de un “modelo” acerca del comportamiento del oponente en el proceso de la formación de las preferencias (atracciones) de los jugadores.

Por último, el modelo cuenta con una variable denominada δ , la cual cumple la función de ponderar los pagos que los jugadores pudieron haber recibido si su elección en el ensayo hubiera sido diferente, manteniendo constante el comportamiento del oponente. Como se mencionó anteriormente, este supuesto forma parte de los modelos de arrepentimiento a los cuales pertenece el modelo de IM, presentado en la sección 1.4.2.

La función propuesta para el modelo de St-EWA, funciona de manera similar a la del modelo de IM, es decir, el valor que se obtiene de esta función siempre es positivo, sin embargo, en el modelo de St-EWA se asume que los pagos que no se recibieron tienen un impacto menor al de su valor original sobre las preferencias. La función que utiliza el modelo es:

$$\delta_i(t) = \begin{cases} \frac{1}{w} & \text{si } h(s_i, s_{-i}(t)) \geq h(t) \\ 0 & \text{en cualquier otro caso} \end{cases}$$

De acuerdo con esta ecuación, sólo los pagos que sean mayores o iguales al del pago en el último ensayo modifican las preferencias del jugador, esto en una magnitud de $1/w$ veces de su valor original⁷.

⁷Al igual que en la función anterior (ver ecuación 1.21), w representa el número de

Con estas variables definidas, las atracciones o preferencias de los sujetos por cada una de sus alternativas se actualizan en cada ensayo de acuerdo con la siguiente función:

$$A_i^j(t) = \frac{\phi_i(t)A_i^j(t-1)N_i(t-1)}{N_i(t)} + \frac{[\delta_i(t) + (1-\delta_i(t))I(s_i^j(t), s_i(t))]h(s_i^j, s_{-i}(t))}{N_i(t)} \quad (1.25)$$

En donde, $A_i^j(t)$ representa la atracción del jugador i por su estrategia j en el ensayo t .

Las atracciones de los sujetos son las que definen la probabilidad de elegir una de las alternativas en cada ensayo. La probabilidad de que el jugador i elija su estrategia j en el ensayo $t + 1$ es:

$$p_i^j(t+1) = \frac{e^{\frac{A_i^j(t)}{\mu}}}{\sum_{k \in S_i} e^{\frac{A_i^k(t)}{\mu}}} \quad (1.26)$$

Esta es la misma función que relaciona la utilidad esperada con una probabilidad de elegir cada estrategia en el modelo de Equilibrio en Respuestas Cuantales (McKelvey y Palfrey, 1995) y en el modelo de SFP (Fudenberg y Kreps, 1993).

De los modelos que han sido presentados, ninguno ha sido diseñado para dar cuenta del comportamiento de los jugadores en juegos en los cuales la función de pagos puede modificarse a través del tiempo, sin embargo, los modelos que suponen que los jugadores basan sus elecciones en el compor-

estrategias que reciben una probabilidad mayor a cero en equilibrio

tamiento del oponente podrían ser suficientes para lidiar con este tipo de cambios, debido a que éste puede ser de mayor importancia en una situación de interacción.

No obstante, vale la pena destacar que en el modelo de St-EWA, si bien no está diseñado para lidiar con este tipo de ambientes, las funciones asociadas al factor de descuento permiten que regule el impacto de la información pasada sobre las preferencias en función del cambio en el comportamiento del oponente.

Hasta este punto únicamente se han tratado algunos de los aspectos formales de los modelos y sus supuestos, sin embargo, estos modelos han sido probados experimentalmente en diversas ocasiones; en algunas se ha tratado de verificar si sus supuestos se relacionan con el comportamiento de los participantes, mientras que en otras se ha comparado su desempeño de manera estadística.

A continuación, se presentaran algunos de los resultados experimentales asociados a estos modelos.

Una de las consecuencias de los modelos de reforzamiento en los cuales la probabilidad de utilizar una estrategia depende del desempeño de la misma, es que se presente una correlación positiva⁸ entre las estrategias que utiliza un jugador a lo largo de los ensayos.

En diversos estudios se ha mostrado que las elecciones de un sujeto a través del tiempo presentan este tipo de correlación (*e.g.* Brown y Rosenthal, 1990; Shachat, 2002). Incluso antes de que Erev y Roth presentaran su modelo, Mookherjee y Shoper (1994), ya habían mostrado que la presentación

⁸Cuando la estrategia utilizada lleva a una ganancia y no a una pérdida.

de información respecto a las ganancias de cada estrategia, modificaba el comportamiento de los jugadores. Este tipo de hallazgos apoyan el supuesto principal de los modelos de reforzamiento, es decir, que las consecuencias de las elecciones de un jugador afectan su comportamiento en los ensayos subsecuentes.

No obstante, en el mismo experimento Mookherjee y Shoper (1994) mostraron que el comportamiento de los jugadores es diferente cuando reciben información acerca del comportamiento de sus oponentes en cada ensayo, lo cual no puede ser explicado por los modelos de reforzamiento.

Una de las consecuencias del modelo de SFP, es que los jugadores reales deberían de comportarse de manera diferente cuando reciben información acerca del comportamiento de su oponente. Esto fue exactamente lo que ocurrió en el experimento realizado por Mookherjee y Shoper (1994). El comportamiento del grupo de jugadores que recibía información acerca del comportamiento de sus oponentes fue diferente al del grupo sin información. Esto sugiere que jugadores humanos son capaces de utilizar la información acerca del comportamiento de otros para llevar a cabo una elección.

Las comparaciones de los modelos de aprendizaje por reforzamiento y los modelos bayesianos, utilizando métodos estadísticos que miden la calidad del ajuste de un modelo a los datos, han mostrado que cuando se utilizan juegos con un equilibrio en estrategias mixtas, los modelos de reforzamiento tienen un mejor ajuste a los datos comparados con los modelos bayesianos. Sin embargo, las diferencias no son grandes, debido a que, por lo general, los resultados que predicen ambos tipos de modelos son bastante similares (Camerer, 2003).

En cuanto a la comparación experimental, no todos los modelos presentados en esta sección han recibido el mismo nivel de atención. Por ejemplo, los modelos de CR y SFP al ser de los primeros en ser propuestos, han sido contrastados utilizando distintos tipos de juegos (*e.g.* Erev y Roth, 1998; Erev, *et. al.*, 2007; Ho, Wang y Camerer, 2008); mientras que el modelo de St-EWA ha sido comparado en menor número de ocasiones (*e.g.* Ho, Camerer y Chong, 2007), por último, el modelo de IM ha sido comparado con los demás sólo en una ocasión (Chmura, Goerg y Selten, 2012).

En su artículo de 2012 Chmura, Goerg y Selten contrastaron el modelo de IM con el de CR (Erev y Roth 1998) y con St-EWA (Ho, Camerer y Chong, 2007). Para comparar estos modelos, los autores utilizaron los datos del experimento de Chmura y Selten de 2008 (ver sección 4, página 12). En dicho estudio, los autores utilizaron dos métodos diferentes para realizar el contraste.

En primer lugar, simularon el comportamiento de cada uno de los modelos para cada juego, después contrastaron el punto estacionario al que llegaban las simulaciones contra la frecuencia relativa con la que los participantes utilizaron cada estrategia. Para contrastar estos dos puntos, los autores utilizaron el estadístico Q presentado anteriormente (ver página 64, ecuación 1.29). Mediante este método, encontraron que los modelos podían ser ordenados de menor a mayor distancia (Q) de la siguiente manera:

- i) Aprendizaje por Igualación de Impulsos (IM).
- ii) Modelo Auto-regulable de Atracciones Ponderadas por la Experiencia (St-EWA).

iii) Reforzamiento Acumulado (CR).

En segundo lugar, utilizaron un estadístico presentado por Brier (1950). Este estadístico se denomina q y busca determinar la congruencia entre una predicción realizada mediante una variable continua (probabilidad) y un evento discreto, a esta función se le conoce como regla de puntuación cuadrática y se define de la siguiente manera:

$$q = \frac{1}{n} \sum_{j=1}^r \sum_{i=1}^t (f_{ij} - E_{ij})^2 \quad (1.27)$$

En donde f_{ij} es la probabilidad que un modelo de asigna al evento j en el ensayo i y la variable E_{ij} toma el valor de uno si ocurrió el evento j en el i –ésimo ensayo y cero en cualquier otro caso.

Selten (1998) mostró que dicho estadístico se reduce a la siguiente función:

$$q(t) = 2\theta_i(t) - \theta_i^2(t) - (1 - \theta_i(t))^2 \quad (1.28)$$

En donde $\theta_i(t)$, representa la probabilidad predicha por el modelo en el ensayo t para la opción que el jugador eligió. Este estadístico muestra la congruencia entre la predicción del modelo y el evento para cada uno de los ensayos, por lo que Chmura, Goerg y Selten (2012) lo utilizaron como criterio para contrastar el desempeño, ensayo a ensayo, de los modelos. El resultado de este contraste reacomoda a dos de los tres modelos. En este caso, el modelo que mejor describió el comportamiento de los sujetos por ensayo fue St-EWA, dejando en segundo lugar a IM.

Uno de los principales problemas de la comparación entre modelos de

aprendizaje en juegos, es que los modelos de reforzamiento y los bayesianos generan predicciones muy similares, lo cual dificulta la discriminación entre ambas familias. Además, existen experimentos que muestran que modelos que utilizan supuestos provenientes de ambas familias, superan a los originales en términos de su calidad en el ajuste a los datos.

Debido a que los modelos de St-EWA (Ho, Camerer y Chong, 2007) e IM (Chmura, Goerg y Selten, 2012) son relativamente nuevos, puede resultar de utilidad comenzar a contrastarlos en distintas situaciones a las que han sido utilizados en el pasado, esto debido a que no tienen el mismo sustento que aquellos modelos que llevan un mayor tiempo en la literatura. Ya que los supuestos que utilizan ambos modelos son muy diferentes de aquéllos que ya están establecidos, contrastarlos con éstos podría arrojar luz en términos de cuáles podrían ser los mecanismos que subyacen al uso de la información en situaciones estratégicas.

El principal objetivo del presente experimento, es el evaluar el comportamiento de los participantes en situaciones de interacción (juegos) cuando las consecuencias de sus elecciones cambian a lo largo del tiempo.

Aunado a esto, se contrastaron 4 modelos de equilibrio que difieren respecto a los supuestos que realizan acerca de cuáles son las variables que guían la elección de los participantes. El objetivo de esta comparación, fue evaluar la capacidad de los modelos para dar cuenta del comportamiento de los participantes cuando únicamente se utilizan las características del juego, sin considerar que los participantes se encuentran en un entorno que cambia o que su comportamiento cambia a través del tiempo.

Por último, se contrastaron 4 modelos de aprendizaje pertenecientes a las

familias de los modelos de reforzamiento y bayesianos. El método implementado en el presente experimento ha mostrado ser de utilidad al comparar este tipo de modelos debido a que las predicciones de ambos pueden diferir en términos de su respuesta a un cambio.

Por ejemplo, en el presente experimento, los modelos de reforzamiento predecían que los sujetos cambiarían su comportamiento de manera gradual después de presentarse un cambio en las consecuencias; por otra parte, los modelos bayesianos predicen que el cambio en el comportamiento de los participantes es abrupto.

Método

Los participantes fueron 70 estudiantes de licenciatura de la Facultad de Psicología de la UNAM de entre 18 y 25 años quienes participaron de manera voluntaria en el estudio. Por su participación a todos los sujetos se les ofreció una recompensa monetaria, sin embargo, algunos de los participantes podían cambiar esta recompensa por un punto en un examen parcial.

Los juegos utilizados en el presente experimento, son una versión generalizada del juego de MP. Esta generalización se realizó alterando los pagos de una de las combinaciones de estrategias, sin embargo, las estrategias disponibles para los jugadores siempre fueron las mismas (“Águila” o “Sol”). La relaciones generales de los pagos presentes en el juego original de MP, se mantuvieron constantes aún después de un cambio¹. Durante la fase experimental, al presentarse un cambio en la matriz de pagos los jugadores permanecieron con el mismo rol², la única diferencia era la cantidad de puntos que podían recibir por una combinación de estrategias. Los juegos utilizados se muestran en las siguientes tablas:

Una de las particularidades de estos juegos, es que los denominados A y

¹Los jugadores ganaban o perdían dada una combinación de estrategias sin importar la matriz de pagos en activo

²Jugador “Fila” o “Columna”

| | | |
|--------|--------|-----|
| | Águila | Sol |
| Águila | 4 | -2 |
| Sol | -2 | 2 |

| | | |
|--------|--------|-----|
| | Águila | Sol |
| Águila | 6 | -3 |
| Sol | -3 | 3 |

| | | |
|--------|--------|-----|
| | Águila | Sol |
| Águila | 17 | -1 |
| Sol | -1 | 1 |

| | | |
|--------|--------|------|
| | Águila | Sol |
| Águila | 8.5 | -0.5 |
| Sol | -0.5 | 0.5 |

Cuadro 2.1: Juegos “A₁”, “A₂”, “B₁” y “B₂”.

| | | |
|--------|--------|-----|
| | Águila | Sol |
| Águila | 1 | -1 |
| Sol | -1 | 1 |

Cuadro 2.2: Juego presentado a todos los sujetos en la fase de entrenamiento.

B comparten una solución. La solución en los juegos A, es la de seleccionar con probabilidad de 0.6 la estrategia “Sol” (para ambos jugadores); mientras que la solución para los juegos B, es elegir con probabilidad de 0.9 la misma estrategia. El uso de estas matrices tuvo dos objetivos. El primero, es que la solución a un juego se alejara de la solución de la siguiente matriz, de manera que fuera posible observar un cambio en la conducta de los sujetos si es que dicho cambio se presentaba. El segundo, fue hacer que, sin importar la combinación de estrategias utilizadas por los participantes, el pago que recibieran fuera completamente diferente al pago recibido en la matriz anterior.

Los participantes nunca se enfrentaron a la misma matriz de pagos en más de una ocasión. El orden de presentación de las matrices para cada pareja se presenta en el cuadro 2.3. Las columnas indican los ensayos en los cuales la

matriz señalada en la segunda fila permaneció en activo, el orden en que los participantes se enfrentaron a cada matriz se mantuvo constante para todos los participantes.

| | | | | |
|--------|-------|---------|---------|---------|
| Ensayo | 1-100 | 101-200 | 201-300 | 301-400 |
| Matriz | A1 | B1 | A2 | B2 |

Cuadro 2.3: Orden de presentación de las matrices de pago a los participantes.

Cabe destacar que en todos los juegos, el jugador Fila tiene una ventaja en términos de la ganancia esperada, tomando esto en cuenta, el pago para todos los participantes fue el mismo al finalizar la sesión, no obstante, las instrucciones que recibían los sujetos indicaban que el pago que recibirían dependía de su desempeño en la tarea.

2.1. Procedimiento

Al ingresar al laboratorio, se les pidió a los participantes que eligieran alguna de las computadoras encendidas. Las computadoras fueron interconectadas previamente de manera aleatoria, esta conexión se mantuvo fija durante todo el experimento. Una vez interconectadas, cada par de computadoras se asignó a uno de dos roles disponibles (“Fila” o “Columna”³). La tarea fue programada utilizando el lenguaje Java.

El estudio se dividió en dos fases, entrenamiento y fase experimental: el entrenamiento consistió en 15 ensayos con una matriz de pagos independiente

³Fila: el jugador gana cuando las monedas elegidas son iguales
Columna: El jugador gana si las caras de las monedas son diferentes

a las de la fase experimental (ver cuadro 2.2, página 66). Antes del inicio de la fase experimental, se les explicó a los participantes cómo funcionaba el programa que utilizarían en el estudio, qué combinaciones de estrategias los harían ganar o perder y cuál debía ser su objetivo durante el experimento (conseguir la mayor cantidad de puntos).

Al inicio de la fase de entrenamiento, se presentaron las instrucciones de manera escrita en la pantalla de cada computadora. Después de leer las instrucciones, los participantes podían iniciar la fase de entrenamiento, sin embargo, si su oponente aún no estaba listo, se presentaba una ventana con el mensaje "Esperando al otro jugador". Cuando ambos jugadores terminaban de leer las instrucciones se iniciaba la fase de entrenamiento.

Al concluir los 15 ensayos de esta fase, se presentaron nuevamente los puntos más importantes de las instrucciones, como la combinación de estrategias que llevaría a ganar (rol del participante). Después de leer el resumen de las instrucciones, los participantes podían iniciar en cualquier momento con la fase experimental, una vez más, si uno de los jugadores terminaba antes que su oponente en la pantalla aparecía el mensaje: "Esperando al otro jugador".

La fase experimental fue dividida en cuatro subfases de 100 ensayos cada una. Estas cuatro fases se diferenciaban entre sí dependiendo de la matriz de pagos en activo. Las matrices de pagos para cada fase se mantuvieron fijas hasta concluir los 100 ensayos previstos (orden de presentación en cuadro 2.3; matrices de pago en cuadro 2.1).

Para ambas fases del experimento, cada ensayo consistió en la elección de una alternativa por parte de ambos jugadores que formaban una pareja. En

ese momento, los participantes observaban una pantalla en la que se presentaban dos monedas (una con la cara en *águila* y otra en *sol*), de las cuales, los jugadores debían elegir sólo una. Una vez que los jugadores llevaban a cabo su elección ésta no podía ser alterada. Además de las monedas, en la pantalla de inicio aparecían señaladas la cantidad de puntos acumulados por el jugador hasta el momento y los pagos del turno anterior. Cuando uno de los jugadores respondía primero, aparecía una ventana con la frase: “Esperando al otro jugador”. Una vez que ambos participantes habían llevado a cabo sus elecciones, se les asignó un pago dependiendo de la combinación de acciones de cada jugador y la matriz de pagos en activo para dicho ensayo.

Al finalizar cada ensayo, se presentó una pantalla en la cual aparecieron las elecciones de cada jugador (*i.e.* qué moneda eligió cada uno): la elección realizada por el oponente en ese ensayo, el pago asignado a cada jugador y la frase “ganaste” o “perdiste”, respectivamente. Durante la aparición de dicha pantalla, los participantes no podían emitir una respuesta durante 1 segundo, después de ese tiempo, los participantes podían regresar a la pantalla de inicio si presionaban un botón, de lo contrario, la ventana de resultado permanecía activa por 3 segundos.

La fase experimental consistió en el uso de 4 matrices de pago diferentes, para cada una de las cuales se presentaron 100 ensayos. La matriz de pagos cambió automáticamente al finalizar los 100 ensayos previstos, sin embargo, a los participantes no se les indicaba en qué punto había cambiado la matriz, por lo que tenían que inferirlo a partir de los pagos recibidos en cada ensayo.

A ninguno de los jugadores se les mostraron las matrices que formaban parte del experimento, esto con la finalidad de que, si existía un cambio en

el comportamiento de los participantes, éste no pudiera ser explicado por la presentación de una matriz diferente de un ensayo al siguiente.

Resultados

En este capítulo se presentan los resultados del experimento así como los análisis estadísticos correspondientes. El capítulo se divide de la siguiente manera:

- i)** Análisis general: Proporción de elecciones y Proporción de elecciones a través del tiempo.
- ii)** Contraste Modelos de Equilibrio.
- iii)** Contraste Modelos de Aprendizaje.

Todos los análisis y gráficos se llevaron a cabo en el programa R (R Core Team, 2014). Además se utilizaron los siguientes paquetes, para el ajuste por métodos de máxima verosimilitud el paquete “stats4”, para solución a sistemas de ecuaciones no lineales “nleqslv” (Hasselman, 2015), para el análisis con modelos gráficos Bayesianos “rjags” y “R2jags” (Plummer, 2015; Yu-Sung y Yajima, 2015) y para otros análisis utilizando los resultados del paquete “R2jags” se utilizaron los paquetes “ggmcmc” y “coda” (Plummer, *et. al.*, 2006; Fernández i Marín, 2016).

3.1. Análisis General

3.1.1. Proporción de Elecciones

Se analizaron primero las proporciones de elecciones de los jugadores para los 4 juegos del presente experimento. En las gráficas que se presentan a continuación los puntos morado, verde, azul y rojo representan a las parejas de jugadores 4, 13, 20 y 34 que se utilizarán a modo de ejemplo en los siguientes apartados.

En la Figura 3.1 se muestran las proporciones de elecciones de la estrategia “Sol” para cada una de las parejas de jugadores (círculos vacíos) así como la proporción media de elecciones de todos los participantes (círculos en gris) para los 4 juegos en el experimento (Juegos “A1”, “A2”, “B1” y “B2”). Los triángulos en color negro representan al equilibrio de Nash para cada juego, en este caso, los juegos “A” comparten una solución de acuerdo con la cual cada pareja debería de elegir su estrategia con una probabilidad de 0.6 “Sol”, mientras que en los juegos “B” la solución es elegir la misma estrategia pero con una probabilidad de 0.9.

Comenzando por el juego “A1” (*ver tabla 3.1*) cuya matriz de pagos se encontró activa de los ensayos 1 a 100, el valor medio de la proporción de elecciones de la estrategia “Sol” para los jugadores Fila se encuentra en 0.517, mientras que la media de los jugadores Columna se encuentra en 0.544 este tipo de asimetría es predicha por todos los modelos conductuales revisados anteriormente pero no por el equilibrio de Nash.

En los márgenes superior y derecho de la gráfica para el juego “A1” se

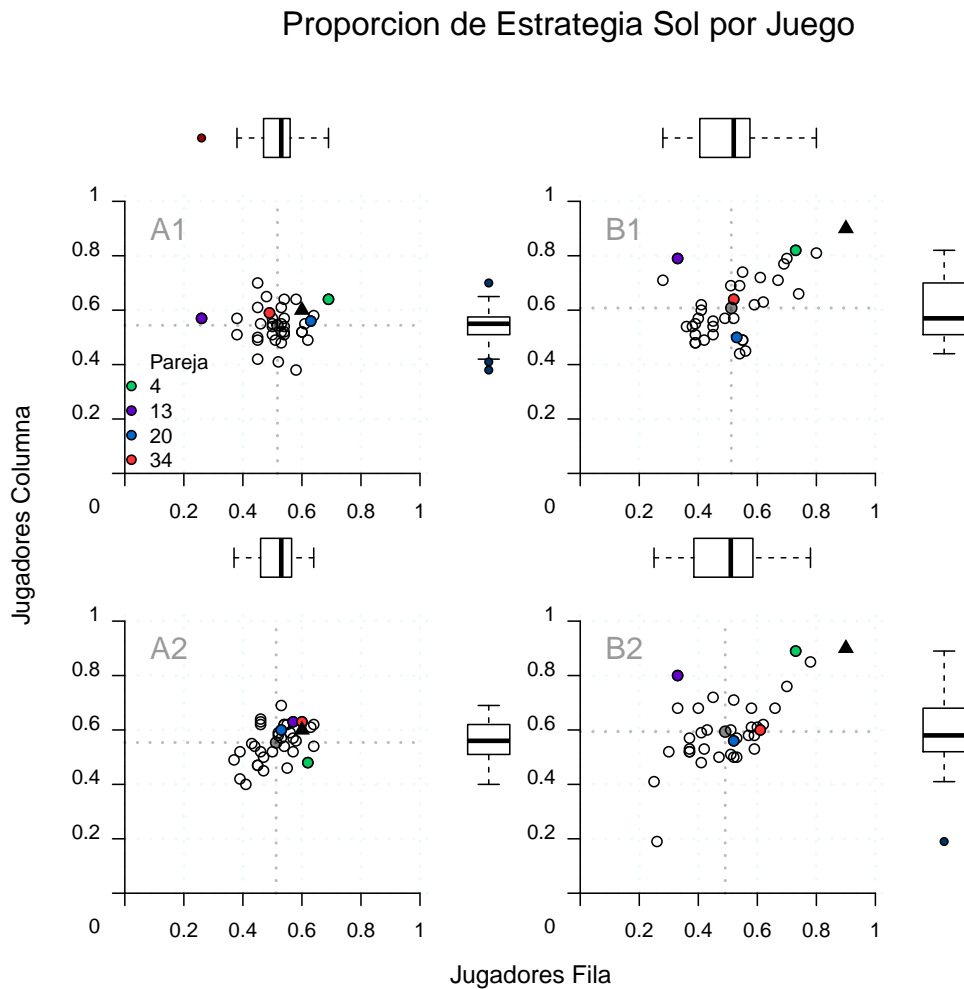


Figura 3.1: Gráficas de la proporción de elecciones “Sol” por parte de los jugadores, en el eje de las abscisas se encuentra la proporción de los jugadores Columna, mientras que en el eje de las ordenadas se encuentra la de los Fila. En cada gráfica los círculos vacíos representan la proporción de “Sol” para cada pareja, en el costado derecho y en la parte superior de cada gráfica se encuentra un gráfico de caja que representa la distribución marginal de las proporciones de cada tipo de jugador. El punto en color gris representa la proporción media de todos los jugadores, el triángulo en color negro representa el equilibrio de Nash en cada uno de los juegos.

| | | |
|--------|--------|-----|
| | Águila | Sol |
| Águila | 4 | -2 |
| Sol | -2 | 2 |

Cuadro 3.1: Juego “A1”, ensayos 1 a 100.

encuentran dos gráficos de caja para la proporción de elecciones de cada tipo de jugador, en este caso es posible observar que las proporciones varían menos en los jugadores Columna, sin embargo, la diferencia no es muy grande (0.007, 0.005) para los jugadores Fila y Columna respectivamente. De las 35 parejas de jugadores sólo 1 jugador fila y 3 jugadores columna se encuentran a una distancia mayor a 1.5 veces el rango intercuartil.

La segunda matriz de pagos (ensayos 101 a 200) fue la correspondiente al juego “B1” (*ver tabla 3.2*) y representa el primer cambio en la matriz de pagos para los participantes. La solución óptima para este juego es que los jugadores elijan con una probabilidad de 0.9 su estrategia “Sol”. Para este juego la proporción de veces en que los jugadores Fila eligieron su estrategia “Sol” fue de 0.512, mientras que para los jugadores Columna fue de 0.608.

| | | |
|--------|--------|-----|
| | Águila | Sol |
| Águila | 17 | -1 |
| Sol | -1 | 1 |

Cuadro 3.2: Juego “B1”, ensayos 101 a 200.

En términos del comportamiento de las parejas es posible observar que en este caso existe una mayor variabilidad ya que algunas de las parejas parecen se concentran al rededor de la equiprobabilidad (0.5) mientras que otras se encuentran distribuidas en dirección del equilibrio de Nash de este juego. Esto puede verse mejor representado en las gráficas de caja que se encuentran en el

margen derecho y superior de la gráfica en las cuales es posible observar que las colas de la caja se alargan más en la dirección correspondiente al equilibrio sin que los datos que se encuentran en esa zona se encuentren fuera del rango de 1.5 veces el rango intercuartil.

A diferencia del juego anterior, en ese caso, al finalizar los 100 ensayos, la proporción de ocasiones en las que las parejas eligieron su estrategia “Sol” presenta una mayor variabilidad, con una varianza de 0.016 y 0.012 para los jugadores Fila y Columna, respectivamente.

La siguiente matriz de pagos a la cual se enfrentaron los participantes fue la matriz “A2” (*ver tabla 3.3*), para este juego la solución óptima para los participantes es elegir su estrategia “Sol” con una probabilidad de 0.6.

| | | |
|--------|--------|-----|
| | Águila | Sol |
| Águila | 6 | -3 |
| Sol | -3 | 3 |

Cuadro 3.3: Juego “A2”, ensayos 201 a 300.

En este juego la proporción media de elecciones de la estrategia “Sol” por parte de los jugadores fila fue de 0.513, mientras que para los jugadores columna fue de 0.55. Al igual que en los casos anteriores se presenta una asimetría en el comportamiento de los jugadores lo cual contrasta con la predicción del equilibrio de Nash.

A diferencia del primero de los juegos (juego “A1”) en este caso existe una menor variabilidad en el comportamiento de los participantes. Es posible observar que la proporción de ocasiones en las cuales cada pareja eligió su estrategia “Sol” se encuentran más cercanas entre sí. Para este juego los gráficos de caja que se encuentran en los márgenes superior y en el derecho

de la gráfica muestran que ninguna de las parejas se encuentra a una distancia mayor a 1.5 veces el rango intercuartil. La varianza de las proporciones en este juego fue de 0.005 para ambos tipos de jugador.

El último de los juegos al cual se enfrentaron los participantes fue el juego “B2” (*ver tabla 3.4*), al igual que el segundo de los juegos, la solución óptima es que cada jugador elija su estrategia “Sol” con una probabilidad de 0.9.

| | | |
|--------|--------|------|
| | Águila | Sol |
| Águila | 8.5 | -0.5 |
| Sol | -0.5 | 0.5 |

Cuadro 3.4: Juego “B2”, ensayos 301 a 400.

Para este juego, la proporción media de elecciones de la estrategia “Sol” para los jugadores Fila fue de 0.49, mientras que para los jugadores Columna fue de 0.59, en este caso, el comportamiento promedio de los jugadores también se encuentra alejado de la solución óptima del juego.

Al igual que en el juego “B1” se presentaron algunas parejas cuya probabilidad de elegir su estrategia “Sol” se aleja de la equiprobabilidad en la dirección predicha por el equilibrio de Nash, es en parte debido a esto que la variabilidad en los juegos “B” es mayor en comparación con los juegos “A”. En el juego “B2” la varianza en las proporciones de los jugadores Fila fue de 0.0175 mientras que en el caso de los jugadores Columna fue de 0.0164. Para este último juego sólo el comportamiento de una de las parejas se encuentra fuera de 1.5 veces el rango intercuartil.

De las gráficas en la figura 3.1 vale la pena destacar que en los juegos que son estratégicamente equivalentes¹ (*i.e.* los juegos “A” tienen como solución

¹Juegos que comparten una solución

elegir la estrategia “Sol” con una probabilidad de 0.6) el comportamiento de los sujetos es bastante similar, en el caso de los juegos “A”, las proporciones de los participantes se encuentran muy cercanas entre sí, mientras que en los juegos “B” las proporciones de la elección “Sol” de los participantes presentan una mayor varianza y se encuentran cercanas a una diagonal entre el equilibrio de Nash y la equiprobabilidad.

3.1.2. Proporción de Elecciones a Través del tiempo

Si bien la proporción de ocasiones en las cuales los participantes eligieron su estrategia “Sol” al finalizar los 100 ensayos correspondientes a un juego otorga una referencia general acerca del comportamiento en el experimento. Esta medida no permite observar si a lo largo de los ensayos existen cambios en la forma en la que los participantes eligen entre sus alternativas.

A continuación se presenta una gráfica con la proporción de elecciones de la estrategia “Sol” por parte de los jugadores de las parejas 4, 13, 20 y 34. Las proporciones mostradas se calcularon utilizando 10 ensayos consecutivos, los datos presentados en cada gráfica corresponden a los 400 ensayos del experimento. Se eligieron a estas 4 parejas ya que presentan tendencias similares a las del resto de la muestra.

Estas gráficas permiten observar el comportamiento de los participantes a través del tiempo. Las líneas verticales representan el ensayo en la sesión en el cual ocurrió un cambio en la matriz de pagos, mientras que las líneas gruesas horizontales representan el punto de equilibrio en dicho juego. En la gráfica, las líneas en color rojo corresponden a la proporción de elecciones de la estrategia “Sol” de los jugadores Fila, mientras que las líneas azules

corresponden a los jugadores Columna.

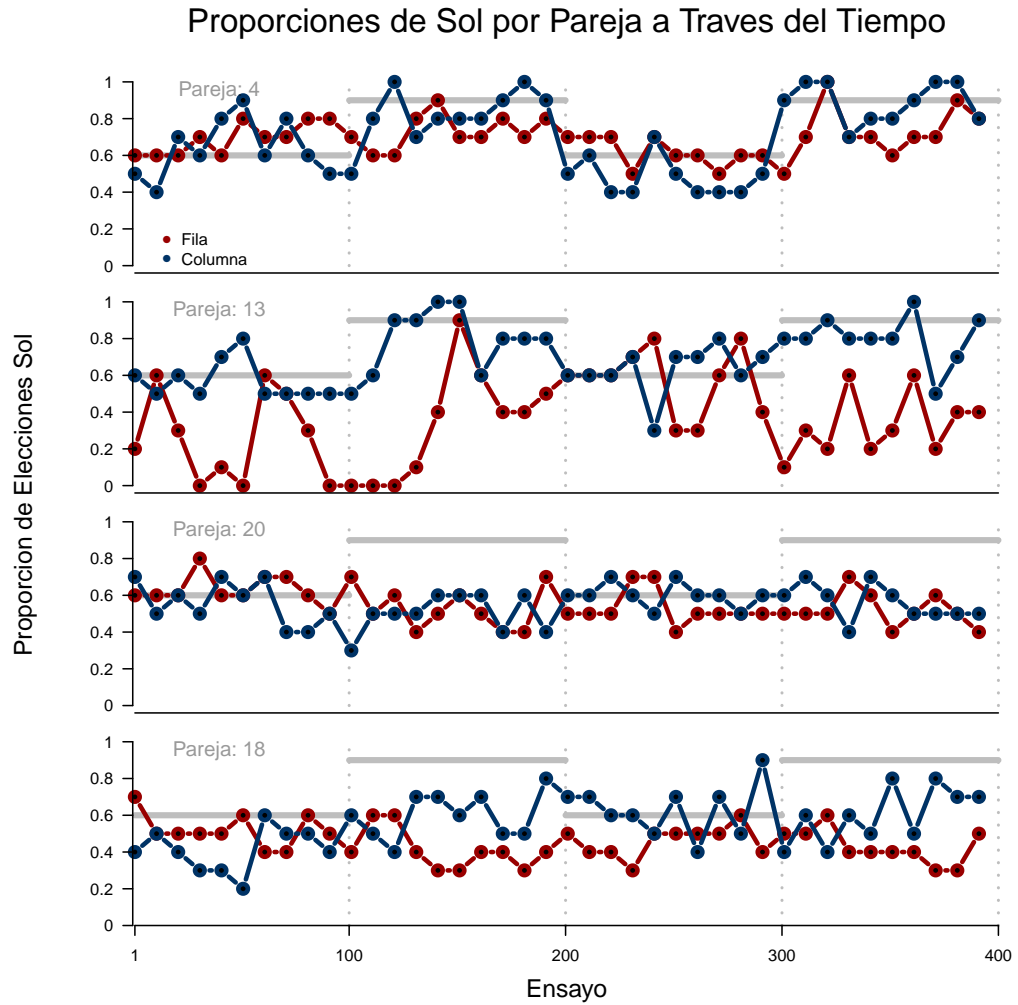


Figura 3.2: Gráficas de la proporción de elecciones “Sol” de 4 parejas en cada 10 ensayos. En el eje de las abscisas se encuentra el número de ensayo, mientras que en el eje de las ordenadas se encuentra la proporción de ocasiones en que los sujetos eligieron su estrategia “Sol” en 10 ensayos consecutivos. Las líneas en color rojo representan a los jugadores Fila mientras que las líneas en color azul representan a los jugadores columna. Las líneas sólidas en color gris representan el equilibrio de Nash para cada uno de los juegos.

Comenzando con la pareja número 4 (primera gráfica en la figura), en

ésta el comportamiento de los jugadores no sólo parece cambiar a lo largo de los ensayos, sino que la proporción de sus elecciones parece aproximarse a las predicciones del equilibrio de Nash cuando ocurre un cambio en la matriz de pagos.

Por ejemplo, al iniciar el primero de los juegos los participantes muestran una proporción de elecciones que se aproxima al equilibrio, sin embargo, al momento de cambiar al segundo juego cuya solución se encuentra en 0.9, al jugador Columna parece tomarle alrededor de 20 ensayos el aproximarse a la solución óptima. Mientras que a su oponente le toma alrededor de 30 ensayos con la nueva matriz de pagos acercarse al equilibrio.

En el caso de esta pareja este patrón se repite para los siguientes cambios en la matriz, es decir, la proporción de ocasiones en las cuales los participantes eligieron su estrategia “Sol” en 10 ensayos consecutivos se aproxima a la predicción del equilibrio de Nash también para los siguientes juegos. Sin embargo, este comportamiento no fue observado en todos los participantes que conforman la muestra.

La pareja 13 representa otro tipo de comportamiento observado en los participantes, en este caso, el comportamiento del jugador Columna se asemeja al de la pareja anterior, es decir, la proporción de veces en las que dicho participante eligió su estrategia “Sol” se aproxima a las predicciones del equilibrio de Nash, en cambio, las proporciones del jugador Fila difieren significativamente del comportamiento óptimo en cada uno de los juegos.

El caso de la pareja 20 es diferente a las 2 anteriores ya que esta pareja parece responder consistentemente al azar durante todo el experimento, sin mostrar cambios en su comportamiento ante los cambios en las matrices

de pagos. Este tipo de comportamiento es el que se esperaría sólo si los participantes se enfocaran en “ganar” o “perder”, es decir, si hicieran caso omiso del valor de cada alternativa.

La última de las parejas en la Figura 3.2 es la pareja 18 en la cual se observa que el comportamiento del jugador Columna es similar al caso de la pareja anterior, es decir, el participante parece no responder a los cambios en la matriz de pagos, pero además, la proporción de ocasiones en la que elige su estrategia “Sol” casi siempre fue menor para prácticamente todas las particiones de 10 ensayos.

En cambio, el participante en el rol de Fila de la misma pareja muestra un comportamiento que se asemeja a las predicciones de dos de los modelos de aprendizaje que se mostrarán más adelante. En este caso, las proporciones de elección de la estrategia “Sol” del participante siempre se encuentran por debajo del equilibrio de Nash, además, al momento del cambio en la matriz de pagos el comportamiento del participante cambia pero en dirección contraria a la predicción del equilibrio de Nash.

Esto sucede para los 3 cambios que conformaron el experimento, en este caso, en los juegos cuyo equilibrio se encontraba en 0.6 el comportamiento de algunos participantes en el rol de Fila se aproxima a la equiprobabilidad, mientras que en los juegos que tienen un equilibrio en 0.9 los participantes presentan proporciones menores al 0.5 y menores también en comparación con los otros 2 juegos.

3.2. Equilibrio

En esta sección se presentan los resultados del ajuste y comparación de los modelos de equilibrio discutidos previamente (*ver sección 1.3*). Los resultados correspondientes se organizan de la siguiente manera:

- i) Distribuciones Posteriores y Distribución posterior Predictiva de los modelos de Equilibrio por Muestreo de Pagos (PSE) y Equilibrio en Respuestas Cuantales (QRE).
- ii) Simulación de proporciones de respuesta por modelo.
- iii) Contraste de modelos mediante una aproximación a las posibilidades a posteriori (“*Posterior odds*”; PO, por sus siglas en inglés).

3.2.1. Distribuciones Posteriores y Posterior Predictiva

Los modelos que se trataran en esta sección son el de Equilibrio por Muestreo de Pagos (PSE) y el Equilibrio en Respuestas Cuantales (QRE). Estos son los únicos modelos de equilibrio que hacen uso de un parámetro libre.

En la figura 3.3 se presenta el modelo de PSE de manera gráfica, las especificaciones que se muestran en dicha figura fueron utilizadas para el análisis de datos en los 4 juegos de manera independiente. Además, las mismas especificaciones fueron utilizadas para la comparación de modelos.

En este tipo de gráficas (*i.e.* ver figura 3.3) las variables continuas se encuentran encerradas en un círculo, mientras que las variables discretas

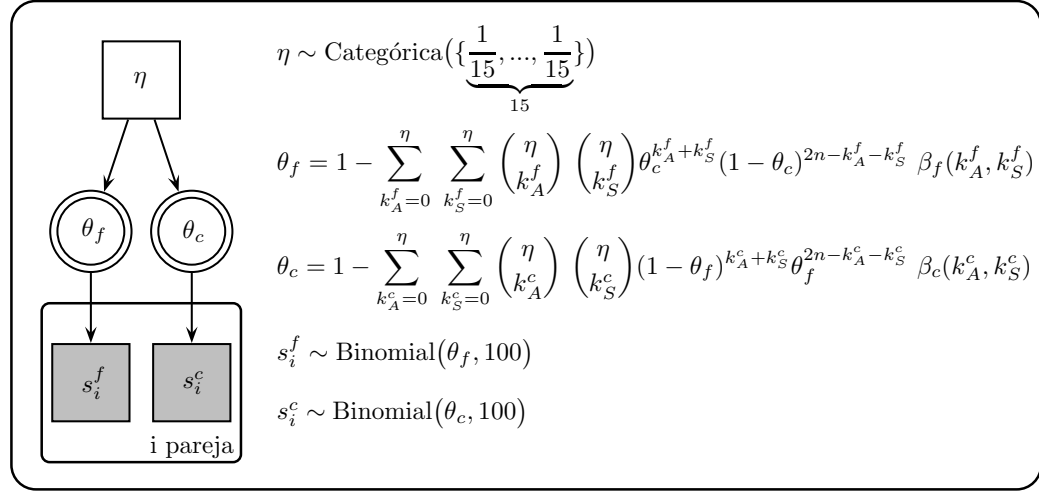


Figura 3.3: Representación gráfica del modelo de PSE, del lado derecho se muestran las relaciones de dependencia entre las variables, del lado izquierdo se muestran las distribuciones utilizadas a manera de información a priori así como la función de verosimilitud.

se encuentran en rectángulos; las variables conocidas (incluyendo los datos) aparecen sombreadas y las variables aleatorias no. Las flechas indican una relación de dependencia entre las variables.

Por último, los rectángulos que envuelven a un conjunto de variables se denominan “platos” y señalan en qué dimensión varía cada una de las variables utilizadas, por ejemplo, las variables s_i^f y s_i^c que representan el número de veces en las que un jugador Fila y Columna eligieron su estrategia “Sol”, se encuentran en el plato con el indicador i que representa a cada pareja $i = \{1, 2, \dots, 35\}$ por lo que en el análisis el modelo separa el número de veces que cada jugador eligió su estrategia “Sol” y trata esos conteos de manera independiente, sin embargo, el parámetro η en este modelo es único para todos los jugadores, por lo que la distribución posterior del parámetro contiene la información que proviene de la muestra completa.

Para los análisis que se llevaron a cabo en ésta y las demás secciones correspondientes a los análisis de los modelos de equilibrio, se utilizó la misma función de verosimilitud, la cual implica que los datos de los sujetos se distribuyen de manera Binomial con parámetros $p_k = \theta_k$ y $n = 100$. De la misma forma, las distribuciones *a priori* para los parámetros de cada modelo fueron siempre las mismas; Categorical para el parámetro η en el modelo de PSE y Gamma para el parámetro μ de QRE.

En la figura 3.4 se encuentra la probabilidad posterior media para cada valor del parámetro η en los 4 juegos, el cual, de acuerdo con el modelo, indica el número de ensayos que los jugadores utilizan para asignarle un valor a cada alternativa y así llevar a cabo sus elecciones (*ver sección 1.3, página 35*); en el presente análisis se restringieron los posibles valores que podría tomar este parámetro debido a que mayores valores implican una mayor proximidad al equilibrio de Nash, sin embargo, las frecuencias acumuladas no mostraron valores cercanos a dicha predicción.

De acuerdo con la figura 3.4, el valor más probable de acuerdo con la distribución posterior del parámetro η para los juegos “A” fue de 3, lo cual ocurre para ambos juegos, con este valor el modelo predice un punto de equilibrio de (0.535, 0.607) para los jugadores Fila y Columna respectivamente; este valor se aproxima a la media observada de las proporciones por parte de los jugadores fila, sin embargo se encuentra más alejada de la observada en los jugadores Columna (0.517, 0.544) con una diferencia de 0.018 y 0.063, respectivamente para el juego “A1”. En el caso del juego “A2”, la proporción media fue de (0.513, 0.55) la cual se aleja de la predicción en 0.22 y 0.057 para los jugadores Fila y Columna respectivamente.

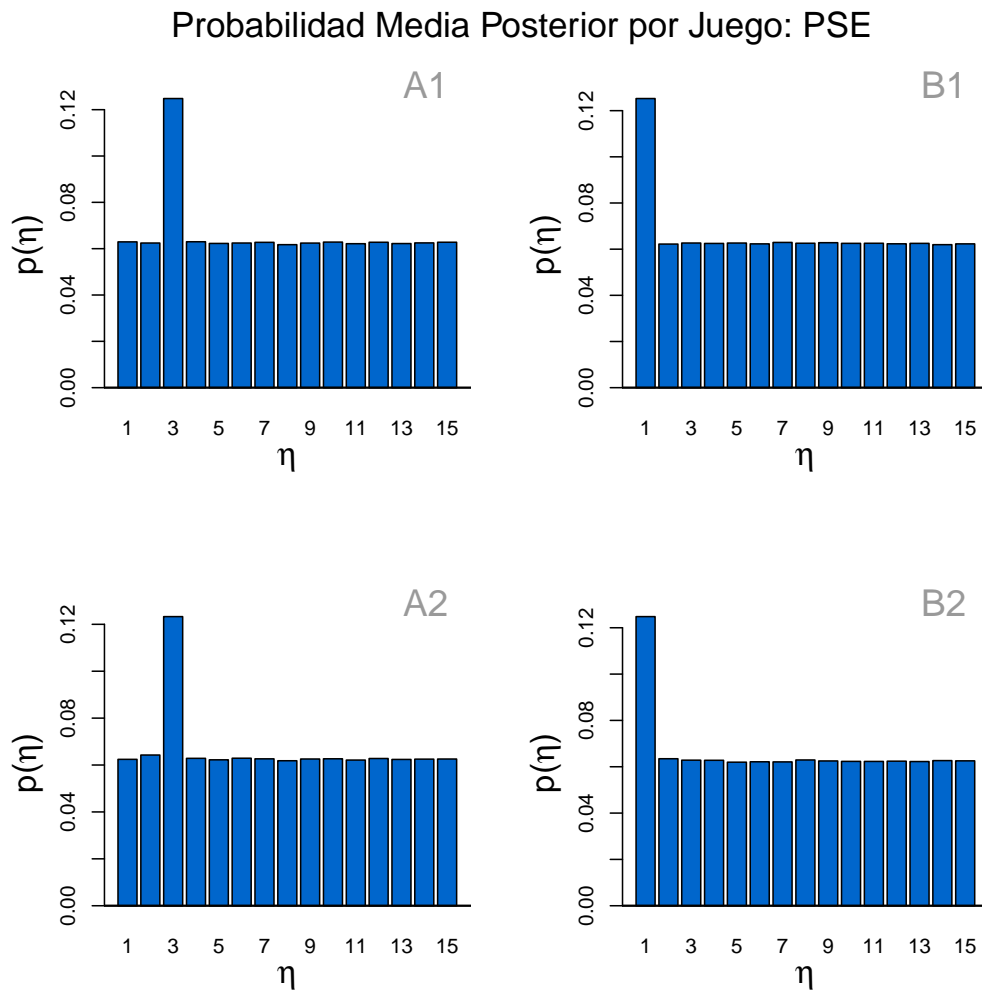


Figura 3.4: Gráfica de la probabilidad posterior media para cada valor del parámetro η en el modelo PSE. La altura de las barras representa la probabilidad posterior media para cada valor.

En el caso de los juegos “B” el valor que cuenta con una mayor probabilidad de acuerdo con los muestreos posteriores de la distribución es el 1, el cual da lugar a un punto de equilibrio de (0.504, 0.621), estos valores se alejan de la proporción media de la estrategia “Sol” en el juego “B1” (0.512, 0.608) por

0.008 y 0.013 para los jugadores Fila y Columna, respectivamente. En el caso del juego “B2” las proporciones se alejan un poco más de la media observada (0.49, 0.59) con un valor de 0.022 para los jugadores Fila y de 0.018 para los jugadores Columna.

A través de la distribución posterior del parámetro η es posible obtener una distribución de las predicciones del modelo en términos del número de ocasiones en las que los jugadores elegirán su estrategia “Sol” en los 100 ensayos de cada juego. La figura 3.5 muestra la distribución posterior predictiva para el modelo de PSE.

Las gráficas en la figura 3.5 se encuentran separadas por juego. En el caso de los juegos “A” (gráfica A1 y A2) se puede observar que, las proporciones de elección de varias de las parejas se encuentran dentro de la zona en la que los cuadrados tienen mayor tamaño², es decir, se encuentran en la zona que el modelo predice con mayor frecuencia, en cambio, en el caso de los juegos “B” (gráficas B1 y B2) esto no ocurre, en este segundo caso la mayoría de los datos se alejan de la zona que el modelo predice con mayor frecuencia.

Una de las características que comparten las predicciones del modelo es que su variabilidad siempre es menor que la variabilidad observada en los datos, es decir, el modelo subestima la variabilidad de las observaciones; esto ocurre de manera más evidente en el caso de los juegos “B”.

A continuación se presentan los resultados correspondientes al análisis del modelo de QRE. En la figura 3.6 se representa el modelo de manera gráfica. Al igual que el modelo anterior éste cuenta con un solo parámetro libre, sin embargo, en este caso el parámetro es continuo y puede tomar valores entre

²A mayor tamaño mayor es la probabilidad que el modelo asigna a dicha predicción.

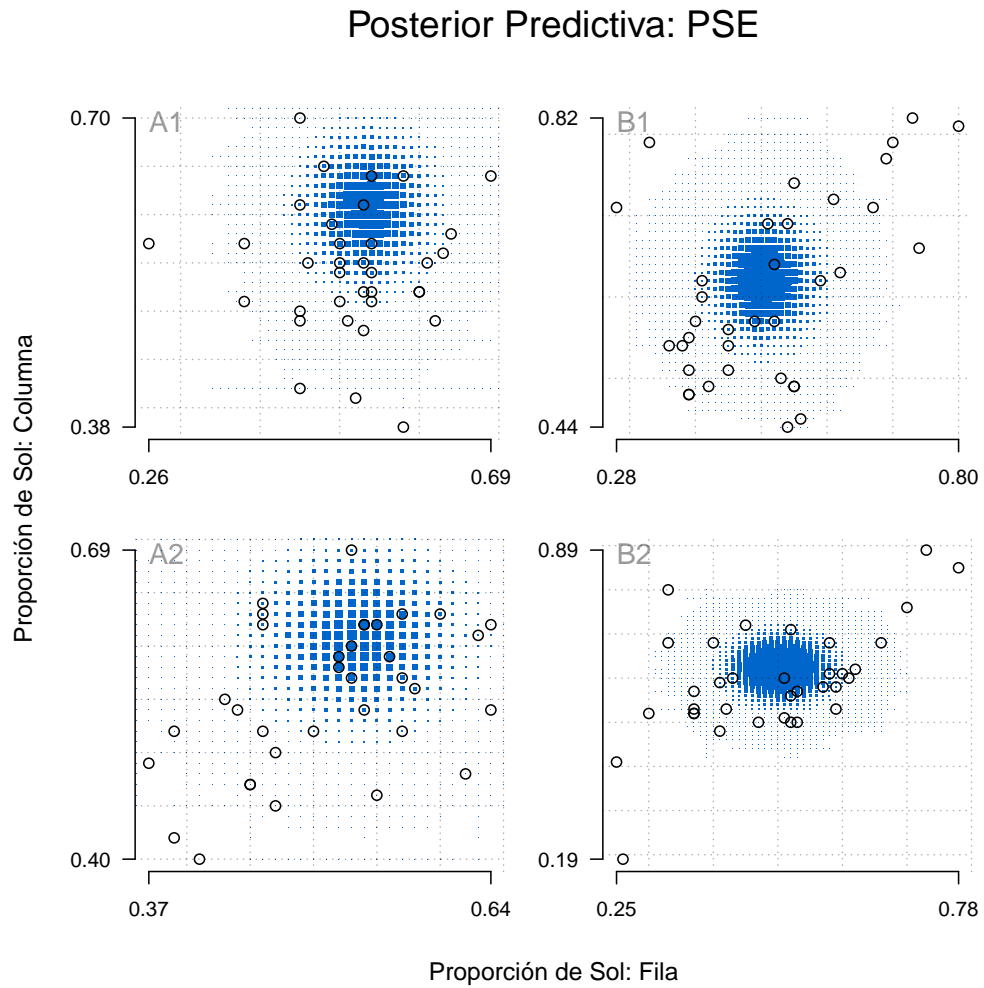


Figura 3.5: Distribución posterior predictiva para el modelo de PSE en cada juego. En el eje de las ordenadas se encuentra la proporción de ocasiones en que los jugadores Columna eligieron su estrategia “Sol”, en el eje de las abscisas se encuentra la de los jugadores Fila. Los cuadrados en color azul representan las predicciones del modelo, el tamaño del cuadrado representa la probabilidad de dicho resultado. Los círculos representan las proporciones de cada pareja en cada juego.

0 e ∞ . En este caso la distribución de probabilidad *a priori* utilizada fue una gamma con parámetros $\alpha = 0.001$ y $\beta = 0.001$; esta distribución se eligió de manera que influyera lo menos posible en las distribuciones posteriores

obtenidas.

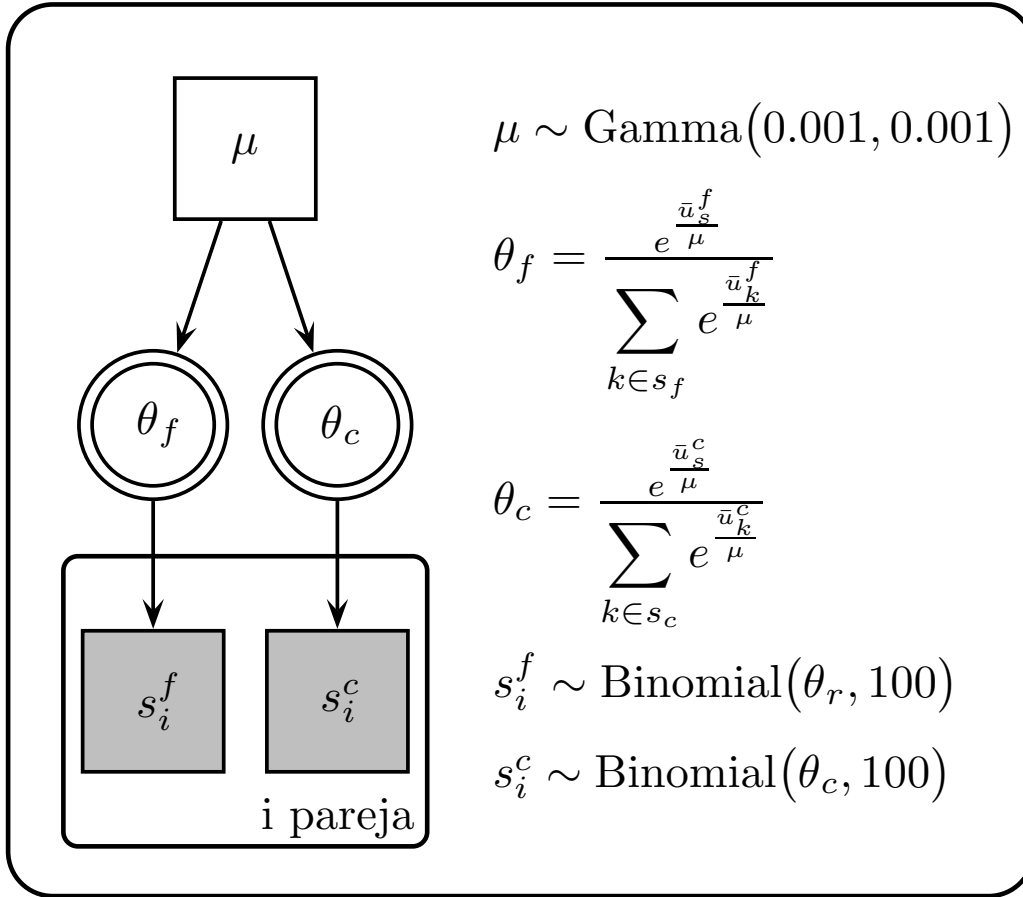


Figura 3.6: Representación Gráfica del modelo de QRE, del lado derecho se muestran las relaciones de dependencia entre las variables, del lado izquierdo se muestran las distribuciones utilizadas a manera de información a priori así como la función de verosimilitud.

En la figura 3.7 se presentan los histogramas de los muestreos posteriores del parámetro μ del modelo QRE. En cada una de las gráficas es posible observar la mediana de la distribución posterior del parámetro así como el intervalo de alta densidad al 95 % (“II”, para el límite inferior e “IS”, para el límite superior).

Densidad Posterior por Juego Modelo: QRE

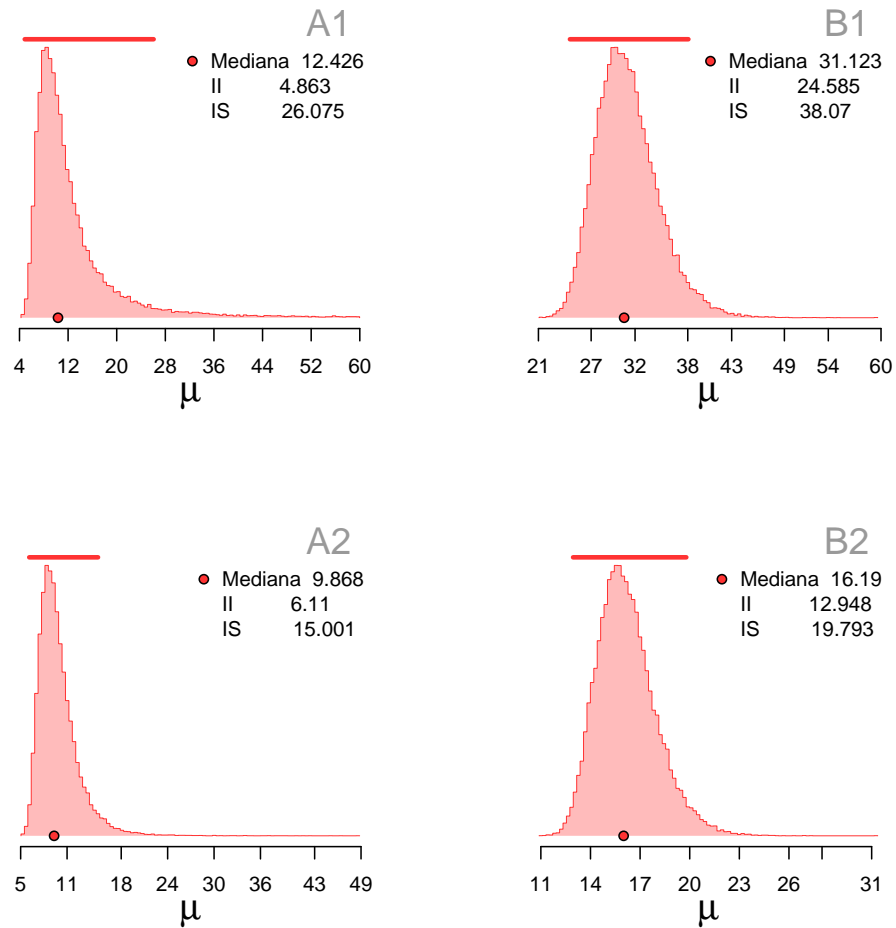


Figura 3.7: Histogramas de los muestreos posteriores del parámetro μ del modelo QRE por juego. Los puntos en color rojo representan la mediana de la distribución posterior, mientras que las rectas que aparecen sobre los histogramas representan el intervalo de alta densidad al 95%. Se presentan también los valores asociados a la mediana de la distribución, así como los límites inferior y superior del intervalo de alta densidad.

Las gráficas en la figura 3.7 muestran los histogramas de los muestreos de la distribución posterior del parámetro μ para cada uno de los 4 juegos. En primer lugar es posible observar que para cada categoría de los juegos

(juegos “A” y “B”) la forma de los histogramas es similar dentro de cada categoría, pero son diferentes para los juegos “A” en comparación con los juegos “B”. Además, es posible observar que los valores más probables del parámetro difieren de un juego al siguiente.

Las diferencias entre los posibles valores que puede tomar el parámetro de un juego al siguiente se explican por la sensibilidad a la magnitud de los pagos que presenta el parámetro μ en el modelo de QRE; un mismo valor para μ en el juego “A2”, por ejemplo 10, implicaría una mayor aproximación del comportamiento de los participantes al equilibrio de Nash en comparación con el juego “A1” ya que los pagos en este juego son menores. En otras palabras, el valor de μ se encuentra asociado a la magnitud de los pagos.

Debido a dicho efecto, las distribuciones posteriores que se encuentran en la figura 3.7 pueden generar predicciones similares para cada uno de los juegos.

Iniciando por las distribuciones correspondientes a los juegos “A” en la gráfica del juego “A1” se puede observar que la media posterior fue de 12.426, mientras que el intervalo de alta densidad al 95 % fue 4.863 y 26.075, para el límite inferior y superior respectivamente. Por otra parte, la media posterior en el juego “A2” fue de 9.868, mientras que el intervalo de alta densidad cuenta con un límite inferior y superior de 6.11 y 15.001, respectivamente.

El punto de equilibrio predicho con la media posterior de la distribución del parámetro μ para los juegos “A1” es 0.485 y 0.523, el cual se aleja de las proporciones observadas en 0.0326 y 0.021 para los jugadores Fila y Columna respectivamente; mientras que para el juego “A2” el punto predicho mediante la media posterior es 0.479 y 0.546, el cual se aleja en 0.033 y 0.008 unidades

de la proporción media observada de elecciones de la estrategia “Sol” en los participantes Fila y Columna.

En este caso, a diferencia del modelo de PSE, el modelo de QRE realiza predicciones diferentes para cada uno de los juegos (esto sucedería aunque se utilizara el mismo valor de μ para ambos), sin embargo, las predicciones del modelo no varían demasiado de un juego al siguiente.

La media posterior para el parámetro μ en el juego “B1” fue la mayor de los 4, con un valor de 31.123, mientras que su intervalo de alta densidad al 95 % es 25 para el límite inferior y 38 para el superior. El punto de equilibrio predicho por el modelo usando la media posterior del parámetro es 0.448 y 0.572 para los jugadores Fila y Columna respectivamente; este punto se alejó de la proporción media observada en 0.064 y 0.036 unidades para cada tipo de jugador.

En cambio, la media posterior para el parámetro μ en el juego “B2” fue de 16.19 mientras que para el intervalo de alta densidad se obtuvieron valores de 13 y 20. Utilizando la media posterior se obtuvo un punto de equilibrio para este juego de 0.449 y 0.569, el cual se aproxima al predicho en el juego “B1”, la distancia respecto a proporciones observadas fue de 0.042 y 0.024 para los jugadores Fila y Columna respectivamente.

En comparación con los juegos “A”, la media posterior y los intervalos de alta densidad correspondientes a los juegos “B” muestran valores más elevados, lo cual, de acuerdo con el modelo, implicaría una mayor distancia respecto al equilibrio de Nash en estos juegos, sin embargo, vale la pena destacar que las magnitudes de los pagos para los juegos “B” son mayores lo cual puede alterar la interpretación de los valores del parámetro.

Vale la pena destacar que la media posterior del parámetro μ en el juego “A2” fue menor a la obtenida en el juego “A1”, lo cual sugiere que en este segundo juego los participantes se aproximaron más al equilibrio de Nash. Lo mismo ocurrió en el caso de los juegos “B”, no obstante, en esos juegos es el juego “B1” el que cuenta con pagos más grandes, por lo tanto, dicho valor no tiene la misma implicación.

En la figura 3.8 se presentan las distribuciones posteriores predictivas del modelo de QRE para cada uno de los juegos.

En la figura 3.8 el tamaño de los cuadrados en color rojo representa la frecuencia relativa de cada resultado de acuerdo con el modelo, a mayor tamaño mayor es la probabilidad con la que el modelo espera dicho resultado. Una vez más, los círculos vacíos representan las proporciones observadas de elecciones de la estrategia “Sol” por parte de los participantes.

Iniciando con las gráficas de los juegos “A1” y “A2”, es posible observar que algunos de los círculos, los cuales reflejan las proporciones observadas para cada pareja, se encuentran en las zonas en las cuales los cuadrados muestran un mayor tamaño, es decir, que varias de las parejas mostraron un comportamiento el cual se aproxima a las predicciones del modelo, sin embargo, al igual que en el caso anterior, las predicciones del modelo muestran una menor varianza en comparación con la varianza observada en las proporciones de los participantes.

Al igual que con el modelo de PSE, las predicciones del modelo de QRE presentan mayores dificultades cuando se trata de los juegos “B”, en los cuales las predicciones del modelo se alejan de los datos, esto ocurrió para ambos modelos, no obstante, como se mostrará más adelante, esto ocurre también

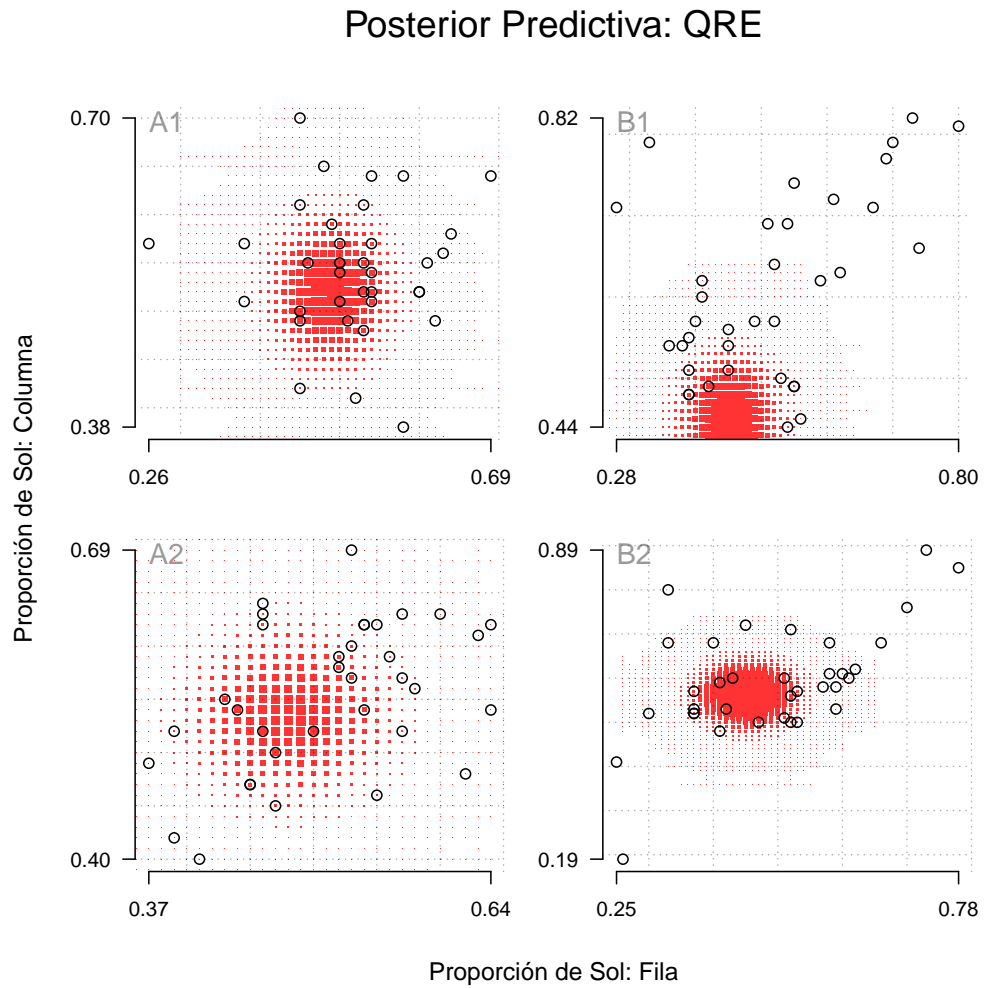


Figura 3.8: Distribución posterior predictiva para el modelo de QRE por juego. En el eje de las ordenadas se encuentra la proporción de elecciones de la estrategia “Sol” de los jugadores Columna, mientras que en el de las abscisas se encuentra la misma proporción pero para los jugadores Fila. Los cuadrados en color rojo representan las predicciones del modelo mientras que los círculos vacíos representan las proporciones de cada pareja.

en el caso de los modelos no paramétricos. De la misma forma que en el caso de los juegos “A”, en estos juegos las predicciones del modelo también muestran una menor varianza en comparación con los datos.

Por separado, las distribuciones posteriores predictivas presentan una noción gráfica acerca del comportamiento de los modelos en el presente experimento, sin embargo, para poder comparar los distintos modelos bajo estudio es necesario realizar un análisis que permita discriminar entre éstos no sólo de manera gráfica. En la siguiente sección se presentan los resultados de la comparación de estos modelos para cada uno de los juegos.

3.2.2. Comparación de Modelos

En este apartado se presentarán los resultados correspondientes a la comparación de los modelos de equilibrio. Además de los dos modelos anteriores se utilizaron también el equilibrio de Nash y el equilibrio por Balance de Impulsos (IBE) los cuales fueron tratados en la sección 1.3.

En la figura 3.9 se presenta el modelo gráfico utilizado para la comparación de los modelos de equilibrio. Para este análisis a los parámetros η y μ de los modelos de PSE y QRE respectivamente se les asignó la misma distribución *a priori* que en los análisis anteriores (Categórica y Gamma).

La variable z_i que aparece en la figura 3.9 tiene la función de indicar para cada pareja i cuál es el modelo que da una mejor cuenta de los datos, por lo que la frecuencia relativa de los valores de dicha variable en los muestreos posteriores funciona como una aproximación a la probabilidad de cada modelo en cada una de las parejas.

Además, se utilizó una distribución *a priori* que asigna igual probabilidad a cada valor de z_i para cada modelo, es decir que, al inicio todos los modelos son igualmente probables para cada pareja. Aunado a esto, se obtuvo un

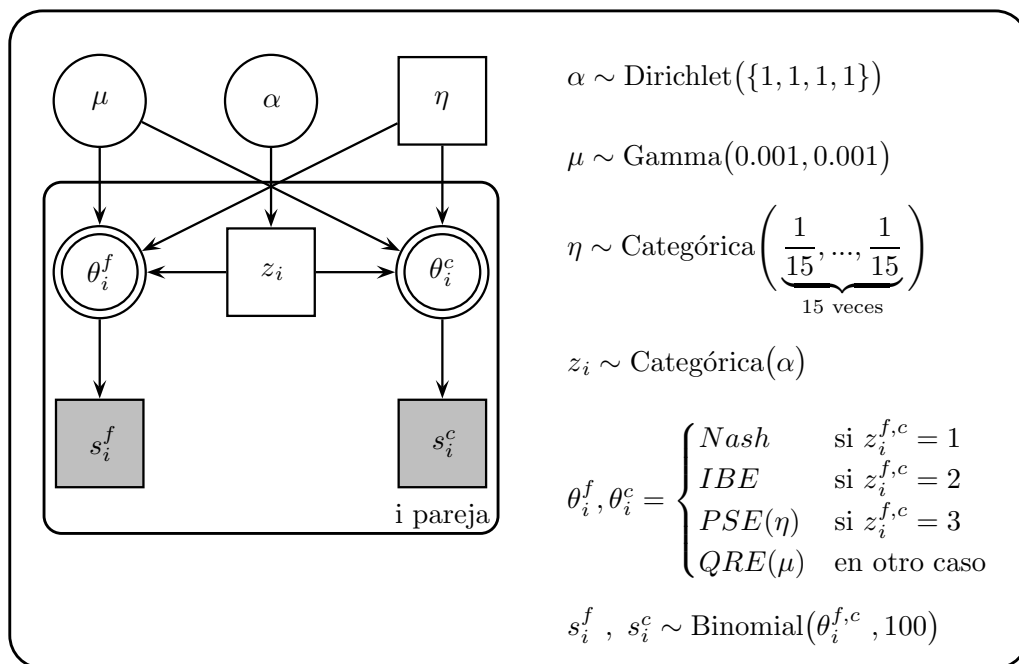


Figura 3.9: En la parte izquierda se presenta gráficamente el modelo para la comparación, mientras que en la derecha se presentan las distribuciones a priori asignadas a cada parámetro así como la función de verosimilitud.

muestreo posterior por juego acerca del comportamiento de la probabilidad de que la variable z tome cada uno de sus posibles valores (1 a 4) en la muestra completa para cada juego, este parámetro denominado α es una aproximación a la probabilidad de cada modelo en la muestra completa.

Antes de presentar los resultados de la comparación, en la figura 3.10 se presentan las predicciones de cada modelo para los 4 juegos, las cuales permiten realizar una comparación gráfica entre los modelos.

La primer fila de gráficas en la figura 3.10 presenta la distribución de las predicciones del equilibrio de Nash en términos de la proporción de ocasiones en que los sujetos elegirían su estrategia “Sol” si se comportaran de acuerdo

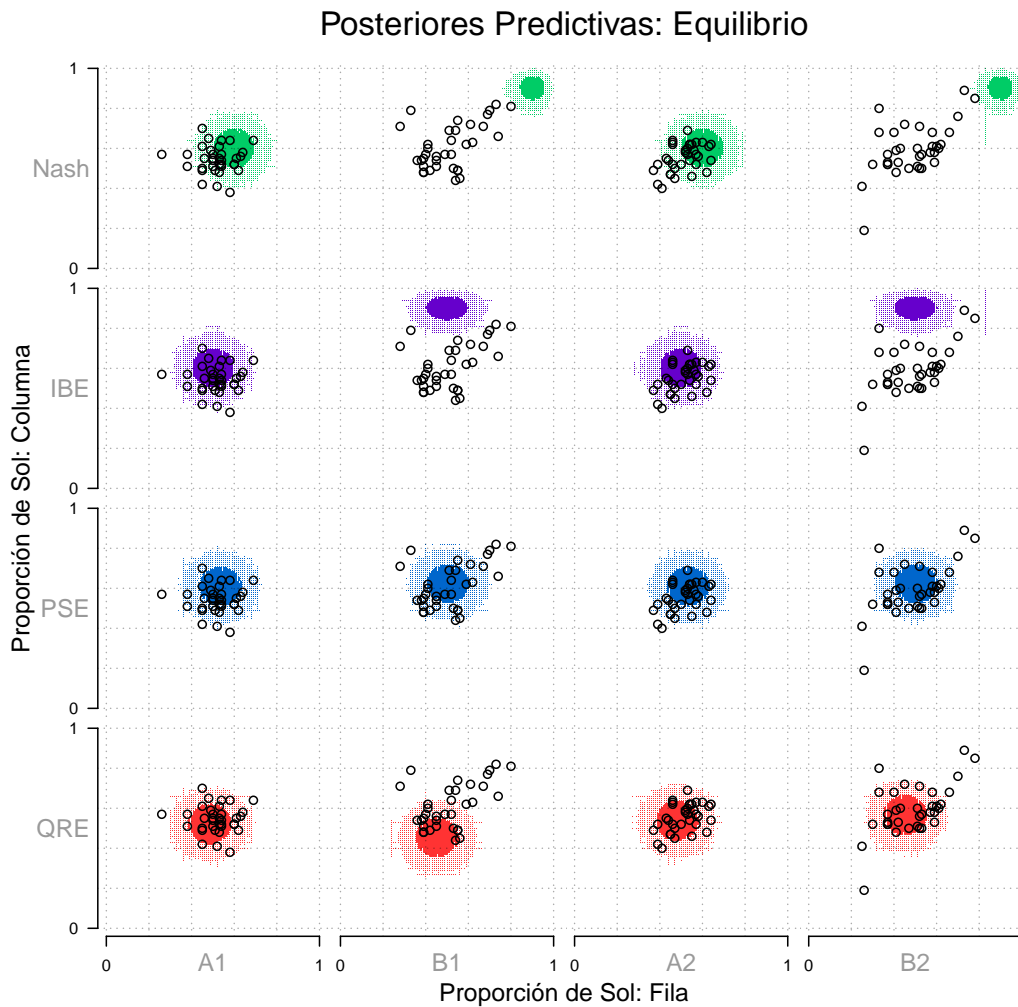


Figura 3.10: Distribuciones predictivas por modelo, en el eje de las abscisas se encuentra la proporción de ocasiones en que los participantes Fila eligieron su estrategia “Sol”, en el eje de las ordenadas se encuentran las proporciones de los participantes en el rol de Columna. Los cuadrados en colores representan los puntos predichos por cada modelo. Cada columna en la figura contiene la gráfica de un solo juego para los cuatro modelos, mientras que cada fila contiene las gráficas de las predicciones de un sólo modelo en los 4 juegos.

con la predicción del modelo.

Estas gráficas permiten observar que el equilibrio de Nash tiene un buen

desempeño cuando sus predicciones se comparan con el comportamiento de los participantes en los juegos “A”, sin embargo, éste no es el caso para los juegos “B”; en éstos, el comportamiento de la mayoría de los participantes se mantiene alejado de las predicciones del equilibrio de Nash.

Al igual que con el equilibrio de Nash, el modelo de IBE parece dar cuenta del comportamiento de los participantes en los juegos “A”, mientras que su desempeño empeora en los juegos “B” ya que en éstos, el comportamiento de la mayoría de los participantes se mantiene cercano a la equiprobabilidad mientras que el modelo predice, al menos para aquellos en el rol de Columna, que elegirán su estrategia “Sol” con una probabilidad de 0.9.

| | A1 | B1 | A2 | B2 |
|---------|-----|-----|-----|-----|
| Fila | 0.5 | 0.5 | 0.5 | 0.5 |
| Columna | 0.6 | 0.9 | 0.6 | 0.9 |

Cuadro 3.5: Puntos de equilibrio del modelo de Balance de Impulsos (IBE).

Los equilibrios predichos por este modelo se presentan en la tabla 3.5; la diferencia entre este modelo y el equilibrio de Nash es que el de IBE predice que los jugadores Fila elegirán con la misma probabilidad entre sus estrategias en los 4 juegos.

Las últimas dos filas en la figura 3.10 pertenecen a los modelo de PSE y QRE, respectivamente. Estas gráficas se mostraron anteriormente de manera más cercana, no obstante, en este caso permiten observar que, en general, sus predicciones no se alejan tanto del comportamiento observado en los participantes en comparación con los otros modelos de equilibrio; esto se debe a que sus predicciones permanecen cercanas a la equiprobabilidad al igual que en varios de los participantes.

Una de las formas en las que se han comparado los modelos de equilibrio recientemente (*i.e.* Selten y Chmura, 2008; Chmura, Goerg y Selten, 2012) es mediante el uso de una medida de distancia la cual se presenta en la ecuación 3.1 :

$$\bar{Q} = \frac{1}{n} \sum_{i=1}^n (\hat{p}_i^f - \theta_f)^2 + (\hat{p}_i^c - \theta_c) \quad (3.1)$$

En donde \hat{p}_i^f y \hat{p}_i^c representan la proporción de elecciones de la estrategia “Sol” por parte de los jugadores Fila y Columna respectivamente mientras que θ_f y θ_c representan la predicción de uno de los modelos.

En la figura 3.11 se presenta la media de las distancias cuadráticas de cada modelo para los 4 juegos. Para calcular la media con respecto de los modelos paramétricos se utilizó una muestra aleatoria de tamaño 5000 de las predicciones obtenidas previamente. Esto se hizo para tomar en cuenta las distintas probabilidades de cada valor de los parámetros en los modelos.

Las gráficas en la figura 3.11 muestran que de los 4 modelos, aquellos cuyas predicciones se aproximan más al comportamiento de los sujetos son las de PSE y QRE. Además, es posible observar que la distancia es mayor para todos los modelos en los juegos “B” en comparación con los juegos “A”.

Por último, las gráficas muestran que de los dos modelos paramétricos, el modelo de QRE presenta una menor distancia respecto a las proporciones observadas solo en los juegos “A”, mientras que el modelo de PSE se aproxima más en los juegos “B” aunque esto no resulta evidente en el juego “B2”.

A continuación se muestran las tablas que contiene los resultados de la comparación de los modelos de equilibrio utilizando la variable indicadora z_i . En estas tablas se presenta la proporción de ocasiones en las cuales dicha

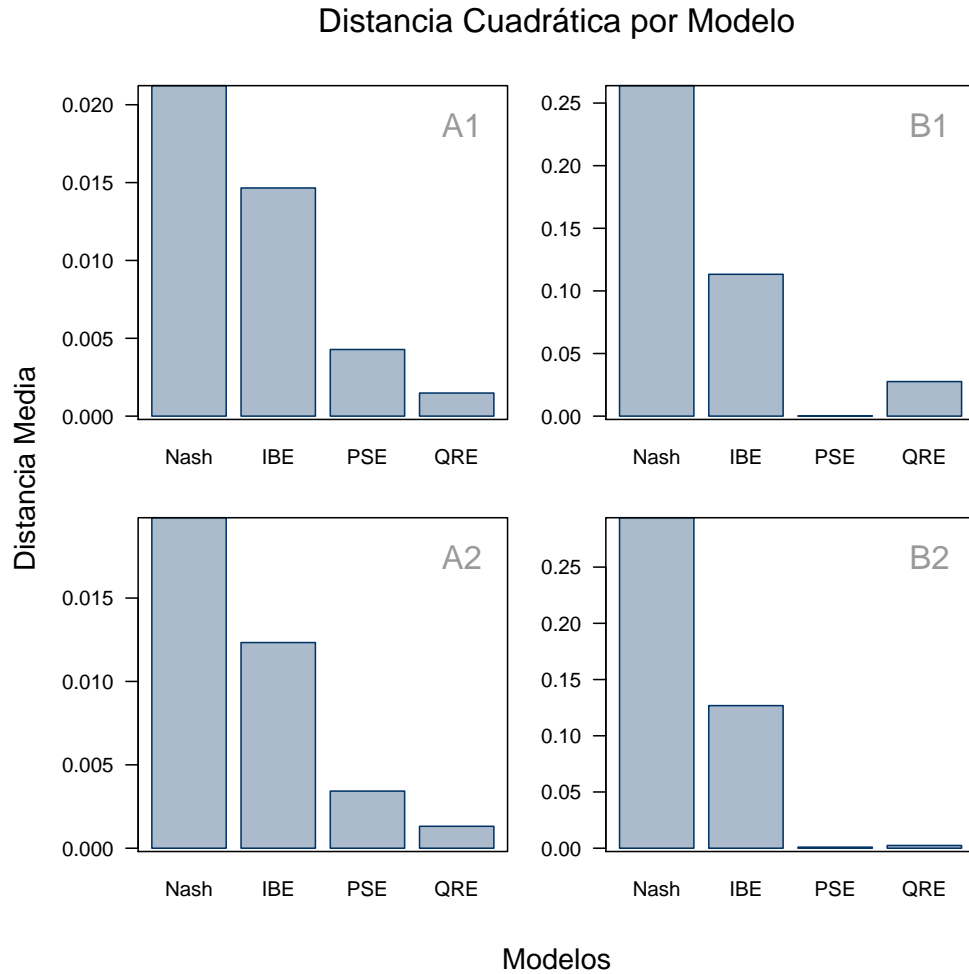


Figura 3.11: Distancia cuadrática por modelo para los 4 juegos.

variable tomó los valores correspondientes a cada modelo de equilibrio por pareja. Además, en la tabla se presenta una aproximación al factor de Bayes que resulta de la comparación de los 2 modelos más probables de acuerdo con las simulaciones de la distribución posterior de la variable para cada pareja.

La tabla 3.6 muestra un resumen de los resultados de la comparación

| Pareja | Nash | IBE | PSE | QRE | BF | Modelo 1 | Modelo 2 |
|--------|-------|-------|-------|-------|---------|----------|----------|
| 1 | 0.623 | 0.073 | 0.128 | 0.176 | 3.544 | Nash | QRE |
| 2 | 0.005 | 0.224 | 0.024 | 0.747 | 3.343 | QRE | IBE |
| 3 | 0.811 | 0.031 | 0.135 | 0.022 | 5.996 | Nash | PSE |
| 4 | 0.89 | 0.005 | 0.105 | 0.001 | 8.492 | Nash | PSE |
| 5 | 0.022 | 0.171 | 0.033 | 0.774 | 4.519 | QRE | IBE |
| 6 | 0.039 | 0.096 | 0.028 | 0.837 | 8.727 | QRE | IBE |
| 7 | 0.101 | 0.165 | 0.066 | 0.669 | 4.062 | QRE | IBE |
| 8 | 0.044 | 0.676 | 0.208 | 0.072 | 3.25 | IBE | PSE |
| 9 | 0.058 | 0.099 | 0.032 | 0.811 | 8.2 | QRE | IBE |
| 10 | 0.001 | 0.032 | 0.002 | 0.966 | 30.274 | QRE | IBE |
| 11 | 0.363 | 0.026 | 0.043 | 0.568 | 1.567 | QRE | Nash |
| 12 | 0.086 | 0.211 | 0.073 | 0.63 | 2.987 | QRE | IBE |
| 13 | 0 | 0.29 | 0.011 | 0.699 | 2.41 | QRE | IBE |
| 14 | 0.028 | 0.225 | 0.046 | 0.702 | 3.12 | QRE | IBE |
| 15 | 0.256 | 0.409 | 0.288 | 0.046 | 1.42 | IBE | PSE |
| 16 | 0.002 | 0.001 | 0 | 0.997 | 411 | QRE | Nash |
| 17 | 0.187 | 0.433 | 0.222 | 0.158 | 1.945 | IBE | PSE |
| 18 | 0.001 | 0.002 | 0 | 0.997 | 539.054 | QRE | IBE |
| 19 | 0.016 | 0.758 | 0.21 | 0.017 | 3.616 | IBE | PSE |
| 20 | 0.762 | 0.041 | 0.129 | 0.068 | 5.925 | Nash | PSE |
| 21 | 0 | 0.059 | 0.002 | 0.939 | 15.837 | QRE | IBE |
| 22 | 0.39 | 0.063 | 0.077 | 0.471 | 1.209 | QRE | Nash |
| 23 | 0.005 | 0.036 | 0.006 | 0.953 | 26.486 | QRE | IBE |
| 24 | 0.013 | 0.61 | 0.101 | 0.276 | 2.214 | IBE | QRE |
| 25 | 0.045 | 0.34 | 0.083 | 0.532 | 1.568 | QRE | IBE |
| 26 | 0.001 | 0.046 | 0.003 | 0.95 | 20.527 | QRE | IBE |
| 27 | 0.008 | 0.072 | 0.012 | 0.908 | 12.665 | QRE | IBE |
| 28 | 0 | 0.331 | 0.021 | 0.647 | 1.955 | QRE | IBE |
| 29 | 0.044 | 0.071 | 0.023 | 0.862 | 12.188 | QRE | IBE |
| 30 | 0.384 | 0.062 | 0.076 | 0.478 | 1.245 | QRE | Nash |
| 31 | 0.177 | 0.281 | 0.137 | 0.405 | 1.439 | QRE | IBE |
| 32 | 0.52 | 0.185 | 0.277 | 0.018 | 1.878 | Nash | PSE |
| 33 | 0.01 | 0.026 | 0.005 | 0.958 | 36.573 | QRE | IBE |
| 34 | 0.041 | 0.467 | 0.112 | 0.38 | 1.229 | IBE | QRE |
| 35 | 0 | 0.003 | 0 | 0.997 | 356.036 | QRE | IBE |

Cuadro 3.6: Probabilidad posterior de cada modelo por pareja en el juego A1. El factor de Bayes se obtuvo de la comparación de los dos modelos más probables por pareja.

de los cuatro modelos de equilibrio para cada una de las parejas. Para este primer juego, el equilibrio de Nash resultó el mejor modelo para 5 de las 35 parejas (14%), mientras que el de IBE fue el mejor para 6 (17%). En este

juego QRE resultó ser el mejor de los modelos para las 24 parejas restantes, es decir, para el 69 % de la muestra.

En los casos de parejas en las cuales el mejor modelo fue el de QRE, el modelo resultó desde 20 %, hasta 500 veces más probable que el segundo mejor de los modelos. Más adelante se presentarán los resultados de la probabilidad de cada valor de la variable indicadora z_i .

Al igual que con la comparación usando la distancia cuadrática, el modelo de QRE resultó ser el mejor modelo para la mayoría de los participantes, sin embargo, el equilibrio de Nash y el modelo de IBE resultaron mejores en comparación con el de PSE sin importar que su distancia cuadrática fuera mayor. Vale la pena destacar que el modelo de PSE fue el segundo más probable sólo para 8 de las 35 parejas (23 %).

En la tabla 3.7 se presentan los resultados para el juego B1. Como se mencionó en la sección anterior (*ver sección 3.2.1*), para este juego, los muestreos posteriores del parámetro η del modelo PSE no mostraron variabilidad, debido a esto se presentaron problemas con la convergencia de las cadenas de Markov para algunos de los participantes.

En este segundo juego, el modelo de QRE también resultó ser el modelo más probable para la mayoría de la muestra con 23 de 35 parejas (66 %), no obstante, de estas 23 en 11 no fue posible estimar el factor de Bayes debido a que los muestreos posteriores de la variable z_i no mostraron variabilidad. Por otra parte, el modelo de PSE resultó ser el mejor para 11 parejas (31 %) y sólo en un caso el mejor modelo fue el equilibrio de Nash.

Al igual que en el juego A1, en éste el mejor modelo no fue el que contaba con una menor distancia cuadrática (PSE), sin embargo, el desempeño del

| Pareja | Nash | IBE | PSE | QRE | Post.odds | Modelo 1 | Modelo 2 |
|--------|-------|-------|-------|-------|-----------|----------|----------|
| 1 | 0 | 0 | 0 | 1 | Inf | QRE | Nash |
| 2 | 0 | 0 | 0 | 1 | Inf | QRE | Nash |
| 3 | 0 | 0 | 0.784 | 0.216 | 3.633 | PSE | QRE |
| 4 | 0.001 | 0.001 | 0.998 | 0 | 718.932 | PSE | Nash |
| 5 | 0 | 0 | 0.969 | 0.031 | 30.746 | PSE | QRE |
| 6 | 0 | 0 | 1 | 0 | 9696.091 | PSE | QRE |
| 7 | 0 | 0 | 1 | 0 | 35555 | PSE | IBE |
| 8 | 0.997 | 0 | 0.003 | 0 | 338.707 | Nash | PSE |
| 9 | 0 | 0 | 0 | 1 | Inf | QRE | Nash |
| 10 | 0 | 0 | 1 | 0 | 35555 | PSE | IBE |
| 11 | 0 | 0 | 0 | 1 | Inf | QRE | Nash |
| 12 | 0 | 0 | 0 | 1 | 106667 | QRE | PSE |
| 13 | 0 | 0.011 | 0.972 | 0.016 | 59.191 | PSE | QRE |
| 14 | 0 | 0 | 0 | 1 | Inf | QRE | Nash |
| 15 | 0 | 0 | 0 | 1 | 2962 | QRE | PSE |
| 16 | 0 | 0 | 0 | 1 | 106667 | QRE | PSE |
| 17 | 0 | 0 | 0.382 | 0.618 | 1.62 | QRE | PSE |
| 18 | 0 | 0 | 0 | 1 | 4101.615 | QRE | PSE |
| 19 | 0 | 0 | 0.918 | 0.082 | 11.135 | PSE | QRE |
| 20 | 0 | 0 | 0 | 1 | Inf | QRE | Nash |
| 21 | 0 | 0 | 0.035 | 0.965 | 27.422 | QRE | PSE |
| 22 | 0 | 0 | 0 | 1 | Inf | QRE | Nash |
| 23 | 0 | 0 | 1 | 0 | 2962 | PSE | QRE |
| 24 | 0 | 0 | 0 | 1 | Inf | QRE | Nash |
| 25 | 0 | 0 | 0.001 | 0.999 | 832.344 | QRE | PSE |
| 26 | 0 | 0 | 0 | 1 | Inf | QRE | Nash |
| 27 | 0 | 0 | 0 | 1 | 53333 | QRE | PSE |
| 28 | 0 | 0 | 0 | 1 | 35555 | QRE | PSE |
| 29 | 0 | 0 | 0.002 | 0.998 | 645.473 | QRE | PSE |
| 30 | 0 | 0 | 0 | 1 | Inf | QRE | Nash |
| 31 | 0 | 0 | 1 | 0 | 2806.053 | PSE | QRE |
| 32 | 0 | 0 | 0.999 | 0.001 | 1870.368 | PSE | QRE |
| 33 | 0 | 0 | 0 | 1 | Inf | QRE | Nash |
| 34 | 0 | 0 | 0.231 | 0.769 | 3.335 | QRE | PSE |
| 35 | 0 | 0 | 0 | 1 | 26666 | QRE | PSE |

Cuadro 3.7: Probabilidad posterior de cada modelo por pareja en el juego B1.

modelo de PSE fue mejor que en el juego anterior.

En la tabla 3.8 se presentan los resultados por pareja de la comparación de modelo en el juego A2, este juego cuenta con el mismo equilibrio que el

juego A1 de acuerdo con el equilibrio de Nash y con el modelo de IBE.

| Pareja | Nash | IBE | PSE | QRE | Post.odds | Modelo 1 | Modelo 2 |
|--------|-------|-------|-------|-------|-----------|----------|----------|
| 1 | 0.002 | 0.19 | 0.027 | 0.781 | 4.116 | QRE | IBE |
| 2 | 0.013 | 0.11 | 0.039 | 0.838 | 7.597 | QRE | IBE |
| 3 | 0.618 | 0.035 | 0.34 | 0.008 | 1.819 | Nash | PSE |
| 4 | 0.296 | 0.023 | 0.086 | 0.595 | 2.009 | QRE | Nash |
| 5 | 0.316 | 0.169 | 0.328 | 0.187 | 1.037 | PSE | Nash |
| 6 | 0.278 | 0.224 | 0.377 | 0.122 | 1.357 | PSE | Nash |
| 7 | 0.115 | 0.284 | 0.248 | 0.352 | 1.239 | QRE | IBE |
| 8 | 0.642 | 0.025 | 0.329 | 0.004 | 1.952 | Nash | PSE |
| 9 | 0.103 | 0.368 | 0.292 | 0.237 | 1.259 | IBE | PSE |
| 10 | 0.19 | 0.318 | 0.427 | 0.065 | 1.345 | PSE | IBE |
| 11 | 0.166 | 0.091 | 0.135 | 0.607 | 3.655 | QRE | Nash |
| 12 | 0 | 0.002 | 0 | 0.998 | 448.348 | QRE | IBE |
| 13 | 0.331 | 0.18 | 0.466 | 0.023 | 1.409 | PSE | Nash |
| 14 | 0 | 0.103 | 0.005 | 0.891 | 8.641 | QRE | IBE |
| 15 | 0.096 | 0.168 | 0.15 | 0.586 | 3.492 | QRE | IBE |
| 16 | 0.001 | 0.246 | 0.033 | 0.72 | 2.923 | QRE | IBE |
| 17 | 0.233 | 0.274 | 0.44 | 0.054 | 1.604 | PSE | IBE |
| 18 | 0.018 | 0.618 | 0.231 | 0.132 | 2.676 | IBE | PSE |
| 19 | 0.018 | 0.633 | 0.252 | 0.097 | 2.516 | IBE | PSE |
| 20 | 0.141 | 0.35 | 0.358 | 0.152 | 1.021 | PSE | IBE |
| 21 | 0.189 | 0.315 | 0.431 | 0.065 | 1.366 | PSE | IBE |
| 22 | 0 | 0.019 | 0.002 | 0.978 | 50.232 | QRE | IBE |
| 23 | 0.353 | 0.135 | 0.308 | 0.204 | 1.146 | Nash | PSE |
| 24 | 0 | 0.037 | 0.001 | 0.962 | 26.159 | QRE | IBE |
| 25 | 0.003 | 0.108 | 0.019 | 0.87 | 8.077 | QRE | IBE |
| 26 | 0.093 | 0.338 | 0.259 | 0.31 | 1.093 | IBE | QRE |
| 27 | 0.013 | 0.016 | 0.011 | 0.96 | 61.339 | QRE | IBE |
| 28 | 0.016 | 0.599 | 0.21 | 0.175 | 2.846 | IBE | PSE |
| 29 | 0.664 | 0.026 | 0.233 | 0.077 | 2.849 | Nash | PSE |
| 30 | 0 | 0.021 | 0.002 | 0.977 | 46.732 | QRE | IBE |
| 31 | 0.135 | 0.333 | 0.525 | 0.007 | 1.577 | PSE | IBE |
| 32 | 0 | 0.001 | 0 | 0.999 | 726.255 | QRE | IBE |
| 33 | 0 | 0.009 | 0.001 | 0.989 | 105.221 | QRE | IBE |
| 34 | 0.474 | 0.088 | 0.429 | 0.009 | 1.105 | Nash | PSE |
| 35 | 0.002 | 0.057 | 0.01 | 0.932 | 16.311 | QRE | IBE |

Cuadro 3.8: Probabilidad posterior de cada modelo por pareja en el juego A2.

Una vez más, para este juego el mejor de los modelos resultó ser el de QRE, siendo el mejor para 17 de las 35 parejas (49%) de la muestra, mientras

que el modelo de PSE fue el mejor para 8 (23%) y por último tanto el equilibrio de Nash como el modelo de IBE resultaron ser los mejores modelos para el 14% y 14% de la muestra respectivamente.

Estos resultados son similares a los del juego A1 en el cual el equilibrio de Nash y el modelo de IBE fueron el mejor modelo en 15% de las parejas, sin embargo, la diferencia entre estos juegos es que el modelo de PSE ahora es el mejor modelo para un 23% de la muestra.

Para este juego, los resultados mediante este método de selección de modelos concuerdan con la conclusión que se podría extrater utilizando únicamente una medida de distancia como la cuadrática.

Para el modelo de QRE, los factores de Bayes aproximados mediante este método indican que el modelo es de 2% hasta 726 veces más probable que el segundo mejor de los modelos utilizados. En los casos en los cuales el mejor fue PSE los valores varían entre 3% y 60%.

Por último, en la tabla 3.9 se muestran los resultados de la comparación de los modelos de equilibrio por pareja para el último de los juegos (B2). Únicamente en uno de los casos el valor de la variable indicadora z_i no presentó variación en los muestreos posteriores a diferencia del juego B1 en el cual se presentaron varios casos similares.

Para este juego, a excepción de dos casos (parejas 3 y 4), los modelos de PSE y QRE son los más probables, el primero de los cuales es el mejor para 11 de las 35 parejas (31%), mientras que el segundo fue el mejor modelo en los 22 casos restantes (63%).

Una vez más, el mejor modelo resultó ser el de QRE, lo cual ocurrió para todos los juegos del presente experimento. Para este último juego los

| Pareja | Nash | IBE | PSE | QRE | Post.odds | Modelo 1 | Modelo 2 |
|--------|------|-------|-------|-------|-----------|----------|----------|
| 1 | 0 | 0 | 0.008 | 0.992 | 126.796 | QRE | PSE |
| 2 | 0 | 0 | 0.044 | 0.956 | 21.831 | QRE | PSE |
| 3 | 1 | 0 | 0 | 0 | 3332.25 | Nash | PSE |
| 4 | 0.44 | 0.523 | 0.036 | 0 | 1.188 | IBE | Nash |
| 5 | 0 | 0 | 0.47 | 0.53 | 1.128 | QRE | PSE |
| 6 | 0 | 0 | 0.563 | 0.437 | 1.288 | PSE | QRE |
| 7 | 0 | 0 | 0.996 | 0.004 | 271.109 | PSE | QRE |
| 8 | 0 | 0.001 | 0.999 | 0 | 739.704 | PSE | IBE |
| 9 | 0 | 0 | 0.185 | 0.815 | 4.409 | QRE | PSE |
| 10 | 0 | 0 | 0.328 | 0.672 | 2.053 | QRE | PSE |
| 11 | 0 | 0 | 0.263 | 0.737 | 2.808 | QRE | PSE |
| 12 | 0 | 0 | 0 | 1 | 9999 | QRE | PSE |
| 13 | 0 | 0.443 | 0.549 | 0.008 | 1.24 | PSE | IBE |
| 14 | 0 | 0 | 0.078 | 0.922 | 11.8 | QRE | PSE |
| 15 | 0 | 0 | 0.441 | 0.559 | 1.269 | QRE | PSE |
| 16 | 0 | 0 | 0.007 | 0.993 | 138.373 | QRE | PSE |
| 17 | 0 | 0 | 0.066 | 0.934 | 14.117 | QRE | PSE |
| 18 | 0 | 0 | 0.262 | 0.738 | 2.811 | QRE | PSE |
| 19 | 0 | 0 | 0.72 | 0.28 | 2.567 | PSE | QRE |
| 20 | 0 | 0 | 0.26 | 0.74 | 2.849 | QRE | PSE |
| 21 | 0 | 0 | 0.967 | 0.033 | 29.098 | PSE | QRE |
| 22 | 0 | 0 | 0.016 | 0.984 | 59.976 | QRE | PSE |
| 23 | 0 | 0 | 0.599 | 0.401 | 1.492 | PSE | QRE |
| 24 | 0 | 0 | 0.41 | 0.59 | 1.44 | QRE | PSE |
| 25 | 0 | 0 | 0.418 | 0.582 | 1.392 | QRE | PSE |
| 26 | 0 | 0 | 0.077 | 0.923 | 12.012 | QRE | PSE |
| 27 | 0 | 0 | 0.982 | 0.018 | 54.633 | PSE | QRE |
| 28 | 0 | 0 | 0.987 | 0.013 | 78.523 | PSE | QRE |
| 29 | 0 | 0 | 0.024 | 0.976 | 41.239 | QRE | PSE |
| 30 | 0 | 0 | 0.016 | 0.984 | 59.606 | QRE | PSE |
| 31 | 0 | 0 | 0.587 | 0.413 | 1.419 | PSE | QRE |
| 32 | 0 | 0 | 0 | 1 | Inf | QRE | Nash |
| 33 | 0 | 0 | 0.076 | 0.924 | 12.236 | QRE | PSE |
| 34 | 0 | 0 | 0.564 | 0.436 | 1.292 | PSE | QRE |
| 35 | 0 | 0 | 0.033 | 0.967 | 29.166 | QRE | PSE |

Cuadro 3.9: Probabilidad posterior de cada modelo por pareja en el juego B2.

factores de Bayes aproximados por pareja variaron entre el 1.12 y 9999, lo cual implica que el modelo es desde 12% o más probable que el segundo modelo más probable.

Para finalizar con la comparación de los modelos de equilibrio en la figura 3.12 se presentan las gráficas de la media posterior de la probabilidad de cada modelo en los 4 juegos, los valores representan únicamente la media posterior de la distribución del parámetro α (ver figura 3.9).

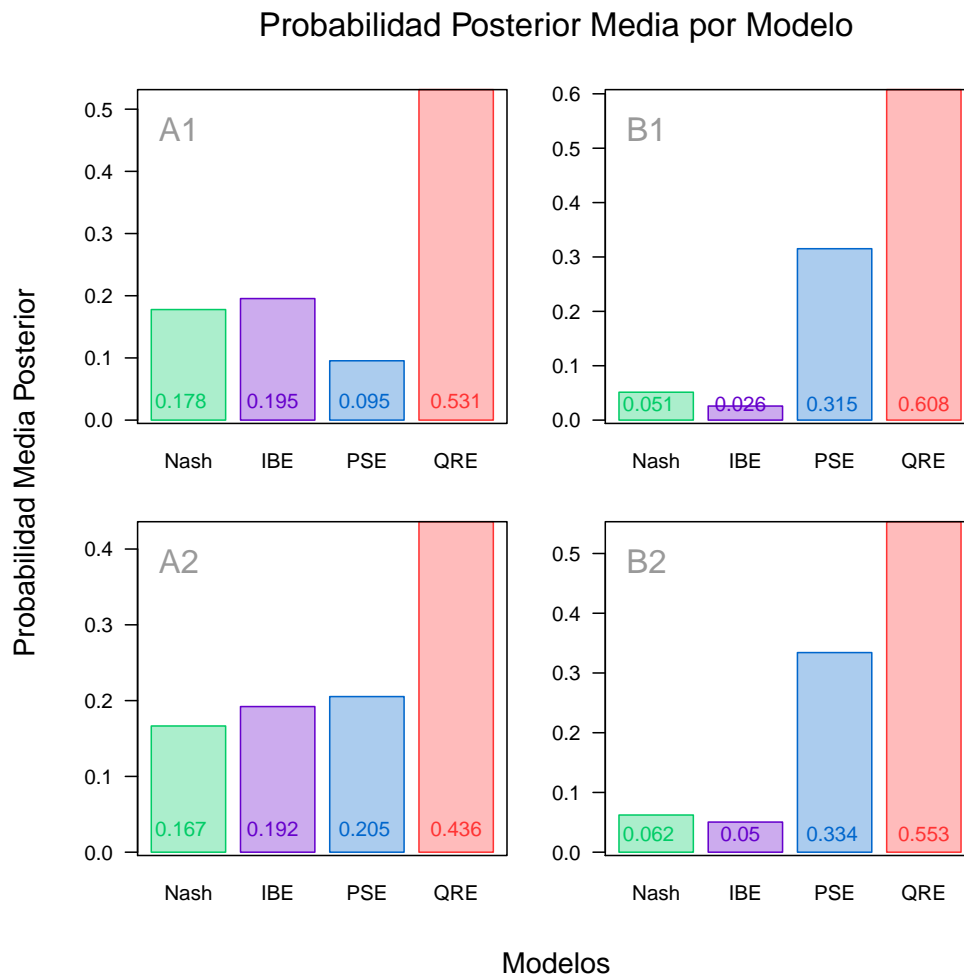


Figura 3.12: Probabilidad posterior media por modelo para los 4 juegos. En el eje de las abscisas se encuentran los modelos, en el eje de ordenadas se encuentra la media posterior para cada componente del parámetro α el cual indica la probabilidad del modelo en la muestra. Los números dentro de cada barra indican el valor de la media.

Como es posible observar en las gráficas de la figura 3.12 para los 4 juegos el modelo más probable fue el de QRE, los valores asociados a dicho modelo oscilan entre 0.43 y 0.608. Seguido de éste, se encontró que el segundo modelo más probable para 3 de los 4 juegos fue el de PSE, a excepción del juego A1 en el cual el modelo de IBE tiene una probabilidad mayor.

Estos resultados son, en su mayoría, consistentes con lo que se encontró utilizando la distancia cuadrática, a excepción de los juegos “B” en los cuales el modelo de PSE tiene una menor distancia pero que en la comparación resultó ser menos probable.

Si bien los modelos de equilibrio resultan de utilidad cuando no se cuenta con la información momento a momento del comportamiento de los sujetos, estos modelos sólo son capaces de dar cuenta del comportamiento de los participantes a un nivel agregado, por lo cual no otorgan información acerca de cómo puede cambiar el comportamiento de los sujetos a lo largo del tiempo.

En la siguiente sección se presentan los resultados del análisis y comparación de los modelos de aprendizaje. Vale la pena destacar que de esos modelos, dos (ST-EWA y SFP) utilizan la misma función de respuesta que el modelo de QRE el cual resultó ser el mejor de los 4 modelos de equilibrio.

3.3. Aprendizaje

En esta sección se presentan los resultados de los análisis de los modelos de aprendizaje que se presentaron en la sección 1.4, los apartados se organizan de la siguiente manera:

i) Distribuciones Posteriores y Distribución posterior Predictiva de los mo-

delos de Aprendizaje por Balance de Impulsos (IBL), Aprendizaje por Reforzamiento (REL), Juego Ficticio Estocástico (SFP) y Aprendizaje Auto-Regulable de Atracciones Ponderadas por la Experiencia (St-EWA).

iii) Contraste de modelos de aprendizaje.

3.3.1. Distribuciones Posteriores y Posterior Predictiva

En este apartado se presentarán los resultados de los análisis con los distintos modelos de aprendizaje, así como las predicciones que realiza cada modelo.

En la figura 3.13 se presentan los resultados de la simulación utilizando el modelo de IBL, este modelo no cuenta con parámetros libres por lo tanto sus predicciones se realizan utilizando únicamente la historia de juego de los participantes.

Las simulaciones en la figura 3.13 se realizaron utilizando únicamente el diseño del experimento, es decir, se utilizaron sólo los juegos y número de ensayos junto con las funciones que conforman al modelo para generar una predicción acerca del comportamiento esperado de los participantes.

De acuerdo con las simulaciones el modelo hace predicciones diferentes respecto a cómo responderá cada tipo de participante (jugador Fila o Columna) ante un cambio en la matriz de pagos. En el caso de los jugadores Columna, la predicción del modelo es que, después del primer cambio (ensayo 101), los jugadores comenzarán a elegir su estrategia “Sol” con una mayor probabilidad conforme pasan los ensayos.

Predicción del Modelo IBL

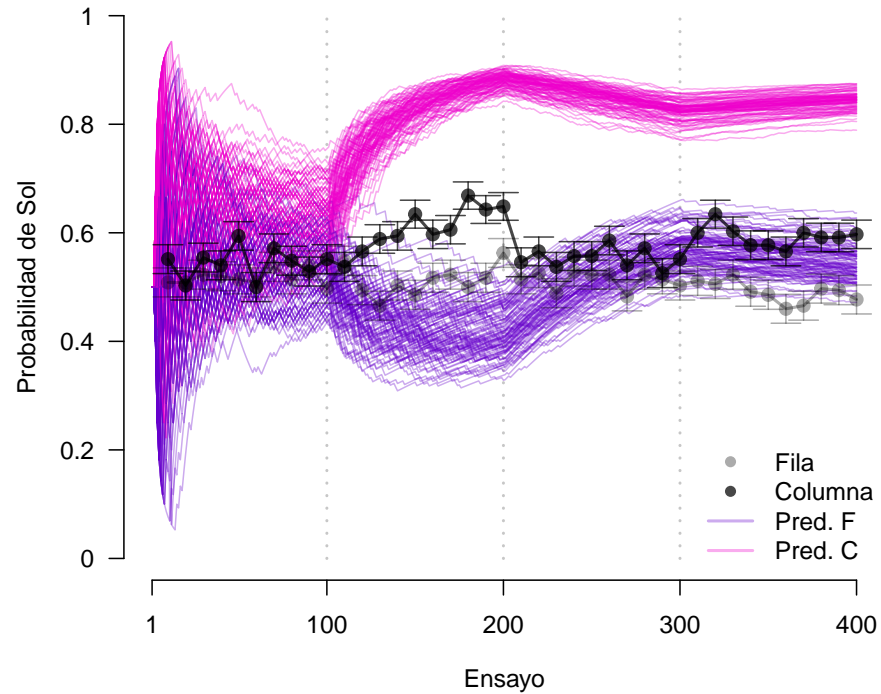


Figura 3.13: Simulación del modelo de IBL, en el eje de las ordenadas se encuentra la probabilidad (proporción) de elegir la estrategia “Sol”, en el de las abscisas se encuentran los ensayos. Las líneas a color representan la probabilidad para cada tipo de jugador de elegir la estrategia “Sol” en cada ensayo en una simulación. Las líneas verticales indican el ensayo en el que ocurrió un cambio en la matriz de pagos. Los puntos en color gris y negro representan la proporción media de elecciones de la estrategia “Sol” de los participantes en cada 10 ensayos.

Por otra parte, en el caso de los jugadores Fila, el modelo predice que su comportamiento cambiará en la dirección opuesta, es decir, el modelo predice

que en el juego B1, los jugadores Fila comenzarán a utilizar su estrategia “Sol” con una menor probabilidad.

Las predicciones acerca del comportamiento de los jugadores Columna no varían mucho después del juego B1 (ensayo 201 en adelante), para los ensayos correspondientes al juego A2 (201 a 300) el modelo predice un ligero decremento en la probabilidad con la que los jugadores Columna elegirán su estrategia “Sol” en comparación con el juego B1, sin embargo, el cambio no es tan pronunciado como en el paso del juego A1 al B1.

En contraste con estas predicciones, para los jugadores Fila la predicción es que en el segundo de los cambios en el juego, la probabilidad con la que los participantes usarán su estrategia “Sol” incrementará, terminando con predicciones similares a las que realiza el modelo para el juego A1 (ensayos 1 a 100), una vez concluidos estos ensayos el modelo predice un ligero cambio en el comportamiento de los participantes en favor de su estrategia “Águila”, sin embargo, dicho cambio es pequeño.

Los puntos en gris y negro en la gráfica representan la proporción media de elecciones de la estrategia “Sol” de los participantes en cada 10 ensayos. Debido a que las simulaciones del modelo muestran mucha variabilidad en el primero de los juegos (ensayos 1 a 100), no es posible realizar una comparación con el comportamiento promedio de los participantes.

Comparando el comportamiento de los jugadores Fila con las predicciones del modelo es posible observar que las proporciones de elección de la estrategia sol de los participantes (puntos en color negro) por lo general se encuentran en la misma zona que las predicciones del modelo, sin embargo, estos puntos no muestran la misma tendencia que las predicciones del mode-

lo, es decir, el comportamiento de los jugadores se muestra más estable de lo que este modelo predice.

En contraste, para el segundo juego el comportamiento de los jugadores Columna muestra una tendencia similar a la de la predicción del modelo, es decir, un incremento en la probabilidad con la que los participantes eligieron su estrategia “Sol”, sin embargo, este incremento no tiene la misma magnitud que la predicción y el promedio de elecciones de los participantes siempre permanece por debajo.

Debido a que la probabilidad predicha por el modelo no se modifica demasiado con los dos cambios restantes en la matriz de pagos, las predicciones se encuentran por encima de la proporción media de elecciones para los juegos A2 y B2, sin embargo, la dirección del cambio predicha por el modelo (menor probabilidad en el juego A2 e incremento de la probabilidad en el juego B2) se asemeja a las tendencias que se pueden observar en los datos.

De los cuatro modelos de aprendizaje, el de IBL es el que mostró una mayor sensibilidad al primero de los cambios, sin embargo, esto no fue así para los cambios subsecuentes. Además, las proporciones de los sujetos no mostraron un comportamiento similar al predicho por el modelo, es decir, un incremento abrupto en el primer cambio y menor sensibilidad en los restantes

En la figura 3.14 se muestra el modelo gráfico utilizado en el análisis del modelo de CR (Erev y Roth, 1998), este modelo cuenta con un parámetro libre que son las atracciones iniciales ($A_{t=0}$). Para poder mostrar las atracciones en cualquier ensayo como una función de las atracciones iniciales se utilizó una variable indicadora ($I(x, y)$) la cual toma un valor de 1 cuando

$x = y$ y de 0 en cualquier otro caso.

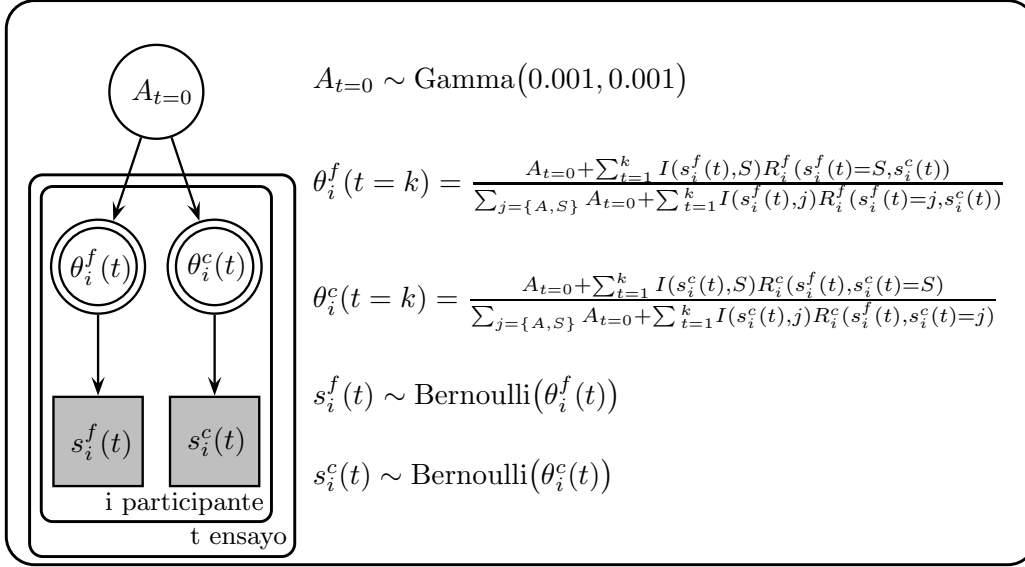


Figura 3.14: Representación gráfica del modelo de Reforzamiento Acumulado (CR; Erev y Roth, 1998). Las funciones completas asociadas con este modelo se encuentran en la sección 1.4.1.

Para este modelo se asignó una distribución *a priori* Gamma para las atracciones iniciales ya que esta distribución cumple con las dos restricciones del modelo (ver sección 1.4.1). Debido a que la probabilidad de que cada jugador elija una estrategia cambia en función de los pagos que ha recibido en el experimento, las probabilidades de que un jugador elija su estrategia “Sol” ($\theta_i^f(t)$, $\theta_i^c(t)$) se encuentran indexadas por pareja y por ensayo. Por último, a diferencia de los modelos de equilibrio, la función de verosimilitud toma la forma de una distribución Bernoulli, ya que la probabilidad de elegir cada estrategia cambia en cada ensayo.

En la figura 3.15 se muestran la distribución posterior de las atracciones iniciales y las predicciones del modelo acerca de la probabilidad con la que

los participantes elegirán su estrategia “Sol” a lo largo de los ensayos.

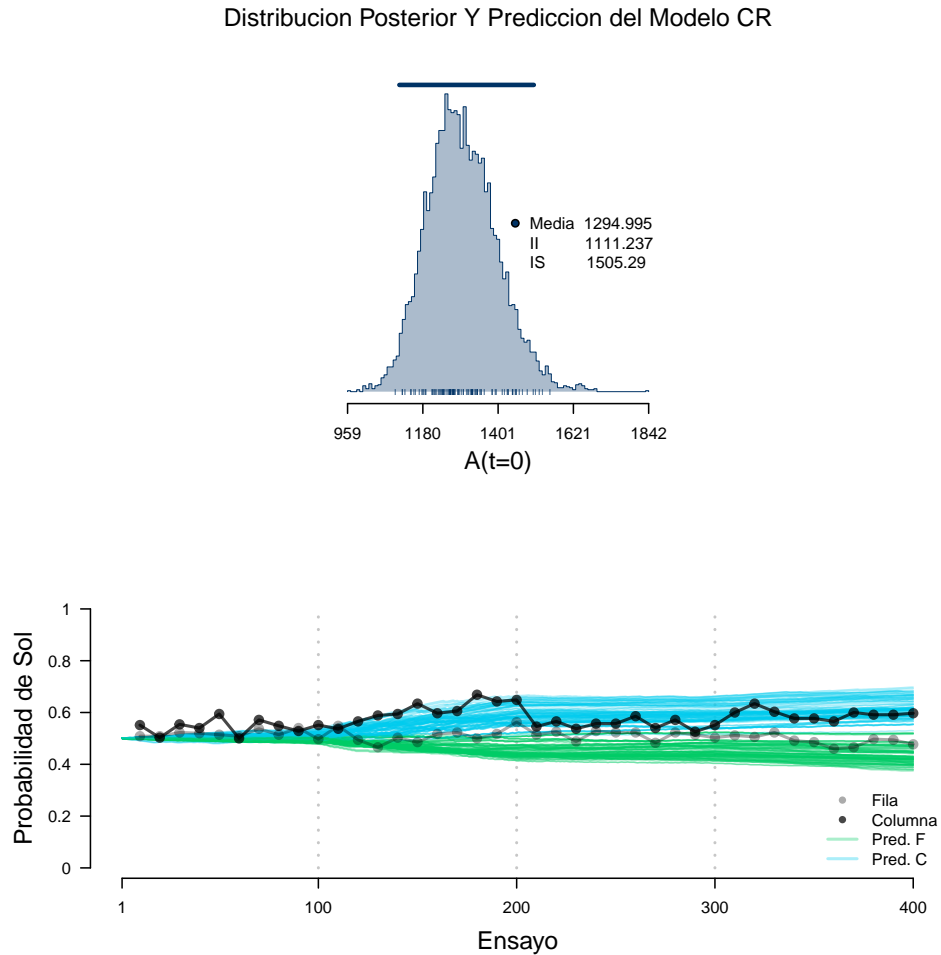


Figura 3.15: En la primera gráfica se presenta un histograma de los muestreos de la distribución posterior de la atracción inicial en el modelo de CR. En la segunda gráfica se presenta una muestra aleatoria de los valores de la variable θ predichos por el modelo dado un valor de la atracción inicial; en color gris se presenta el promedio cada 10 ensayos de las elecciones de los jugadores Columna, en negro se presenta el promedio de los jugadores Fila.

La gráfica en la parte superior de la figura 3.15 muestra los muestreos

posteriores de la atracción inicial en el modelo de CR. En este caso, la media posterior fue de 1294.995, mientras que los valores del intervalo de alta densidad al 95 % fueron de 1111.237 para el intervalo inferior y de 1505.29 para el superior. Estos valores son elevados e implican que el impacto del resultado de cada ensayo sobre las probabilidades de elección de los sujetos es pequeño, debido a esto, las predicciones del modelo en los puntos en los que cambió la matriz de pagos podrían ser conservadoras en comparación con las predicciones del modelo de IM.

En la parte inferior de la misma gráfica se presentan unas líneas verticales en color azul, estas líneas muestran los valores de la atracción inicial asociadas con los valores predichos de la variable θ que aparecen en la segunda gráfica de la misma figura.

En la segunda gráfica de la figura 3.15 se presenta una muestra aleatoria de los valores de la variable θ dado un valor de la atracción inicial. Estos valores son la predicción del modelo acerca de la probabilidad con la que los participantes elegirán su estrategia “Sol” a lo largo de los ensayos.

Como se muestra en la gráfica, las predicciones del modelo de CR respecto a la probabilidad con la que los jugadores elegirán su estrategia “Sol” son más conservadoras en comparación con las predicciones del modelo anterior. En este caso, el modelo predice que la probabilidad con la que los jugadores llevarán a cabo sus elecciones inicia cerca de la equiprobabilidad y que permanece prácticamente estable durante los primeros 100 ensayos.

Una vez que se presenta el primer cambio en la matriz de pagos el modelo predice que la probabilidad con la que los jugadores Columna elegirán su estrategia “Sol” incrementará a través del tiempo, mientras que en los

jugadores Fila el cambio se da en la dirección contraria, en términos generales este es el mismo comportamiento que muestran los demás modelos, sin embargo, este modelo predice que las probabilidades para ambos tipos de jugador no cambiarán demasiado, aunque continúan separándose conforme aumentan los ensayos.

Sin tomar en cuenta los valores predichos para la probabilidad por los modelos de CR e IM, ambos presentan características similares, es decir, ambos modelos muestran un incremento (decremento) en la probabilidad de elegir la estrategia “Sol” para los jugadores en el primer cambio después del cual muestran una menor sensibilidad a los cambios en las matrices de pagos.

En la figura 3.16 se presenta el modelo gráfico utilizado en el análisis del modelo SFP, en este caso el parámetro libre en el modelo es μ .

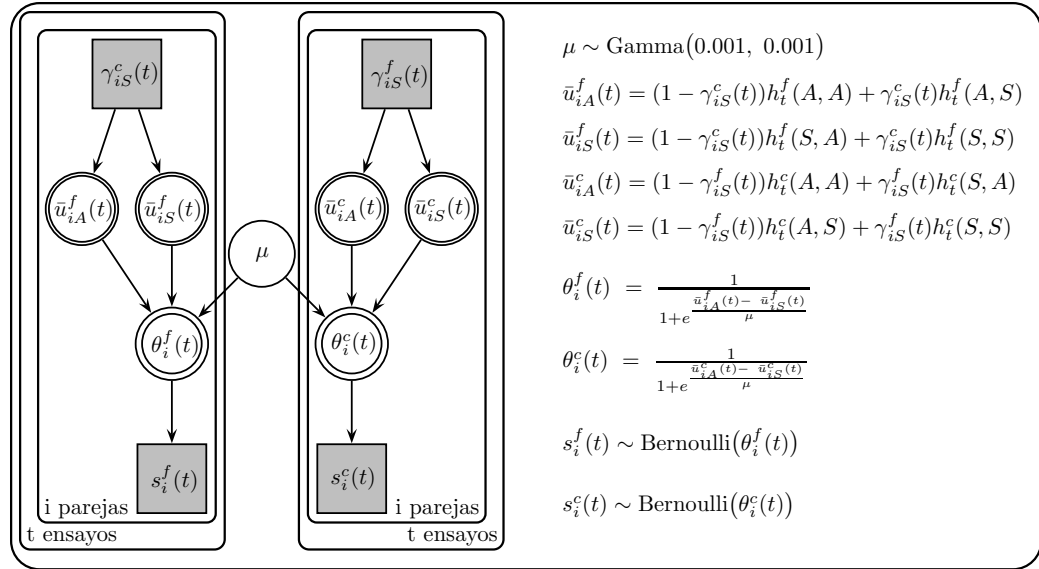


Figura 3.16: Representación gráfica del modelo de Juego Ficticio Estocástico (SFP; Fudenberg y Kreps, 1993). Las funciones completas asociadas con este modelo se encuentran en la sección 1.4.3.

Al parámetro μ se le asignó una distribución de probabilidad *a priori* Gamma con parámetros $\alpha = 0.001$ y $\beta = 0.001$; por otra parte, la función de verosimilitud que se utilizó en este modelo fue la misma que en el caso anterior, es decir, las elecciones de cada jugador se distribuyen Bernoulli.

En la primer gráfica de la figura 3.17 se muestra un histograma de los muestreos posteriores del parámetro μ del modelo de SFP; de acuerdo con estos muestreos, la media posterior del parámetro fue de 24.598. La línea horizontal que se muestra sobre el histograma representa el intervalo de credibilidad al 95 % el cual se encontró entre los valores de 21.178 para el límite inferior y 28.109 para el superior. Estos valores indican que los sujetos, en general, se encuentran más cercanos a la equiprobabilidad que a una solución óptima tomando en cuenta los valores esperados generados a través de la función de aprendizaje.

No obstante vale la pena destacar que el valor de este parámetro es dependiente de la magnitud de los valores o utilidades esperadas, por lo tanto, aún un mismo valor del parametro puede indicar distintos niveles de proximidad a la equiprobabilidad para diferentes sujetos, no obstante, esto dependerá de la historia de juego de cada uno.

En la segunda gráfica en la figura 3.17 se muestran las predicciones del modelo para distintos valores del parámetro μ , los valores asociados a estas predicciones aparecen marcados en el histograma mediante las líneas verticales.

Las predicciones que realiza este modelo difieren de las observadas en los dos modelos pertenecientes a la familia de modelos de Reforzamiento (CR e IM), en primer lugar por que presentan una menor variabilidad en todos los

Distribucion Posterior Y Prediccion del Modelo SFP

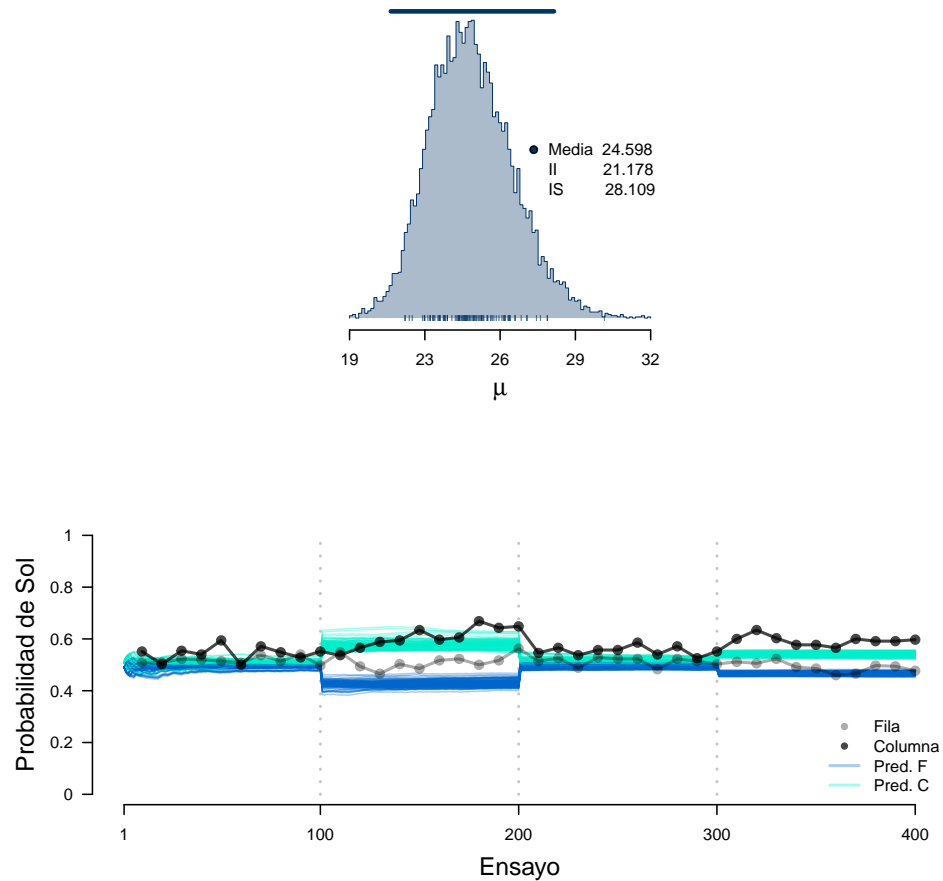


Figura 3.17: En la primer gráfica se muestra un histograma de los muestreos posteriores del parámetro μ . En la segunda se presentan los valores de la variable θ que se obtienen a partir de una muestra aleatoria del parámetro μ . Cada línea representa la predicción en términos de probabilidad para un jugador, en gris se encuentra el promedio de elecciones de la estrategia “Sol” de todos los jugadores Columna cada 10 ensayos, en negro se presenta la misma medida para los jugadores Fila.

juegos, es decir, las líneas se encuentran más cercanas entre sí; en segundo lugar, este modelo predice cambios abruptos y de baja magnitud, para las probabilidades con las que los jugadores elegirán su estrategia “Sol” después

de cada cambio en la matriz de pagos.

A diferencia del modelo de CR, este tipo de cambios en las probabilidades se asemejan al comportamiento observado en algunos de los participantes (*e.g.* ver figura 3.2, sección 3.1.2). Además, a diferencia del modelo de IM, en este caso, la proporción de elecciones de los sujetos en cada 10 ensayos normalmente se encuentra dentro de la misma zona que las predicciones del modelo, a excepción del juego 2 en el caso de los jugadores Fila y el juego 4 en el caso de los jugadores Columna.

En la figura 3.18 se presenta el modelo gráfico utilizado para el análisis del modelo de St-EWA. Como se mencionó en la sección 1.4.4, esta forma de aprendizaje combina supuestos de los modelos de Reforzamiento (CR e IM) con supuestos provenientes de los modelos Bayesianos (SFP).

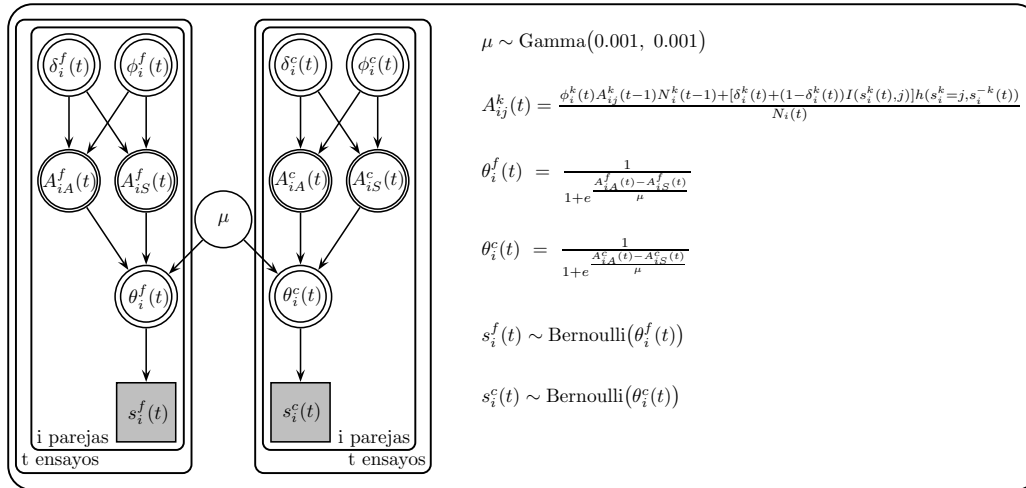


Figura 3.18: Representación Gráfica del modelo del Modelo Autorregulable de Atracciones Ponderadas por la Experiencia (St-EWA; Ho, Camerer y Chong, 2007). Las funciones completas asociadas con este modelo se encuentran en la sección 1.4.4.

Al igual que con el modelo de SFP, en éste se asignó una distribución de

probabilidad *a priori* Gamma al parámetro μ con un valor para α y β de 0.001. La función de verosimilitud tiene la misma forma que en los 2 modelos anteriores.

En la figura 3.19 se presenta un histograma con los muestreos de la distribución posterior del parámetro μ , así como las predicciones en términos de la probabilidad de cada tipo de jugador de elegir “Sol” a lo largo de los ensayos.

A diferencia del modelo anterior, en este caso, el valor de la media posterior para el parámetro μ fue menor, con un valor de 15.516 mientras que el límite inferior para el intervalo de alta densidad al 95 % fue de 13.569 y el superior de 17.656. Además de mostrar valores menores, el intervalo también tiene una menor longitud para este modelo (4.09) en comparación con el modelo de SFP.

Lo anterior sugiere que las elecciones de los participantes son menos aleatorias tomando en cuenta las atracciones generadas por el modelo de St-EWA en comparación con los valores esperados del modelo de SFP, sin embargo, como se mencionó anteriormente este parámetro se encuentra asociado al valor de las alternativas que se comparan (atracciones en un modelo y valores esperados en otro), por lo que sería necesario comparar dichos valores a lo largo de los ensayos para cada jugador de manera que sea posible saber si los valores del parámetro son comparables.

Por otra parte, este modelo presenta una mayor variabilidad en términos de sus predicciones en comparación con el anterior, aunque, en general, las tendencias son similares, es decir, el modelo predice una cercanía a la equiprobabilidad para el primero de los juegos y después del primer cambio en la

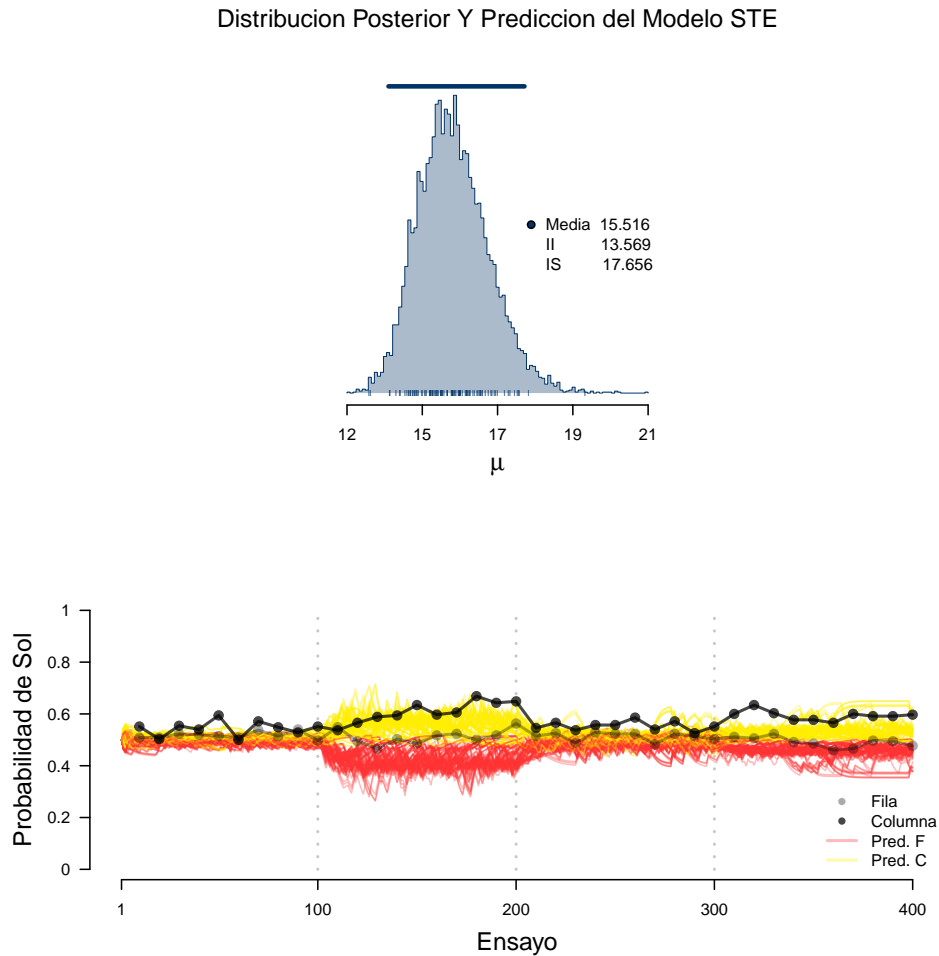


Figura 3.19: En la primer gráfica se muestra el histograma de los muestreos de la distribución posterior del parámetro μ . En la segunda se presentan los valores de la variable θ que se obtienen a partir de una muestra aleatoria del parámetro μ . Cada línea representa la predicción en terminos de probabilidad para un jugador, en gris se encuentra el promedio de elecciones de la estrategia “Sol” de todos los jugadores Columna cada 10 ensayos, en negro se presenta la misma medida para los jugadores Fila.

matriz de pagos (ensayo 101) predice un cambio pequeño en la dirección del equilibrio de Nash para los jugadores Columna y un cambio en la dirección opuesta para los jugadores Fila.

Los cambios predichos en la probabilidad utilizando este modelo, al igual que los modelos de Reforzamiento, son graduales, lo cual contrasta con los cambios abruptos predichos por el modelo anterior. Sin embargo, una diferencia importante es que este modelo aún muestra una sensibilidad a los siguientes cambios en las matrices de pagos, de los 3 modelos anteriores esto sólo ocurre con el modelo de SFP el cual pertenece a la familia de los modelos Bayesianos.

La proporción de elecciones de la estrategia “Sol” de los participantes permanece dentro de las zonas que predice el modelo, a excepción de los primeros ensayos en el juego “B2” en los cuales se subestima la probabilidad de que los jugadores Columna elijan su estrategia “Sol”. Además, en el juego “B1” el modelo subestima la probabilidad de que los sujetos Fila elijan esa misma estrategia.

Las comparaciones que se han mostrado hasta este punto son únicamente gráficas y no es posible extraer conclusiones claras a partir de las mismas, no obstante, permiten contrastar de manera cualitativa el comportamiento de los modelos y sus predicciones.

En la siguiente sección se presenta una comparación de estos modelos de aprendizaje a través del método utilizado para comparar modelos de equilibrio. Se presentará la comparación mediante una aproximación al factor de Bayes utilizando modelos gráficos Bayesianos.

3.3.2. Comparación de Modelos

En esta sección se presentaran los resultados de la comparación de modelos de aprendizaje; para comparar el desempeño de los modelos se utilizaron dos métodos, el primero es el uso de la función de puntuación cuadrática (*ver sección 1.4.4*), el segundo fue utilizar una aproximación al factor de Bayes mediante el uso de modelos gráficos Bayesianos.

En la figura 3.20 se presentará la puntuación cuadrática media para todos los jugadores y por separado para cada tipo de jugador para los 4 modelos de aprendizaje presentados anteriormente.

Los puntajes que aparecen en la gráfica de la figura 3.20 se obtuvieron utilizando las predicciones de los modelos que se presentaron en la sección anterior. La línea horizontal que aparece en la gráfica representa la puntuación obtenida por un modelo que supone que la respuesta en todos los ensayos es al azar ($\theta(t) = 0.5$). Los resultados se encuentran separados también para cada tipo de jugador.

Las primeras barras en la gráfica representan la puntuación asociada al modelo de IM, de los cuatro éste es el único cuyas puntuaciones no superan a las de un modelo aleatorio, por lo tanto, de acuerdo con este metodo es el modelo que presenta un peor ajuste al comportamiento de los participantes. Esto es un reflejo de lo que se observó en la figura 3.13, en este caso, el modelo sobreestima la probabilidad de que los jugadores Columna elijan su estrategia “Sol” y subestima la probabilidad de los jugadores Fila de elegir la misma estrategia, aunque en menor medida. Estas dos características de la relación entre las predicciones del modelo y las proporciones medias observadas, se

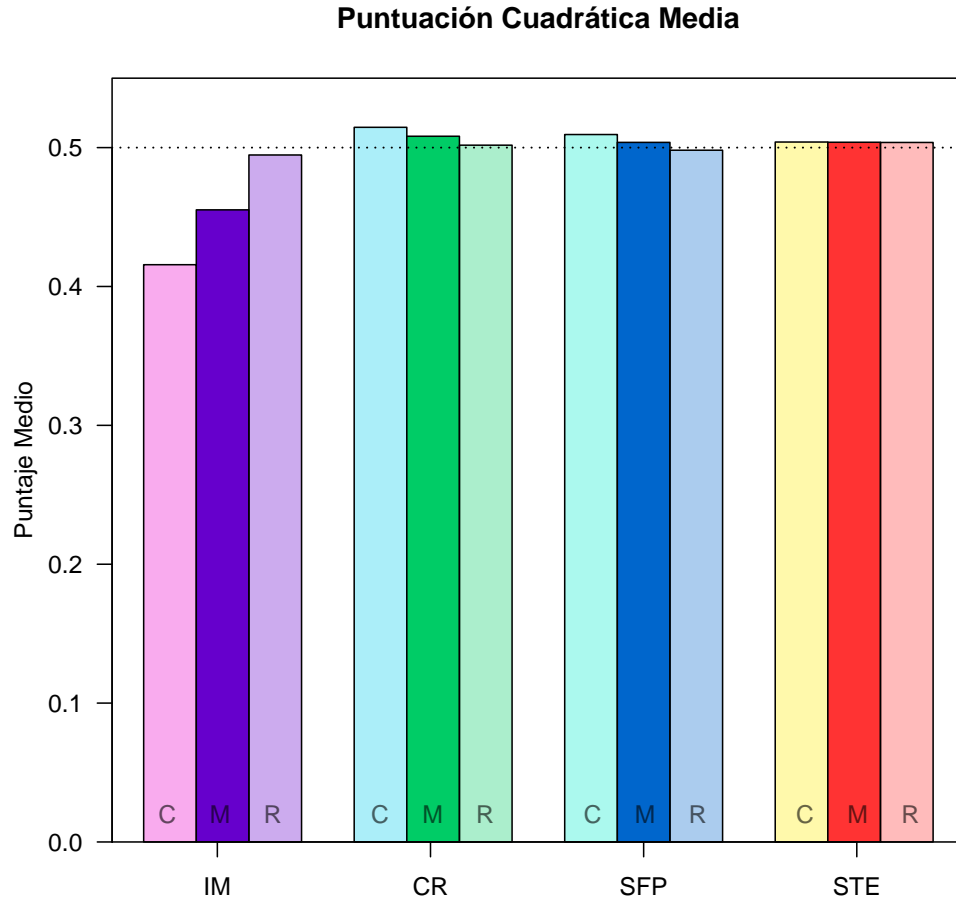


Figura 3.20: En la gráfica se muestra la media general de la puntuación cuadrática así como la media por tipo de jugador para cada modelo de aprendizaje.

reflejan en las puntuaciones cuadráticas del modelo, las cuales son mayores en el caso de los jugadores Columna en comparación con los jugadores Fila.

A diferencia del modelo de IM los modelos de CR y SFP muestran una puntuación similar entre sí, en primer lugar, ambos modelos tienen puntajes mayores a los que se esperaría de un modelo aleatorio para las 3 medidas.

Además, ambos modelos presentan un mayor puntaje en el caso de los jugadores Columna en comparación con el de los jugadores Fila. No obstante, ambos tienen diferentes predicciones tanto en magnitud como en varianza. Debido a que las puntuaciones se encuentran cercanas entre sí a pesar de estas diferencias en las predicciones de los modelos, utilizar una medida de puntuación no es suficiente para comparar estos modelos entre sí.

Por último, el modelo de St-EWA, también superó las puntuaciones de un modelo aleatorio en sus predicciones para el comportamiento de los jugadores Fila y Columna.

Para poder comparar estos modelos tomando en cuenta la variabilidad que presentan sus predicciones se utilizó el mismo método que en el caso de los modelos de equilibrio, es decir, se utilizó una aproximación al factor de Bayes mediante el uso de un modelo gráfico. En la figura 3.21 se presenta el modelo gráfico utilizado para la comparación de los modelos de aprendizaje que obtuvieron una puntuación mayor al azar.

El modelo en la figura 3.21 contiene el supuesto de que, a lo largo de los 400 ensayos, cada sujeto puede utilizar sólo 1 de 5 reglas posibles, es decir, los jugadores podrían estar eligiendo ya sea al azar ($z_i^k = 1$), siguiendo la regla del modelo de CR ($z_i^k = 2$), la de SFP ($z_i^k = 3$), la de St-EWA ($z_i^k = 4$) o una regla que permite asignar a cada dato su propio parámetro ($z_i^k = 5$). En este caso, la primer regla (elegir al azar) y la última (cada elección se se explica utilizando un parámetro diferente) se utilizaron para evaluar la probabilidad de los modelos de aprendizaje de una manera más general y no sólo en relación a los modelos propuestos en el experimento.

De la misma forma que en el caso de los modelos de equilibrio, el pará-

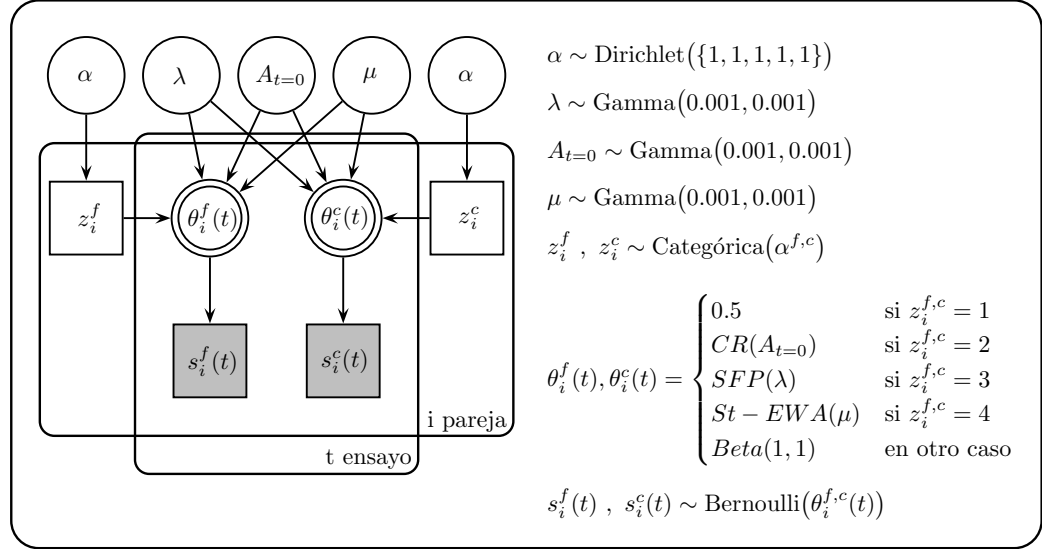


Figura 3.21: En la parte izquierda se presenta el modelo gráfico utilizado para la comparación de los 3 modelos de aprendizaje; del lado derecho se presentan las distribuciones a priori asignadas a cada parámetro en el modelo así como la función de verosimilitud.

metro α indica la probabilidad de que los participantes sigan una de estas 5 reglas, es decir, funciona como una aproximación a la probabilidad de cada modelo.

En las tablas 3.10 y 3.11 se muestran los resultados del análisis del modelo presentado en la figura 3.21. Los valores en la tabla representan la proporción de los muestreos de la distribución posterior en los que la variable z_i^k tomó los valores asociados a cada modelo.

En la tabla 3.10 es posible observar que para 19 de los 35 participantes en el rol de Fila el modelo con mayor probabilidad es el que predice que los participantes elegirán con igual probabilidad entre sus estrategias (RAN), lo cual representa el 54 % de la muestra, para este modelo los valores aproximados del factor de Bayes indican que este modelo es desde 11 % más probable

| Pareja | RAN | REL | SFP | STE | FULL | BF | Modelo 1 | Modelo 2 |
|--------|-------|-----|-------|-------|------|----------|----------|----------|
| R1 | 0.034 | 0 | 0.01 | 0.956 | 0 | 28.214 | STE | RAN |
| R2 | 0.667 | 0 | 0 | 0.332 | 0 | 2.009 | RAN | STE |
| R3 | 1 | 0 | 0 | 0 | 0 | 7999 | RAN | STE |
| R4 | 0.99 | 0 | 0 | 0.01 | 0 | 104.263 | RAN | STE |
| R5 | 0.942 | 0 | 0 | 0.058 | 0 | 16.391 | RAN | STE |
| R6 | 0.988 | 0 | 0 | 0.012 | 0 | 79.808 | RAN | STE |
| R7 | 0.976 | 0 | 0 | 0.024 | 0 | 40.885 | RAN | STE |
| R8 | 0.999 | 0 | 0 | 0.001 | 0 | 1332.333 | RAN | STE |
| R9 | 0.964 | 0 | 0.001 | 0.035 | 0 | 27.755 | RAN | STE |
| R10 | 0 | 0 | 0 | 1 | 0 | Inf | STE | RAN |
| R11 | 1 | 0 | 0 | 0 | 0 | Inf | RAN | REL |
| R12 | 0 | 0 | 0 | 1 | 0 | Inf | STE | RAN |
| R13 | 0 | 0 | 0.986 | 0.014 | 0 | 67.966 | SFP | STE |
| R14 | 0.005 | 0 | 0.131 | 0.864 | 0 | 6.605 | STE | SFP |
| R15 | 0.995 | 0 | 0 | 0.005 | 0 | 215.216 | RAN | STE |
| R16 | 0 | 0 | 0 | 1 | 0 | Inf | STE | RAN |
| R17 | 0.372 | 0 | 0 | 0.628 | 0 | 1.689 | STE | RAN |
| R18 | 0.948 | 0 | 0 | 0.052 | 0 | 18.321 | RAN | STE |
| R19 | 0.998 | 0 | 0 | 0.002 | 0 | 532.333 | RAN | STE |
| R20 | 1 | 0 | 0 | 0 | 0 | 7999 | RAN | STE |
| R21 | 0 | 0 | 0.913 | 0.087 | 0 | 10.527 | SFP | STE |
| R22 | 0 | 0 | 0 | 1 | 0 | 1999 | STE | RAN |
| R23 | 0.988 | 0 | 0 | 0.012 | 0 | 85.957 | RAN | STE |
| R24 | 0 | 0 | 0.371 | 0.628 | 0 | 1.692 | STE | SFP |
| R25 | 0.987 | 0 | 0 | 0.013 | 0 | 77.431 | RAN | STE |
| R26 | 1 | 0 | 0 | 0 | 0 | 2665.667 | RAN | STE |
| R27 | 1 | 0 | 0 | 0 | 0 | Inf | RAN | REL |
| R28 | 0.098 | 0 | 0.001 | 0.901 | 0 | 9.194 | STE | RAN |
| R29 | 0.003 | 0 | 0 | 0.997 | 0 | 346.826 | STE | RAN |
| R30 | 0 | 0 | 0 | 1 | 0 | 1999 | STE | RAN |
| R31 | 0.138 | 0 | 0 | 0.862 | 0 | 6.233 | STE | RAN |
| R32 | 0 | 0 | 0 | 1 | 0 | Inf | STE | RAN |
| R33 | 0.966 | 0 | 0 | 0.034 | 0 | 28.74 | RAN | STE |
| R34 | 0.531 | 0 | 0 | 0.469 | 0 | 1.133 | RAN | STE |
| R35 | 0.268 | 0 | 0 | 0.732 | 0 | 2.728 | STE | RAN |

Cuadro 3.10: Probabilidad posterior por modelo de aprendizaje en los jugadores Fila, en la columna “BF” se muestra una aproximación al factor de Bayes entre los dos modelos con mayor probabilidad.

que el segundo mejor modelo. En comparación, para los participantes en el rol de Columna (tabla 3.11), el mismo modelo fue el mejor sólo para el 25 %

| Pareja | RAN | REL | SFP | STE | FULL | BF | Modelo 1 | Modelo 2 |
|--------|-------|-----|-------|-------|------|---------|----------|----------|
| C1 | 0.86 | 0 | 0 | 0.14 | 0 | 6.149 | RAN | STE |
| C2 | 0.142 | 0 | 0 | 0.858 | 0 | 6.036 | STE | RAN |
| C3 | 0 | 0 | 1 | 0 | 0 | Inf | SFP | RAN |
| C4 | 0 | 0 | 1 | 0 | 0 | Inf | SFP | RAN |
| C5 | 0.001 | 0 | 0.8 | 0.198 | 0 | 4.037 | SFP | STE |
| C6 | 0 | 0 | 0.996 | 0.004 | 0 | 234.265 | SFP | STE |
| C7 | 0 | 0 | 1 | 0 | 0 | Inf | SFP | RAN |
| C8 | 0 | 0 | 1 | 0 | 0 | Inf | SFP | RAN |
| C9 | 0.978 | 0 | 0 | 0.022 | 0 | 45.512 | RAN | STE |
| C10 | 0 | 0 | 1 | 0 | 0 | Inf | SFP | RAN |
| C11 | 0.138 | 0 | 0 | 0.862 | 0 | 6.246 | STE | RAN |
| C12 | 1 | 0 | 0 | 0 | 0 | Inf | RAN | REL |
| C13 | 0 | 0 | 1 | 0 | 0 | Inf | SFP | RAN |
| C14 | 0.106 | 0 | 0 | 0.894 | 0 | 8.456 | STE | RAN |
| C15 | 0.157 | 0 | 0.002 | 0.842 | 0 | 5.374 | STE | RAN |
| C16 | 0.004 | 0 | 0 | 0.996 | 0 | 241.424 | STE | RAN |
| C17 | 0.888 | 0 | 0.023 | 0.09 | 0 | 9.904 | RAN | STE |
| C18 | 0.029 | 0 | 0.006 | 0.965 | 0 | 32.979 | STE | RAN |
| C19 | 0 | 0 | 0.99 | 0.01 | 0 | 104.263 | SFP | STE |
| C20 | 0.187 | 0 | 0 | 0.813 | 0 | 4.355 | STE | RAN |
| C21 | 0 | 0 | 0.994 | 0.006 | 0 | 176.778 | SFP | STE |
| C22 | 0.993 | 0 | 0 | 0.007 | 0 | 134.593 | RAN | STE |
| C23 | 0 | 0 | 0.982 | 0.018 | 0 | 55.338 | SFP | STE |
| C24 | 0.013 | 0 | 0 | 0.987 | 0 | 76.67 | STE | RAN |
| C25 | 0.007 | 0 | 0 | 0.993 | 0 | 134.593 | STE | RAN |
| C26 | 0.711 | 0 | 0 | 0.289 | 0 | 2.456 | RAN | STE |
| C27 | 0.224 | 0 | 0 | 0.776 | 0 | 3.464 | STE | RAN |
| C28 | 0.001 | 0 | 0 | 0.999 | 0 | 1598.8 | STE | RAN |
| C29 | 0.022 | 0 | 0 | 0.978 | 0 | 45.5 | STE | RAN |
| C30 | 0.991 | 0 | 0 | 0.009 | 0 | 114.942 | RAN | STE |
| C31 | 0 | 0 | 1 | 0 | 0 | Inf | SFP | RAN |
| C32 | 1 | 0 | 0 | 0 | 0 | Inf | RAN | REL |
| C33 | 0.92 | 0 | 0 | 0.08 | 0 | 11.52 | RAN | STE |
| C34 | 0.001 | 0 | 0.177 | 0.823 | 0 | 4.654 | STE | SFP |
| C35 | 0.482 | 0 | 0 | 0.518 | 0 | 1.074 | STE | RAN |

Cuadro 3.11: Probabilidad posterior por modelo de aprendizaje en los jugadores Columna, en la columna “BF” se muestra una aproximación al factor de Bayes entre los dos modelos con mayor probabilidad.

de los participantes.

Los modelos de SFP y St-EWA fueron el mejor modelo para 12 y 14 de

los participantes en el rol de Columna respectivamente, lo cual representa el 34 y 40 % de los participantes en dicho rol para cada modelo.

En el caso de los jugadores en el rol de Fila, el modelo de SFP resultó ser el mejor modelo para 2 de los 35 participantes (5 %) mientras que el modelo de St-EWA fue el mejor modelo para 14 de los participantes, lo cual equivale a un 40 %.

| | 1 | 2 | 3 | 4 | 5 |
|------|----|----|----|----|----|
| RAN | 28 | 21 | 5 | 0 | 16 |
| CR | 0 | 0 | 0 | 16 | 54 |
| SFP | 14 | 3 | 14 | 39 | 0 |
| STE | 28 | 31 | 11 | 0 | 0 |
| FULL | 0 | 15 | 40 | 15 | 0 |

Cuadro 3.12: Suma de las posiciones de cada modelo para los 70 participantes en el experimento de acuerdo a la frecuencia de la variable z_i^k .

En la tabla 3.12 se presenta el número de participantes para los cuales cada modelo terminó en una de las 5 posiciones, siendo 1 el más probable y 5 el menos probable de acuerdo con los muestreos posteriores de la variable z_i^k .

En la tabla se puede observar que los modelos RAN y St-EWA son los más probables para una mayor cantidad de los participantes en el experimento, con 40 % de la muestra para cada uno, sin embargo, el modelo de St-EWA resultó ser el segundo mejor para 31 de los 70 participantes (44 %) mientras que el modelo que predice que los sujetos elijen con la misma probabilidad entre sus opciones es el segundo mejor en sólo 30 % de los casos.

Por otra parte, los modelos FULL y CR no fueron el mejor modelo para ninguno de los participantes, no obstante el modelo completo (FULL) fue el

segundo para 15 de los 70 participantes (21%). Aun cuando el modelo de CR tiene la mayor puntuación cuadrática de los 3 modelos de aprendizaje, de acuerdo con este procedimiento, no es capaz de dar cuenta de los datos.

En la figura 3.22 se muestra la media posterior del parámetro α el cual funciona como una aproximación a la probabilidad de cada modelo dentro de la muestra.

Los números que aparecen en la parte inferior de la figura 3.22 indican el valor de la media posterior del parámetro α . De acuerdo con estos valores, el modelo más probable es el que supone que los sujetos eligen con igual probabilidad entre sus estrategias (RAN) con una probabilidad de 0.396, en segundo lugar se encuentra el modelo de St-EWA con una probabilidad de 0.371.

De estos resultados vale la pena destacar que el número de sujetos para los cuales el modelo de RAN resultó ser el mejor modelo dependió del tipo de jugador, ya que fue el mejor para el 54% de los participantes en el rol de Fila, mientras que para los jugadores Columna fue el mejor sólo para el 25% de los participantes.

El modelo gráfico que se presenta en la figura 3.23 tiene por objetivo comparar las probabilidades de cada uno de los modelos en función del rol de los sujetos, la diferencia principal con el modelo anterior, es que el parámetro α se encuentra separado para cada tipo de jugador. Las distribuciones *a priori* para cada uno de los modelos, así como la función de verosimilitud fueron las mismas que en el anterior.

En la figura 3.24 se presenta una gráfica de barras con los resultados de la media posterior del parámetro α para cada tipo de jugador por modelo.

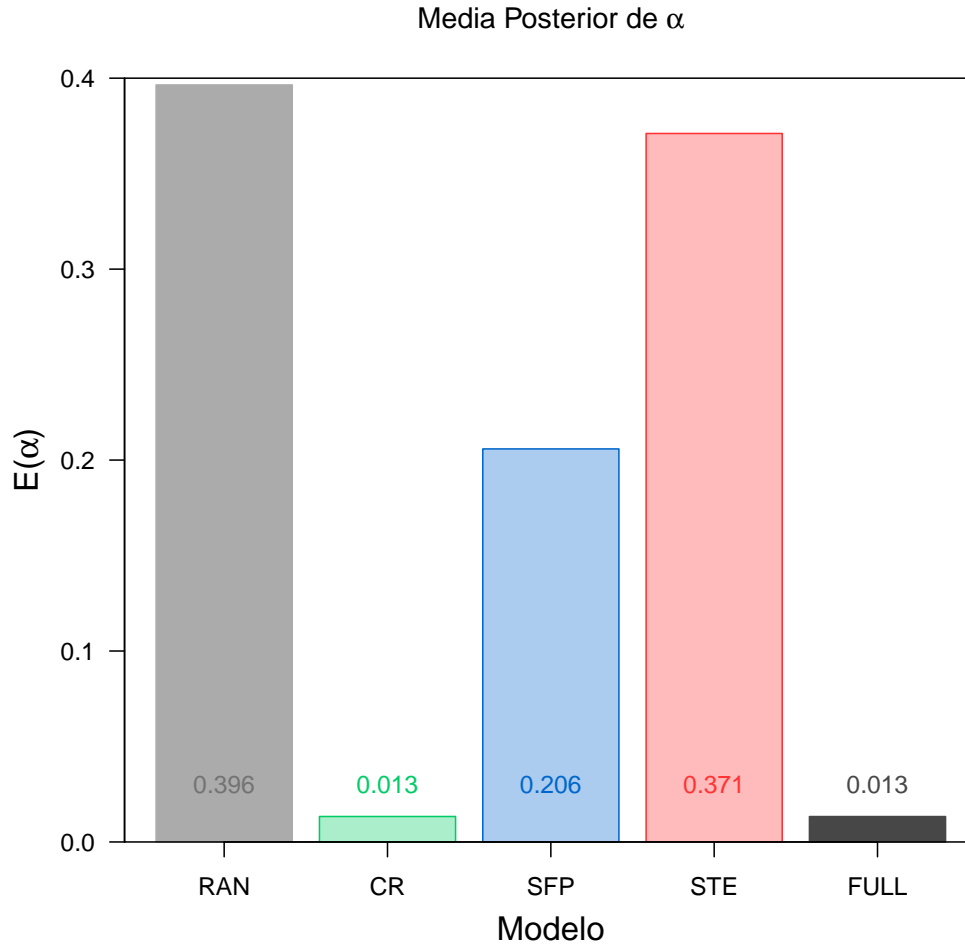


Figura 3.22: Las barras muestran la media de la probabilidad posterior (α) para cada uno de los modelos, en la parte baja de la gráfica se muestra numéricamente el valor obtenido para cada modelo.

Comenzando por los modelos de CR y el modelo completo (FULL), de acuerdo con la media posterior del parámetro α , estos modelos son los menos probables, con una media de 0.025 para ambos. De los modelos utilizados, estos son los únicos que no presentan diferencias en función del tipo de juga-

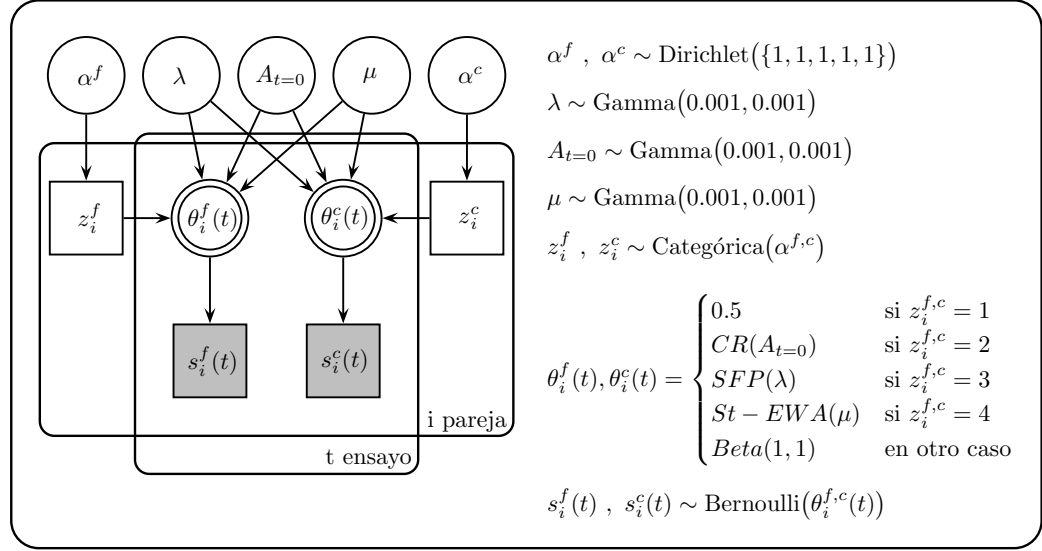


Figura 3.23: En la parte izquierda se presenta el modelo gráfico utilizado para la comparación de los 3 modelos de aprendizaje; del lado derecho se presentan las distribuciones a priori asignadas a cada parámetro en el modelo así como la función de verosimilitud.

dor.

En el caso del modelo de St-EWA tampoco se observan diferencias marcadas en su probabilidad dependiendo del tipo de jugador, sin embargo, la media es mayor en el caso de los jugadores Fila (0.373) en comparación con los jugadores Columna (0.368).

De los modelos que muestran mayores diferencias en función del rol del participante son el modelo RAN, el cual supone que los sujetos eligen con la misma probabilidad entre sus alternativas, y el modelo de SFP. Para el primero, la media posterior del parámetro α para los participantes en el rol de Fila fue de 0.504, en cambio, para los participantes Columna fue de 0.253.

En cambio, la media posterior del mismo parámetro asociada al modelo de SFP tuvo un comportamiento opuesto, es decir, el modelo es menos pro-

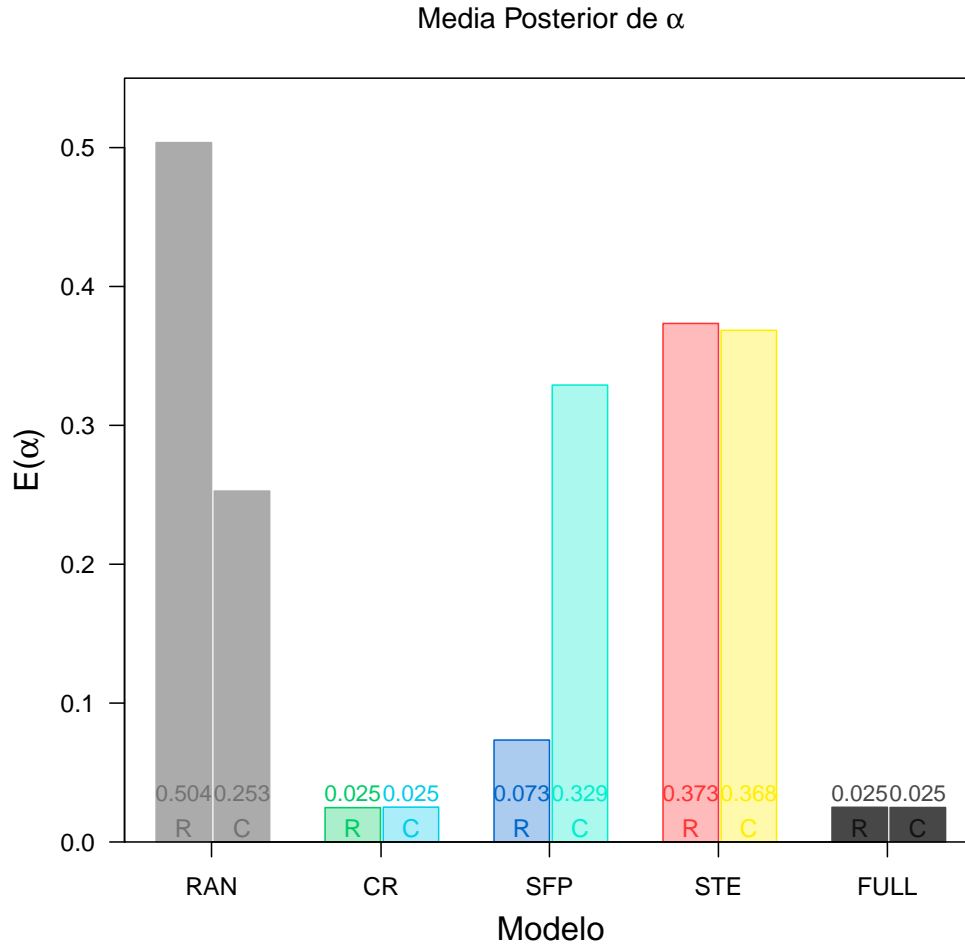


Figura 3.24: Las barras muestran la media de la probabilidad posterior (α) para cada uno de los modelos y tipos de jugador, en la parte baja de la gráfica se muestra numéricamente el valor obtenido para cada modelo.

bable en el caso de los jugadores Fila 0.073 en comparación con los jugadores Columna 0.329.

Para calcular las posibilidades posteriores de que la probabilidad de los modelos RAN y SFP difiera dependiendo del tipo de jugador se utilizó una

aproximación al factor de Bayes a través del método Savage-Dickey (Lee y Wagenmakers, 2013) el cual sirve para la comparación de modelos anidados. En este caso los modelos que se comparan son uno en el cual el valor de α asociado al modelo RAN es el mismo para ambos tipos de jugador con uno que en el que son diferentes. El caso es el mismo para el modelo de SFP.

De acuerdo con este método, la hipótesis de que el modelo RAN es más probable en los jugadores Fila en comparación con los jugadores columna es 7.43 veces más probable. Por otra parte, la hipótesis de que el modelo de SFP es más probable para los jugadores Columna que en los jugadores fila es 37.12 veces más probable que la hipótesis de que la probabilidad del modelo para ambos tipos de participantes es la misma.

Estos resultados muestran que el comportamiento de los participantes no sólo difiere en función del rol en el que se encuentran, lo cual va en contra de las predicciones del equilibrio de Nash, sino que los participantes difieren en la regla que usan para elegir entre las alternativas disponibles a lo largo del tiempo.

Discusión

El objetivo principal del presente experimento fue estudiar cómo cambia el comportamiento de participantes humanos cuando se enfrentan a una situación de conflicto en la cual los pagos por sus elecciones cambian a través del tiempo. Además, se utilizaron 4 modelos de equilibrio para comparar sus predicciones acerca del comportamiento general de los participantes bajo las 4 funciones de pagos en el experimento. Por último, se compararon las predicciones de cuatro modelos de aprendizaje con el comportamiento de los participantes a través del tiempo.

Al igual que en el experimento realizado por Ochs (1995), el comportamiento de los participantes mostró una gran heterogeneidad, no obstante, a diferencia de ese mismo estudio, se presentaron casos en los cuales el comportamiento de los participantes cambió en la dirección predicha por el equilibrio de Nash, lo cual no ocurrió con los participantes en el experimento mencionado. Este tipo de cambio se presentó principalmente en los participantes en el rol de Columna.

En cuanto al objetivo principal del experimento los datos de los participantes mostraron una gran heterogeneidad en términos de su respuesta a los cambios en las matrices de pago. Por ejemplo, se observaron parejas en

las cuales la probabilidad de elegir la estrategia “Sol” cambiaba en la dirección predicha por el equilibrio de Nash, mientras que otras parejas elegían de manera equiprobable entre sus alternativas a lo largo del experimento (*ver figura 3.2*).

De los experimentos que han utilizado cambios en las matrices de pagos sólo en el experimento de Sanabria y Thrailkill (2009) se habían observado resultados similares (cambios en la dirección predicha por el equilibrio de Nash), sin embargo, este tipo de comportamiento no había sido reportado con participantes humanos.

Las proporciones de elección de la estrategia “Sol” por parte de los participantes mostraron que en los juegos “A” el comportamiento de los sujetos se aproximaba al equilibrio de Nash el cual se encontraba en 0.6, en cambio, en los juegos “B” el comportamiento de los participantes mostró una mayor variabilidad. En estos juegos las proporciones de algunos de los participantes se alejaban de la equiprobabilidad en la dirección del equilibrio, no obstante, ninguna de las parejas mostro las proporciones predichas por el Equilibrio de Nash.

Para dar cuenta del comportamiento de los participantes en términos de la proporción de ocasiones en las que las parejas utilizaban su estrategia “Sol” en los 100 ensayos en los cuales estuvo activa cada matriz de pagos, se utilizaron 4 modelos de equilibrio los cuales fueron comparados a través de una aproximación al factor de Bayes utilizando métodos gráficos.

Una de las características de los modelos conductuales de equilibrio, es que, a diferencia del equilibrio de Nash, predicen una asimetría en el comportamiento de los jugadores, es decir, la predicción de los modelos difiere para

los jugadores en el rol de Columna y Fila. Este tipo de asimetría se observó en las proporciones de elección de las parejas en el experimento.

Al comparar los modelos de equilibrio para los 100 ensayos correspondientes a cada juego por separado, el modelo de “*Equilibrio en Respuestas Cuantales*” (QRE) resultó ser el modelo más probable para los 4 juegos en el experimento, con probabilidades de 0.531 en el juego “A1”, 0.436 en el juego “A2”, 0.608 y 0.553 en los juegos “B1” y “B2”, respectivamente.

Estos resultados concuerdan con lo que encontraron Brunner, Camerer y Goeree (2011; BCG), al reexaminar los resultados de Selten y Chmura (2008). De acuerdo con los autores originales (Selten y Chmura 2008; SC), el modelo de “*Equilibrio por Balance de Impulsos*” (IBE) era el que mejor daba cuenta del comportamiento de los participantes en diversos juegos, incluidos los juegos de suma constante¹. Sin embargo, al reexaminar los datos BCG encontraron que el modelo de QRE era mejor en este tipo de juegos.

Las comparaciones realizadas por BCG y por SC (2008) se basaron en una medida de distancia denominada “*Distancia Cuadrática Media*”. Esta misma medida fue utilizada en este experimento y los resultados son similares a los descritos por BCG. En primer lugar, la distancia media fue menor en los 4 juegos para el modelo de QRE en comparación con el modelo de IBE, en segundo lugar, el modelo de “*Equilibrio por Muestreo de Pagos*” (PSE) presentó una media menor a la del modelo QRE para los juegos “B”.

Si las comparaciones entre modelos se realizaran usando únicamente este estadístico existe la posibilidad de que los resultados fuesen los mismos a los reportados en BCG, sin embargo, los resultados de la comparación de

¹Los juegos del presente experimento son una instancia de estos últimos.

modelos utilizando una aproximación al factor de Bayes mediante modelos gráficos difieren ligeramente de los resultados de la comparación de la distancia cuadrática.

Los resultados de las comparaciones mediante modelos gráficos mostraron que, el modelo más probable en los 4 juegos fue el de QRE, mientras que el modelo de PSE fue el segundo mejor modelo sólo en 3 de los cuatro juegos en el experimento.

Las distribuciones posteriores predictivas de los modelos de QRE y PSE (ver figuras 3.8 y 3.5) son similares entre sí, sin embargo, las del modelo de QRE se encuentran más cercanas a los datos en los juegos “A” en comparación a las del modelo de PSE y al mismo tiempo, son muy similares en el caso de los juegos “B”.

El mejor ajuste del modelo QRE implica que, la mejor forma de describir el comportamiento de los participantes es a través del supuesto de que la preferencia de los sujetos por sus alternativas es una variable aleatoria (McKelvey y Palfrey, 1995).

En cuanto a los modelos de aprendizaje, en el presente experimento se utilizaron 4 modelos provenientes de dos familias diferentes, los modelos de “*Reforzamiento Acumulado*”(CR) y “*Aprendizaje por Igualación de Impulsos*” (IM) que pertenecen a la familia de los modelos de reforzamiento, el de “*Juego Ficticio Estocástico*” que pertenece a la familia de los modelos bayesianos y por último el modelo de “*Aprendizaje Auto Regulable de Atracciones Ponderadas por la Experiencia*” (St-EWA) el cual es un híbrido entre ambos tipos de familias de modelos.

A excepción del modelo de SFP, estos modelos han sido comparados en

una sola ocasión (Chmura, Goerg y Selten, 2012), en la cual se encontró que el modelo de St-EWA era el que mejor daba cuenta del comportamiento de los participantes a través del tiempo seguido por el modelo de IM.

En el presente experimento se extrajeron primero las predicciones de los modelos para compararlas con la proporción media de elecciones de los participantes. Los resultados mostraron que, en primer lugar, el modelo de IM sobre estimaba la magnitud del cambio en el comportamiento de los sujetos cuando se presentaba el primer cambio en la matriz, para después mostrar una baja sensibilidad en los siguientes cambios.

El modelo de CR mostró un comportamiento diferente, en primer lugar, el modelo predice que, después de un cambio, el comportamiento de los participantes cambiará de manera gradual, en segundo lugar, predice que la proporción de elecciones de los participantes permanecerá cerca de la equiprobabilidad la mayor parte del experimento (ensayos 1 a 300). Al igual que el modelo de IM éste también mostró una baja sensibilidad a los cambios subsecuentes en el experimento

La baja sensibilidad de estos modelos a los cambios en las matrices de pagos se relaciona con uno de los supuestos que subyacen a ambos, en particular se relaciona con el supuesto de que las preferencias de los sujetos se forman a través de la acumulación de la información acerca del desempeño de las estrategias, es decir, la baja sensibilidad a los cambios se debe a la ley de poder de la práctica, de acuerdo a la cual el impacto de las experiencias más recientes se reduce en función de la experiencia que se ha acumulado.

Por otra parte, el modelo de SFP presentó un comportamiento diferente. Para este modelo, la predicción es que la proporción de elecciones de los

participantes se encontrará cercana a la equiprobabilidad para los juegos “A” y cambiará de manera abrupta al presentarse un cambio en la matriz de pagos; en el caso de los jugadores en el rol de Columna el modelo predice que la proporción de elecciones de la estrategia “Sol” incrementará cuando se presenta un cambio hacia los juegos “B”. En cambio, el modelo predice que la proporción con la que los jugadores Fila elegirán la misma estrategia se reduce para esos juegos. En otras palabras, el modelo predice un cambio en la dirección del equilibrio de Nash para los jugadores Columna y, al mismo tiempo, un cambio en la dirección opuesta para los jugadores Fila.

Por último, las predicciones del modelo de St-EWA acerca del comportamiento de los participantes muestran características similares a las de los modelos de Reforzamiento y a las del modelo de SFP. Por un lado, de manera similar a la predicción del modelo de SFP, St-EWA predice que el comportamiento de los participantes cambiará cuando se presente un cambio en la matriz de pagos, por otra parte, la predicción del modelo es que este cambio es gradual, lo cual es similar a la predicción en los modelos de reforzamiento.

Al igual que en los modelos de aprendizaje, a nivel agregado, los datos muestran diferentes tendencias dependiendo del tipo de jugador, en el caso de los jugadores Fila, las proporciones de elecciones de la estrategia sol para estos jugadores se mantiene relativamente estable y en valores cercanos al 0.5, en cambio, el comportamiento de las proporciones de elección de la estrategia “Sol” por parte de los jugadores Columna parece cambiar gradualmente cuando se presenta un cambio en la matriz de pagos. Además, estas proporciones de elección muestran una sensibilidad a todos los cambios y no sólo al inicial como lo predicen los modelos de Reforzamiento.

Estos resultados concuerda con las predicciones del modelo de St-EWA, al menos en el caso del comportamiento de los jugadores Columna, sin embargo, vale la pena destacar que el cambio gradual observado en las proporciones de elecciones de los participantes puede deberse a que los datos se encuentran agregados.

Para comparar el desempeño de las predicciones de los modelos se utilizó inicialmente una medida de la “correspondencia” entre las predicciones del modelo y las elecciones de los sujetos, esta medida se conoce como “*Puntuación Cuadrática*” (PC) y fue presentada por Brier (1950). De acuerdo con la media de la PC el modelo de IM fue el único cuyas predicciones no superaron al azar (0.5), en cambio, los modelo de CR, SFP y St-EWA lo superaron con un margen pequeño.

Además, de acuerdo con la PC obtenida para cada modelo y para cada tipo de participante, se encontró que los modelos de CR y SFP presentaban un mayor puntaje en sus predicciones para los jugadores Columna en comparación con sus predicciones respecto a los jugadores Fila.

Para comparar los modelos que habían superado la puntuación que puede ser recibida por un modelo que supone que los participantes eligen al azar en cada ensayo (aunque fuese por un margen reducido) se utilizó el mismo método que en el caso de los modelos de equilibrio, sin embargo, en esta ocasión se utilizaron dos modelos adicionales para poder comparar la eficiencia de los modelos de una forma más general. De Los modelos utilizados fueron uno supone que los participantes eligen de manera equiprobable (0.5) entre sus alternativas, sin importar los pagos del juego, mientras que el segundo utiliza un parámetro diferente para cada elección de los participantes.

De acuerdo con este método, el modelo que resultó más probable fue el modelo que supone que los sujetos responden de manera equiprobable con una probabilidad de 0.396 en la muestra, el segundo fue el de St-EWA con una probabilidad de 0.371. No obstante, al revisar que modelo era mejor para los participantes se observó una discrepancia entre los jugadores en el rol de Fila y Columna, en este caso se encontró que el modelo aleatorio era el mejor para el 54 % de los participantes en el rol de Fila, pero lo mismo ocurría sólo para el 11 % de los jugadores en el rol de Columna.

De manera que fuera posible comparar si la probabilidad de los modelos dependía del rol del participante, se separaron los parámetros que controlaban la probabilidad de cada modelo para cada tipo de jugador. Esto resulta importante ya que, como se mencionó anteriormente, los modelos de aprendizaje predicen un comportamiento diferente para cada tipo de jugador.

Al realizar el análisis con este modelo se encontró que era aproximadamente 7 veces más probable que la probabilidad del modelo aleatorio fuera mayor para los jugadores Fila que para los jugadores Columna. Mientras que resultó 37 veces más probable que la probabilidad del modelo de SFP fuera mayor en el caso de los jugadores Columna en comparación con los jugadores Fila. Los otros modelos no mostraron una diferencia en su probabilidad en función de cada tipo de participante.

Este resultado es de importancia en el presente experimento ya que los modelos que resultaron ser los más probables para los jugadores en el rol de Columna fueron aquellos que predicen que, cuando cambian las matrices de pagos, el comportamiento de los jugadores cambiará en la dirección predicha por el equilibrio de Nash aunque en una magnitud menor.

A partir de estos resultados, y al igual que en los resultados presentados por Chmura, Goerg y Selten (2012), es posible afirmar que, de los modelos de aprendizaje comparados en el presente experimento el que más se asemeja al comportamiento que presentaron los participantes de ensayo a ensayo es el de St-EWA.

El presente experimento tiene dos limitaciones importantes, la primera, es que para los participantes no existía una relación clara entre las ganancias que podían obtener ensayo a ensayo y una recompensa al finalizar el experimento, por lo tanto, los participantes podrían estar menos motivados durante el experimento.

La segunda se relaciona con los pagos utilizados en los juegos, en este caso, debido a que los 4 juegos utilizados son una instancia del juego de MP, los participantes podrían dejar de poner atención a los pagos recibidos y concentrarse únicamente en ganar o perder, esto podría explicar por qué algunas de las parejas eligen de forma equiprobable en todos los juegos del experimento. El mayor problema en este caso, es que no es posible discriminar entre ésta y otras explicaciones en el caso de parejas que eligen entre sus alternativas con una probabilidad cercana a 0.5.

De acuerdo con el mejor modelo de aprendizaje, el ajuste de los sujetos ante un cambio en la matriz de pagos depende del rol en el que se encuentren. En este caso, de los jugadores Fila se espera que la proporción con la que eligen su estrategia “Sol” permanezca cercana a la equiprobabilidad, o que la probabilidad con la que eligen entre sus estrategias sea menor cuando se presenta un cambio en la matriz de pagos. En cambio, de los sujetos en el rol de Columna el modelo indica que la probabilidad con la que eligen su

estrategia “Sol” cambia en la dirección predicha por el equilibrio de Nash, sin embargo la magnitud del cambio es menor a la que se espera de acuerdo con este modelo.

Conclusión

Una de las principales propiedades de los ambientes es que cambian a través del tiempo, esto no es diferente para las situaciones de interacción, en este tipo de situaciones siempre puede presentarse algún evento que cambie las consecuencias del comportamiento de los sujetos involucrados.

Este experimento ha buscado responder a la pregunta de si los sujetos son capaces de adaptar su comportamiento a las nuevas consecuencias que tienen sus elecciones en una situación de interacción, mostrando que, en general, las personas son capaces de ajustar su comportamiento cuando se presenta un cambio de esta naturaleza. Sin embargo, este cambio no siempre tiene la dirección o magnitud que se esperaría de un agente racional.

En este sentido, el comportamiento de los participantes en el presente experimento presentó una gran heterogeneidad, con algunas parejas que no presentaban una respuesta ante el cambio, mientras que otras respondían en la dirección esperada de acuerdo con el equilibrio de Nash.

Se ha mostrado también que, utilizando supuestos acerca la información con la que cuentan las personas, es posible dar cuenta del comportamiento de los sujetos de manera general aislando cada una de las situaciones con las que se encuentran a lo largo del tiempo aún cuando han tenido experiencia

con situaciones similares. Es decir, aun cuando las personas cuentan con experiencia es posible aproximar su comportamiento asumiendo que el mismo se dirige a un mismo punto para todos, un punto que es posible denominar como equilibrio.

Además, los resultados sugieren que no sólo es posible generar predicciones acerca del comportamiento de las personas asumiendo que todos llegaran al mismo punto, sino que es posible dar cuenta de la forma en la que sus elecciones cambiarán a través del tiempo, esto utilizando supuestos acerca del tipo de información que las personas utilizan para tomar una decisión. Los resultados obtenidos a partir de la comparación de los modelos de aprendizaje muestran que las personas podrían estar combinando la información que tienen de su entorno y de lo que han recibido a consecuencia de sus acciones a lo largo del tiempo para llevar a cabo sus elecciones.

Si bien los efectos de las consecuencias de las acciones de los organismos sobre su comportamiento a futuro han sido ampliamente estudiados, el efecto de la estructura del entorno no ha recibido el mismo tipo de atención. Por lo tanto, el estudio de la estructura y el "comportamiento" del entorno podría ser fundamental para la comprensión del comportamiento de los organismos no sólo cuando se estudian entornos que cambian a través del tiempo, sino también cuando se estudia el comportamiento en ambientes que se mantienen constantes. Algunos de los resultados de este experimento indican que el papel que puede jugar la estructura del entorno (que en el caso particular de este experimento era otra persona), puede ser igual de importante que las consecuencias de las elecciones de los organismos.

Por último, el presente experimento y los resultados del mismo abren el

paso a nuevas preguntas de investigación que podrían resultar de interés, por ejemplo, si existe una relación entre las consecuencias con las que se puede encontrar un organismo en una situación de interacción y el tipo de información en el cual los organismos basan sus elecciones; o si incorporar supuestos acerca de cómo los sujetos perciben las consecuencias de sus elecciones puede ayudar a mejorar las predicciones de un modelo acerca del comportamiento de una población.

En general, el presente experimento ha mostrado que los organismos son capaces de ajustar su comportamiento cuando existen cambios en las consecuencias de sus acciones aún en ambientes tan complejos como los de una situación de interacción y que, además, existen supuestos teóricos acerca de la información que utilizan los organismos para llevar a cabo sus elecciones que permiten predecir la dirección y magnitud de dichos ajustes.

Referencias

- Bailey, J. T., y Mazur, J. E. (1990). *Choice behavior in transition: development of preference for the higher probability of reinforcement*. Journal of the Experimental Analysis of Behavior, 53(3), 409-422.
- Beggs, A. W. (2005). *On the convergence of reinforcement learning*. Journal of Economic Theory, 122(1), 1-36.
- Behrens, T., Woolrich, M., Walton, M. y Rushworth, M. (2007). *Learning the Value of Information in an Uncertain World*. Nature Neuroscience, 10: 1214-1221.
- Brier, G. (1950). *Verification of Forecasts Expressed in Probabilistic Terms of Probability*. Monthly Weather Review, (78), 1-3.
- Brown, J. N., y Rosenthal, R. W. (1990). *Testing the Minimax Hypothesis: A Re-Examination of O'Neill's Game*. Econometrica, 58(5), 1065-1081.
- Brunner, C., Camerer, C. F., y Goeree, J. K. (2011). *Stationary concepts for experimental 2 x 2 games: Comment*. American Economic Review, 101(2), 1029-1040.
- Camerer, C. (2003). *Behavioral Game Theory: Experiments in strategic interaction*. New Jersey: Princeton University Press.

- Camerer, C. y Ho, T-H. (1998). *Experience-Weighted Attraction Learning in Coordination Games: Probability Rules, Heterogeneity, and Time-Variation*. New Jersey: Princeton University Press.
- Camerer, C. y Ho, T-H. (1999). *Experience-Weighted Attraction Learning in Normal Form Games*. *Econometrica*, 67, 827-874.
- Chmura, T., Goerg, S. J., y Selten, R. (2012). *Learning in experimental 2x2 games*. *Games and Economic Behavior*, 76(1), 44-73.
- Eliasz, K., y Rubinstein, A. (2011). *Edgar Allan Poe's riddle: Framing effects in repeated matching pennies games*. *Games and Economic Behavior*, 71(1), 88-99.
- Erev, I., y Roth, A. (1998). *Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria*. *American Economic Review*, 88(4), 848-881.
- Erev, I., Roth, A., Slonim, R., y Barron, G. (1998). *Learning and Equilibrium as Useful Approximations: Accuracy of Prediction on Randomly Selected Constant Sum Games*. *Economic Theory*, 33, 29-51.
- Fernández i Marín, X. (2016). *ggmcmc: Analysis of MCMC Samples and Bayesian Inference*. *Journal of Statistical Software*, 70(9), 1-20.
<doi:10.18637/jss.v070.i09>
- Fudenberg, D., y Kreps, D. M. (1993). *Learning mixed equilibria*. *Games and Economic Behavior*, 5, 320-367.
- Fudenberg, D., y Levine, D. (2009). *Learning and Equilibrium*. *Annual Review of Economics*, 1, 385-419.
- Gallistel, C., Mark, T., King, A. y Latham, P. (2001). *The Rat Approximates*

- an Ideal Detector in Changes in Rates of Reward: Implications for the Law of Effect*. *Journal of Experimental Psychology: Animal Behavior Processes*, 27(4), 354-372.
- Goeree, J. K., Holt, C. A., y Palfrey, T. R. (2003). *Risk averse behavior in generalized matching pennies games*. *Games and Economic Behavior*, 45(1), 97-113.
- Hasselman, B. (2015). *nleqslv: Solve Systems of Nonlinear Equations*. R package version 2.9.
<http://CRAN.R-project.org/package=nleqslv>
- Ho, T., y Camerer, C. (1999). *Experience-Weighted Attraction Learning in Normal Form Games*. *Econometrica*, 67(4), 827-874.
- Ho, T., Camerer, C., y Chong, J. K. (2007). *Self-Tuning Experience Weighted Attraction Learning in Games*. *Journal of Economic Theory*, 133(1), 177-198.
- Ho, T., Wang, X., y Camerer, C. (2007). *Individual Differences in EWA Learning With Partial Payoff Information*. *The Economic Journal*, 37-59.
- Kass, R. R. E., y Raftery, A. E. A. (1995). *Bayes Factors*. *Journal of the American Statistical Association*, 90(430), 773- 795.
- Lee, M. y Wagenmakers, E-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.
- Mark, T. y Gallistel, C. (1994). *Kinetics of Matching*. *Journal of Experimental Psychology. Animal Behavior Processes*, 20(1), 79-95.
- Mazur, J. E. (1992). *Choice behavior in transition: development of preference*

- with ratio and interval schedules*. Journal of Experimental Psychology. Animal Behavior Processes, 18(4), 364-378.
- Mazur, J. E. (1997). *Effects of rate of reinforcement and rate of change on choice behaviour in transition*. The Quarterly Journal of Experimental Psychology. B, Comparative and Physiological Psychology, 50(2), 111-128.
- Mookherjee, D. y Sopher, B. (1994). *Learning Behavior in an Experimental Matching Pennies Game*. Games and Economic Behavior, 7, 62-91.
- Mckelvey, R. D., y Palfrey, T. R. (1995). *Quantal Response Equilibria for Normal Form Games*. Games and Economic Behavior, 10, 6-38.
- Nash, J. (1950). *Equilibrium Points in N-Person Games*. En H. W. Kuhn y S. Nasar, The Essential John Nash (2002). Princeton, New Jersey: Princeton University Press.
- Ochs, J. (1995). *Games With Unique Mixed Strategy Equilibria: An Experimental Study*. Games and Economic Behavior, 10, 202-217.
- O'Neill, B. (1987). *Nonmetric test of the minimax theory of two-person zero-sum games*. Proceedings of the National Academy of Sciences of the United States of America, 84(7), 2106-2109.
- O'Reilly, J. X. (2013). *Making predictions in a changing world: inference, uncertainty, and learning*. Frontiers in Neuroscience, 7: 1-10.
- Osborne, M. J., y Rubinstein, A. (1998). *Games with Procedurally Rational Players*. American Economic Review, 88(4), 834-847.
- Plummer, M. (2015). *rjags: Bayesian Graphical Models using MCMC*. R package version 3-15.

<http://CRAN.R-project.org/package=rjags>

Plummer, M., Best, N., Cowles, K. y Vines, K. (2006). *CODA: Convergence Diagnosis and Output Analysis for MCMC*, R News, vol 6, 7-11

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.

Sanabria, F., y Thrailkill, E. (2009). *Pigeons (Columba livia) approach Nash equilibrium in experimental Matching Pennies competitions*. Journal of the Experimental Analysis of Behavior, 91(2), 169-183.

Selten, R. (1998). *Axiomatic Characterization of the Quadratic Scoring Rule*. Experimental Economics, (1), 43-62.

Selten, R., Abbink, K., y Cox, R. (2005). *Learning direction theory and the winner's curse*. Experimental Economics, 8(1), 5-20.

Selten, R., y Chmura, T. (2008). *Stationary concepts for experimental 2x2 Games*. American Economic Review, 98(3), 938-966.

Shachat, J. M. (2002). *Mixed Strategy Play and the Minimax Hypothesis*. Journal of Economic Theory, 104(1), 189-226.

Shoham, Y., Powers, R., y Grenager, T. (2007). *If Multi-Agent Learning is The Answer, What is The Question?*. Artificial Intelligence, 171(7), 365-377.

Staddon, J. (2001). *Adaptive Dynamics*. Cambridge, Massachusetts: The MIT Press.

Tversky, A., y Kahneman, D. (1992). *Advances in Prospect Theory: Cumulative Representation of Uncertainty*. Journal of Risk and Uncertainty,

(5), 297-323.

Von Neumann, J. (1928). *On the Theory of Games of Strategy*. En A. W.

Tucker y R. D. Luce (1959), *Contributions to the theory of games*.

Princeton, New Jersey: Princeton University Press.

Von Neumann, J., y Morgenstern, O. (1945). *Theory of Games and Economic*

Behavior. Princeton, New Jersey: Princeton University Press.

Young, H. P. (2002). *Bounded Rationality and Learning: On The Limits to*

Rational Learning. *European Economic Review*, 46, 791-799.

Young, P. (2004). *Strategic Learning and its limits*. New York: Oxford Uni-

versity Press.

Young, H. P. (2007). *The Possible and the Impossible in Multi-Agent Lear-*

ning. *Artificial Intelligence*, 171(7), 429-433.

Yu-Sung, S. y Yajima, M. (2015). *R2jags: Using R to Run 'JAGS'*. R package

version 0.5-7.

<http://CRAN.R-project.org/package=R2jags>