



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

---

---

FACULTAD DE PSICOLOGÍA  
DIVISIÓN DE ESTUDIOS PROFESIONALES

**ESTUDIOS CON EL PROBLEMA DE  
MONTY HALL**

T E S I S

QUE PARA OBTENER EL GRADO DE:  
**LICENCIADO EN PSICOLOGÍA**

PRESENTA:  
**JOSÉ LUIS BAROJA MANZANO**

DIRECTOR DE TESIS:  
**DR. ARTURO BOUZAS RIAÑO**

REVISOR:  
**DR. GERMÁN PALAFOX PALAFOX**

SINODALES:  
**DR. FLORENTE LÓPEZ RODRÍGUEZ**  
**DR. VLADIMIR ORDUÑA TRUJILLO**  
**DR. ÓSCAR ZAMORA ARÉVALO**



CIUDAD DE MÉXICO

MARZO, 2015



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.



**ESTUDIOS CON EL PROBLEMA DE  
MONTY HALL**

por

JOSÉ LUIS BAROJA MANZANO

Tesis presentada para obtener el grado de

LICENCIADO EN PSICOLOGÍA

en la

FACULTAD DE PSICOLOGÍA

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Ciudad de México. Marzo, 2015





*A la memoria de Emmanuel*



# Agradecimientos

Este trabajo probablemente no hubiera llegado a término si no hubiera tenido un comienzo. Pero incluso el comienzo hubiera servido de poco sin las porras y el apoyo de Germán Palafox y Emmanuel Guzmán durante las etapas iniciales. De alguna manera que aún no entiendo, en alguna de las pláticas con Chet Baker o King Crimson de fondo, lograron convencerme de que el proyecto valía la pena.

Ya encarrerado no perdí el impulso y eso se lo debo a más de una persona. Eduardo Parra fue el primero en escucharme hablar sobre el Problema de Monty Hall y también el primero en mostrarme, como siempre ha sido su costumbre, cuán interesante puede ser una idea sencilla. Tuve la fortuna de contar con la compañía de Alejandra González en más de una sesión de recolección de datos. Los consejos y comentarios de Michael Lee y Manuel Villarreal me ayudaron a comprender y mejorar diversos aspectos del proceso de inferencia. Darío Trujano ha sido mi apoyo principal para entender modelos de reforzamiento y elección estocástica, y agradezco esa capacidad suya para encontrar la manera de explicar las cosas clara y detalladamente, encima de mis comentarios sarcásticos e impacientes. Elena Villalobos recibió el primer manuscrito del trabajo; las razones por las que el texto final es mucho mejor que aquel borrador accidentado y confuso fueron sus sugerencias y correcciones. El apoyo y presión paciente pero constante de Vladimir Orduña y Alejandro Segura resultaron cruciales para no olvidar y tener presente en todo momento que había un final que alcanzar, que no se valía doblarse.

En la Facultad trabaja el mejor equipo de investigación que conozco, repartido en varios laboratorios, y durante todo el proyecto, y antes, y después, ha sido un privilegio formar parte de él.

Debo mi formación profesional, y también buena parte de la otra, al Dr. Arturo Bouzas. Siempre, pero en especial ahora, he tenido ganas de preguntarle cuánto he cambiado desde que nos conocimos. (Él sigue igualito.)

A ustedes, a mi familia y a mis amigos: No sé cómo agradecerles.



# Índice general

1. Introducción y Panorama General	1
2. El Problema de Monty Hall	3
3. Los Modelos	9
4. Recolección de Datos	19
5. Resultados	23
6. Discusión	55
7. Conclusiones	67
8. Referencias	69



# Estudios con el Problema de Monty Hall

por

José Luis Baroja Manzano

## Resumen

Este trabajo evalúa la capacidad de tres modelos de aprendizaje por reforzamiento para explicar el desempeño humano en una versión repetida del Problema de Monty Hall (PMH).

Reportes recientes sugieren que dichos modelos proporcionan explicaciones razonables sobre la manera en la que participantes humanos aprenden sobre y explotan alternativas inciertas a medida que tienen contacto repetido con sus consecuencias.

Nuestra evaluación incluye dos técnicas Bayesianas de comparación de modelos recientemente desarrolladas por Steingroever et al. (2014, 2013b), diseñadas para evaluar el desempeño absoluto de un modelo. La primera técnica, conocida como el *método de ajuste post hoc absoluto*, mide el ajuste del modelo respecto de un conjunto de datos, mientras que la segunda, el *método de simulación*, evalúa la capacidad del modelo para generar el conjunto de datos.

Nuestros resultados sugieren que una versión extendida del modelo de *Aprendizaje de Prospecto de Valencia*, que se caracteriza por combinar la función de valor de Teoría de Prospecto (Tversky y Kahneman, 1992) y la Regla de Actualización por Decaimiento (Erev y Roth, 1998), es el más adecuado entre los tres que evaluamos para explicar el patrón de elecciones en el PMH repetido.



# Estudios con el Problema de Monty Hall

by

José Luis Baroja Manzano

## Abstract

This work evaluates three reinforcement learning models' capacity to explain human performance in a repeated version of the Monty Hall Problem.

Recent reports have suggested that the models we evaluate provide adequate explanations on the way human participants learn about and exploit uncertain alternatives as they repeatedly experience the consequences of each option.

Our evaluation comprises two Bayesian model comparison techniques recently developed by Steingroever et al. (2014, 2013b), designed to evaluate a model's absolute performance. The first technique, known as *post hoc absolute fit*, measures the model fit to certain dataset, while the second, the *simulation method*, evaluates the model's capacity to generate the dataset.

Our results suggest that an extended version of the *Prospect Valence Learning* model, which combines Prospect Theory's utility function (Tversky and Kahneman, 1992) and the Decay Learning Rule of expectancy updating (Erev y Roth, 1998), is the best model among the three we evaluate to explain human response patterns in the repeated Monty Hall Problem.

# Capítulo 1

## Introducción y Panorama General

El Problema de Monty Hall (PHM) es una de las tareas más estudiadas y menos comprendidas en la literatura contemporánea sobre decisión. En este problema el participante debe hallar un premio valioso que se encuentra oculto detrás de una de tres puertas. En la primera etapa de elección el participante selecciona una de las puertas pero no tiene acceso a su contenido. Posteriormente se revela una de las puertas que no ha sido elegida por el participante y que no contiene el premio. Finalmente el participante debe decidir entre conservar la puerta que eligió al principio o cambiar a la otra que permanece cerrada.

La mejor decisión consiste en cambiar de puerta porque al hacerlo el participante dobla la probabilidad de encontrar la recompensa. Sin embargo, los participantes humanos generalmente muestran una fuerte tendencia a quedarse con la primera puerta elegida la primera vez que enfrentan el problema, aunque dicha tendencia disminuye a medida que el participante resuelve el PMH repetidamente. En este trabajo proponemos una explicación de cómo y por qué participantes humanos aprenden a cambiar de puerta en el PMH cuando tienen experiencia extensiva en la tarea.

Nuestra explicación toma la forma de una comparación entre tres modelos formales de decisión desarrollados para explicar el comportamiento humano en una tarea de elección que tiene similitudes importantes con el PMH: La Tarea de Selección de Iowa (TSI). Dichos modelos especifican cómo se modifica la probabilidad de elegir diferentes alternativas inciertas a medida que el participante experimenta las consecuencias de elegir cada una.

En las secciones siguientes presentamos el PMH, su solución, y los principales hallazgos em-

píricos reportados a la fecha. Después discutimos en qué sentido el PHM es similar a la TSI. Posteriormente describimos los modelos a evaluar, especificando su forma funcional y la interpretación psicológica de cada uno de sus componentes. Finalmente analizamos cuál de ellos ofrece la mejor explicación de las observaciones recolectadas utilizando dos técnicas de evaluación de modelos recientemente sugeridas por Steingroever et al. (2013b, 2014): 1) El *método de ajuste post hoc absoluto* y 2) el *método de simulación*.

Nuestros resultados sugieren que el modelo  $\mathcal{M}_{\text{PVL} \neq 0}$  ofrece la mejor explicación de los patrones de elección en el PMH observados en una muestra de estudiantes de licenciatura. Dicho modelo se caracteriza por utilizar la función de valor de Teoría de Prospecto (Tversky y Kahneman, 1992) y la Regla de Actualización por Decaimiento (Ever y Roth, 1998). Los valores inferidos sobre los parámetros de ambos componentes sugieren que en la versión repetida del PMH, 1) las pérdidas son ignoradas y todos los montos de ganancia son subjetivamente similares al margen de su valor absoluto, y 2) la experiencia acumulada en cada opción tiene un peso relativo mucho mayor que el de las consecuencias recientes de cada alternativa.

## Capítulo 2

# El Problema de Monty Hall

La versión clásica del PMH consiste en cuatro etapas:

1. El anfitrión, referido de ahora en adelante como *Monty*, oculta un premio valioso detrás de una de tres puertas y en las dos restantes oculta premios indeseables o simplemente las deja vacías. Las puertas son indistinguibles y no es posible conocer el contenido de cada una mientras estén cerradas.
2. Un participante, referido como *Connie*<sup>1</sup>, elige una de las puertas como *primera elección* pero su contenido no es revelado.
3. Monty revela el contenido de una puerta que no ha sido elegida por Connie y que tampoco contiene el premio. Si tiene la posibilidad de abrir más de una puerta (i.e., si Connie ha elegido la puerta premiada en la primera elección) elige entre ellas al azar.
4. Connie decide entre “quedarse” con la primera puerta elegida o “cambiar” a la otra que permanece cerrada. El contenido de esta *segunda elección* es revelado y entregado a Connie.

La mejor estrategia para Connie consiste en cambiar de puerta sin importar qué puerta eligió primero ni qué puerta fue revelada por Monty. Como mostraremos a continuación, la decisión “cambiar” gana con  $Pr = \frac{2}{3}$ , mientras que “no cambiar” (i.e., “quedarse”) gana con  $Pr = \frac{1}{3}$ .

Suponiendo que Connie eligió la puerta 1 y que Monty mostró la puerta 2, podemos calcular

---

<sup>1</sup>Los nombres *Monty* y *Connie* han sido propuestos por A. V. Gnedin (2012) para facilitar la explicación del problema.

la probabilidad condicional de que la puerta 3 esté premiada,  $Pr(p = 3|c = 1, m = 2)$ , que corresponde con la probabilidad de ganar al cambiar de puerta.

Por Regla de Bayes:

$$\begin{aligned}
 Pr(p = 3|c = 1, m = 2) &= \frac{Pr(c = 1, m = 2|p = 3)Pr(p = 3)}{Pr(c = 1, m = 2)} \\
 &= \frac{Pr(c = 1, m = 2|p = 3)Pr(p = 3)}{\sum_{i=1}^3 Pr(c = 1, m = 2|p = i)Pr(p = i)} \\
 &= \frac{\frac{1}{3} \cdot \frac{1}{3}}{\frac{1}{6} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}} = \frac{\frac{1}{9}}{\frac{1}{18} + \frac{1}{9}} \\
 &= \frac{2}{3}
 \end{aligned}$$

en donde  $p$ ,  $c$ , y  $m$  son variables aleatorias que representan la ubicación del premio, la puerta elegida por Connie, y la puerta mostrada por Monty, respectivamente.

De manera similar podemos calcular la probabilidad de que la puerta elegida esté premiada, es decir, la probabilidad de ganar sin cambiar de puerta:

$$\begin{aligned}
 Pr(p = 1|c = 1, m = 2) &= \frac{Pr(c = 1, m = 2|p = 1)Pr(p = 1)}{\sum_{i=1}^3 Pr(c = 1, m = 2|p = i)Pr(p = i)} \\
 &= \frac{\frac{1}{6} \cdot \frac{1}{3}}{\frac{1}{6} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{1}{3} + 0 \cdot \frac{1}{3}} = \frac{\frac{1}{18}}{\frac{1}{18} + \frac{1}{9}} \\
 &= \frac{1}{3}
 \end{aligned}$$

Este resultado se sostiene en cuatro supuestos centrales:

- Es igualmente probable que cualquier puerta esté premiada.
- Es igualmente probable que Connie elija cualquier puerta como primera elección.
- La primera puerta elegida por Connie es independiente de la puerta premiada.
- Cuando Monty puede decidir qué puerta mostrar (i.e., cuando  $p = c$ ) elige entre ambas opciones al azar, con igual probabilidad sobre cada una.

Cuando estos supuestos son el caso el resultado del análisis anterior es idéntico para todas las combinaciones posibles de puertas elegidas y puertas mostradas: “Cambiar” gana con  $Pr = \frac{2}{3}$  y “quedarse” con  $Pr = \frac{1}{3}$ .

Aunque es posible rastrear los orígenes del PMH al menos hasta mediados del Siglo XX (Rosenhouse, 2009; Selvin, 1975a, 1975b), el problema fue popularizado hasta inicios de la década de los 90's en una serie de columnas publicadas por Marilyn vos Savant en la revista *Parade* (vos Savant, 1990a, 1990b, 1991). La primera de dichas publicaciones apareció como respuesta a la siguiente pregunta, propuesta por el lector Craig Whitaker:

Supón que te encuentras en un programa de concursos y puedes elegir entre tres puertas. Detrás de una puerta hay un automóvil, detrás de las otras, cabras. Escojes una puerta, digamos la número 1, y el anfitrión, que sabe qué hay detrás de cada puerta, abre otra puerta, digamos la número 3, que tiene una cabra. Después te pregunta, “¿Quieres elegir la puerta número 2?”. ¿Te conviene cambiar de puerta?

La respuesta de vos Savant (1990):

Sí; deberías cambiar de puerta. La primera puerta tiene  $1/3$  de probabilidad de ganar, pero la segunda puerta tiene  $2/3$  de probabilidad. Esta es una buena manera de visualizar qué pasa: Supón que hay un millón de puertas y tú eliges la puerta número 1. El anfitrión, que sabe qué hay detrás de cada puerta y que siempre evitará la puerta con el premio, abre todas excepto la puerta número 777,777. Cambiarías a esa puerta rápidamente, ¿o no?

Este inocente intercambio inauguró uno de los debates más intensos en teoría de probabilidad de que se tiene registro. En reacción a la solución propuesta, vos Savant recibió literalmente varios miles de cartas en respuesta. La mayoría de ellas no sólo cuestionaba la solución sino que la desafiaba, sugiriendo, erróneamente, que una vez que Monty mostraba el contenido de una puerta sin el premio las dos puertas restantes tenían la misma probabilidad de estar premiadas. Notablemente, gran porcentaje de las “correcciones” provenían de altos grados académicos, muchos de ellos especialistas en estadística, probabilidad o matemáticas (von Savant, 1996; Rosenhouse, 2009).

Desde entonces, el PMH se ha vuelto una de las tareas más estudiadas en diversas disciplinas (Rosenhouse, 2009). En matemáticas y en teoría de probabilidad, por ejemplo, el PMH ha motivado numerosos debates sobre la especificación correcta del problema y/o sobre la solución más adecuada. En economía y en ciencia cognitiva el PMH se ha convertido en una tarea experimental extensivamente utilizada con humanos y recientemente también con animales no humanos. Muchos de los estudios empíricos sobre el PMH pueden clasificarse en dos grandes grupos: 1) Los estudios que buscan responder a la pregunta: *¿Por qué la mayoría de participantes decide no cambiar de puerta la primera vez que enfrentan la tarea?*, y 2) los estudios que buscan averiguar *si los participantes aprenden a cambiar de puerta cuando enfrentan el PMH repetidamente.*

Los primeros reportes sobre el desempeño humano en el PMH fueron publicados por Donald Granberg y colaboradores hace casi 20 años (Granberg y Brown, 1995; Granberg y Dorr, 1998; Granberg, 1999a, 1999b). Esta primera serie de estudios observó experimentalmente dos de los hallazgos más robustos en el PMH: 1) Participantes humanos tienen una fuerte tendencia a “quedarse” con la primera puerta elegida en lugar de “cambiar” a la otra puerta cerrada cuando enfrentan el PMH por primera vez, y 2) la propensión a cambiar de puerta aumenta a medida que los participantes resuelven el problema en repetidas ocasiones. Ambos resultados han sido replicados en numerosas ocasiones desde entonces, en diferentes grupos de edad y en diferentes grupos étnicos (Granberg, 1999a).

El primer estudio que comparó el desempeño de humanos con el de una especie no humana en el PMH fue reportado por Herbranson y Schroeder (2010). Estos investigadores midieron la tasa de “cambios” de palomas en una versión del PMH adaptada a cajas operantes estándar durante varios días. En la primera sesión la tasa de “cambios” media de seis palomas fue cercana

a 33 %, lo cual reproduce el resultado comúnmente hallado en humanos. Sin embargo, después de 30 sesiones la tasa de “cambios” media del pequeño grupo de pichones fue superior a 95 %. En contraste, la tasa media de “cambios” de un grupo de estudiantes de licenciatura después de 200 ensayos en el PMH fue cercana a 60 %, aunque mostró un aumento gradual pequeño a lo largo de la sesión. Este sorprendente resultado parece sugerir que las palomas, a diferencia de los humanos, se desempeñan mejor en el PMH cuando tienen experiencia repetida con el problema. Sin embargo hay varias diferencias entre el tratamiento que recibieron las palomas y el que recibieron los humanos en los experimentos de Herbranson y Schroeder (2010). Entre ellas, quizá la más importante es que las palomas tuvieron entrenamiento extensivo en el problema (cerca de 3000 ensayos), mientras que el de los humanos fue considerablemente menor en términos del número de ensayos.

Para evaluar el tamaño del efecto debido al número de ensayos de entrenamiento, Mazur y Kahlbaugh (2012) replicaron el diseño original de Herbranson y Schroeder (2010), pero analizaron el desempeño de pichones y de humanos en la misma etapa de la tarea. Sus resultados demuestran que la tasa de “cambios” media de un grupo de 36 estudiantes de licenciatura es superior a la de un grupo de 12 palomas cuando ambas especies son comparadas entre los ensayos 175 y 200 ( $\approx 60\%$  y  $\approx 40\%$ , respectivamente). Además, en condiciones adicionales Mazur y Kahlbaugh (2012) manipularon la tasa de recompensas condicional a “cambiar” volviéndola más extrema (hasta  $90\%$ ). En ambas especies la tasa de “cambios” fue mayor en las condiciones en las que “cambiar” pagaba a una tasa más alta, lo cual sugiere que los organismos son sensibles a la probabilidad de reforzamiento por “cambiar”. Uno de los hallazgos recurrentes en esta serie de estudios es que tanto en pichones como en humanos, incluyendo niños de 3 a 5 años, existe una gran variabilidad individual en la tasa de “cambios” en todas las etapas del experimento: Hay individuos que “cambian” desde los primeros bloques de ensayos, individuos que aprenden a “cambiar” gradualmente, individuos cuya tasa de “cambios” se mantiene constante a lo largo de la sesión, e incluso individuos que tienden a “cambiar” menos conforme el experimento avanza.

Tanto la tasa media de “cambios” como la alta variabilidad individual reportada por Mazur y Kahlbaugh han sido observadas en varias réplicas recientes. Klein et al. (2013), por ejemplo, midieron la tasa de “cambios” de 15 estudiantes y de 7 monos Rhesus en 500 ensayos del PMH. Las tasas medias de “cambios” de ambas especies incrementaron a lo largo de los ensayos,



aunque ningún grupo alcanzó niveles óptimos. Nuevamente, estos investigadores reportan una alta variabilidad individual: Es posible identificar individuos en ambas especies con una tasa de “cambios” alta y otros con tasas bajas a lo largo de la tarea.

Un resultado similar fue reportado recientemente por Stagner y Zentall (2015), quienes utilizaron palomas y ratas. En este experimento, aparte de presentar la versión clásica del PMH, Stagner y Zentall incluyeron dos condiciones adicionales: La primera consistió en una versión invertida del PMH en la que “quedarse” pagaba el 67 % de las veces, mientras que “cambiar” sólo el 33 %; la segunda fue una versión control en la que tanto “cambiar” como “quedarse” pagaban el 50 % de los ensayos. Stagner y Zentall reportan que tanto ratas como palomas parecen preferir en promedio la opción con la tasa de reforzamiento mayor en cada condición (indiferencia en la condición control), pero los sujetos de ambas especies muestran grandes diferencias individuales en todas las condiciones del experimento.

En resumen, numerosos reportes previos sugieren 1) que en promedio participantes humanos y animales aprenden a cambiar de puerta en el PMH cuando resuelven el problema repetidamente, y 2) que existe una gran variabilidad individual en la tasa de “cambios” en el problema (Granberg y Brown, 1995; Granberg y Dorr, 1998; Granberg, 1999a, 1999b; Herbranson y Schroeder, 2010; Mazur y Kahlbaugh, 2012; Klein et al., 2013; Stagner y Zentall, 2015; Friedman, 1998; Palacios-Huerta, 2003).

Pese a que estos hallazgos han sido verificados en numerosas ocasiones y a que existen muchos conjuntos de datos que pueden servir como base para evaluar diferentes modelos, a la fecha no hay explicaciones formales sobre *cómo* los organismos aprenden a “cambiar” o a “quedarse” a medida que tienen contacto con las consecuencias de cada opción en el PMH. En este trabajo evaluaremos tres posibles explicaciones que buscan dar cuenta de cómo se modifica el patrón de “cambios” y de “no cambios” a medida que participantes humanos exploran ambas alternativas de respuesta y tienen contacto con las consecuencias de cada una. Las explicaciones que evaluamos tienen la forma de modelos de aprendizaje por reforzamiento. La evaluación cuantitativa de modelos formales de decisión puede ser una herramienta útil para entender el desempeño humano en diferentes situaciones de elección, y en nuestro caso puede ayudar a evaluar y mejorar diferentes explicaciones sobre por qué es tan difícil cambiar de puerta en el Problema de Monty Hall.

## Capítulo 3

# Los Modelos

Los modelos que evaluamos en este trabajo tienen su origen en la literatura de aprendizaje por reforzamiento y fueron desarrollados para explicar el desempeño humano en una tarea similar a la versión repetida del PMH: La Tarea de Selección de Iowa. En las secciones siguientes describimos la TSI y después presentamos los modelos en detalle.

La Tarea de Selección de Iowa (Bechara, Damasio, Damasio y Anderson, 1994) fue diseñada para estudiar la manera en la que participantes humanos aprenden y deciden en ambientes inciertos. En esta tarea cada participante recibe cierto monto de dinero ficticio al inicio del experimento, generalmente \$2000. Posteriormente debe elegir repetidamente una de cuatro posibles barajas a lo largo de la sesión, que usualmente se compone de 100 ensayos. Las cartas de cada baraja tienen impreso un monto de ganancia y uno de pérdida que controlan el monto acumulado del participante. Por ejemplo, si en determinado ensayo un participante elige una carta rotulada  $+100 -150$ , gana 100 puntos pero pierde 150 y el monto acumulado total decrementa 50 unidades al final del ensayo.

Dos de las cuatro barajas, tradicionalmente rotuladas como A y B, son barajas desventajosas porque aunque ofrecen ganancias inmediatas grandes (100 puntos por ensayo) ocasionalmente ofrecen pérdidas mayores, por lo que en caso de elegir las repetidamente las ganancias totales disminuyen. Las dos barajas restantes, C y D, son barajas ventajosas porque, pese a ofrecer ganancias inmediatas pequeñas (50 puntos por ensayo), ocasionalmente ofrecen pérdidas todavía menores y a largo plazo el monto total acumulado aumenta. En este escenario, un participante interesado en obtener la mayor cantidad de puntos debería elegir consistentemente las barajas

C y/o D y evitar las barajas A y/o B.

Adicionalmente, dos barajas de la tarea (A y C) ofrecen pérdidas relativamente frecuentes (en 5 de cada 10 cartas), mientras que las dos restantes (B y D) ofrecen pérdidas infrecuentes (en 1 de cada 10 cartas). El cuadro 3-1 presenta un resumen del arreglo clásico de pagos en la TSI (ver Steingroever et al., 2014; Bechara et al., 1994).

	Baraja A	Baraja B	Baraja C	Baraja D
	<b>Desventajosa</b>	<b>Desventajosa</b>	<b>Ventajosa</b>	<b>Ventajosa</b>
	Pérdidas frecuentes	Pérdidas infrecuentes	Pérdidas frecuentes	Pérdidas infrecuentes
Ganancia en cada ensayo	\$100	\$100	\$50	\$50
Número de pérdidas en 10 ensayos	5	1	5	1
Monto de pérdida	-\$250	-\$1250	-\$50	-\$250
Pérdida esperada en 10 ensayos	-\$1250	-\$1250	-\$250	-\$250
Ganancia total esperada en 10 ensayos	<b>-\$250</b>	<b>-\$250</b>	<b>\$250</b>	<b>\$250</b>

Tabla 3-1: Características de cada baraja en la Tarea de Selección de Iowa.

Al inicio de la sesión los participantes no saben cuál es el arreglo de pagos de las diferentes barajas y tienen que descubrirlo y adaptarse a él conforme avanza la tarea y experimentan las consecuencias de elegir cada opción. La pregunta importante es: ¿Cómo lo hacen?

Una aproximación que permite responder esta pregunta consiste en especificar modelos for-

males que hacen explícito el papel que diferentes variables ambientales y psicológicas juegan en el proceso de aprendizaje. La mayoría de dichos modelos formaliza la idea intuitiva de que a medida que el participante gana experiencia en la tarea y tiene contacto con las consecuencias de cada opción su patrón de elección se modifica hasta mostrar una preferencia estable por una o más barajas.

La propuesta de esta tesis es que el PMH puede ser visto como un problema de aprendizaje similar a la TSI porque al parecer participantes sin experiencia no saben que las opciones “cambiar” o “quedarse” pagan a tasas diferentes, pero pueden aprender sobre ellas a medida que las exploran repetidamente. Bajo este supuesto, y asumiendo que el mismo proceso de aprendizaje opera en ambas tareas, razonamos que un modelo que capture correctamente el desempeño humano en la TSI también debería capturarlo en el PMH.

En los últimos 15 años se han propuesto numerosos modelos que buscan explicar el patrón de elecciones humanos en la TSI. Tres de los más importantes son: 1) El modelo de *Valencia Esperada* ( $\mathcal{M}_{EV}$ , por sus siglas en inglés) (Bussemeyer y Stout, 2002), 2) el modelo de *Aprendizaje de Prospecto de Valencia* ( $\mathcal{M}_{PVL}$ ) (Ahn et al., 2008, 2011), y un híbrido de ambos, 3) el modelo de *Aprendizaje de Prospecto de Valencia-delta* ( $\mathcal{M}_{PVL\delta}$ ) (Ahn et al., 2008; Fridberg et al., 2010). En este trabajo evaluamos la capacidad de estos modelos para dar cuenta del patrón de elecciones humano en el PMH. Elegimos dichos modelos porque de acuerdo con reportes recientes constituyen las mejores explicaciones disponibles de la TSI a la fecha.

Los tres modelos están constituidos por tres componentes: 1) Una *función de utilidad*, que especifica el valor subjetivo de las consecuencias de cada ensayo, 2) una *regla de aprendizaje*, que describe cómo la utilidad de cada ensayo se relaciona con el valor esperado de cada alternativa de respuesta, y 3) una *regla de elección*, que calcula la probabilidad de elegir cada alternativa en función de su valor esperado relativo.

A continuación se presentan los detalles formales de los tres modelos acompañados de la interpretación psicológica de cada uno de sus componentes.

## Modelo $\mathcal{M}_{EV}$

La función de utilidad del modelo  $\mathcal{M}_{EV}$  está dada por<sup>1</sup>:

$$u_k(t) = (1 - \mathbf{w}) \cdot W(t) + \mathbf{w} \cdot L(t) \quad \text{EV-1}$$

$$\mathbf{w} \in [0, 1]$$

En donde  $u_k(t)$  representa la utilidad de la opción<sup>2</sup>  $k$  en el ensayo  $t$ ,  $W(t)$  corresponde al monto ganado al elegir la opción  $k$  en el ensayo  $t$ , y  $L(t)$  corresponde al monto perdido por elegir  $k$  en  $t$ . La ecuación (EV-1) indica que la utilidad en el ensayo  $t$  es una suma ponderada de las ganancias y las pérdidas obtenidas en dicho ensayo que depende del parámetro  $w$ : Cuando  $w$  toma un valor alto (i.e.,  $w > 0.5$ ), las pérdidas del ensayo tienen más peso que las ganancias, y cuando  $w < 0.5$  las ganancias reciben más peso que las pérdidas. En este sentido, un participante con  $w = 0.8$  puede ser clasificado como *averso a las pérdidas*, mientras que un participante con  $w = 0.2$  puede describirse como *buscador de ganancias*. Como  $w$  controla la atención relativa entre ganancias y pérdidas este parámetro se conoce como *parámetro de ponderación de atención*.

De acuerdo con el modelo  $\mathcal{M}_{EV}$ , la utilidad subjetiva de la opción  $k$ , elegida en el ensayo  $t$ , actualiza el valor esperado de la opción  $k$  de acuerdo con:

$$Ev_k(t) = Ev_k(t - 1) + \mathbf{a} \cdot (u_k(t) - Ev_k(t - 1)) \quad \text{EV-2}$$

$$\mathbf{a} \in [0, 1]$$

En donde  $Ev_k(t)$  corresponde con el valor esperado de la opción  $k$  actualizado después de observar las consecuencias de elegir  $k$  en el ensayo  $t$ ,  $Ev_k(t - 1)$  representa en valor esperado de dicha opción hasta el ensayo anterior, y  $u_k(t)$  es la utilidad obtenida al elegir  $k$  en  $t$ . La regla de aprendizaje (EV-2), conocida como *regla delta*, define el valor esperado de cada opción en el ensayo actual como el resultado de agregar la diferencia entre el valor esperado hasta el ensayo

---

<sup>1</sup>Al presentar cada ecuación resaltamos en negritas los parámetros libres y especificamos sus rangos.

<sup>2</sup>En la TSI cada baraja se interpreta como una opción diferente y por lo tanto  $k \in [1, 2, 3, 4]$ . En el presente análisis del PMH nos concentramos en las opciones de respuesta “quedarse” y “cambiar”, y  $k \in [0, 1]$ . No modelamos ni la primera ni la segunda puerta elegida porque nuestro principal interés se limita a entender cómo se aprende a cambiar (al margen de qué puerta se abandona), aunque consideramos que un modelo completo del PMH debería especificar el patrón de elecciones sobre cada puerta en ambas etapas.

anterior  $Ev_k(t-1)$  y la utilidad observada en el ensayo actual  $u_k(t)$  al valor esperado de la opción hasta el ensayo anterior. Cuando la utilidad obtenida en el ensayo ( $t$ ) es mayor que la esperada hasta ( $t-1$ ), el valor esperado de la alternativa aumenta, pero cuando la utilidad obtenida es menor que la esperada, el valor esperado decrementa. El parámetro  $a$ , llamado *parámetro de tasa de actualización*, controla la importancia relativa de dicha diferencia: Un valor alto de  $a$  otorga mucho peso a las consecuencias inmediatamente anteriores e indica que el valor esperado de cada opción cambia drásticamente entre ensayos. En contraste, un valor de  $a$  pequeño sugiere que toda la experiencia previa juega un papel dominante al actualizar el valor esperado de cada alternativa y que las consecuencias recientes tienen poco peso relativo.

El modelo  $\mathcal{M}_{EV}$  supone que después del ensayo ( $t$ ) sólo se actualiza el valor esperado de la alternativa elegida. Las demás opciones conservan su valor esperado acumulado hasta el ensayo ( $t-1$ ).

La probabilidad de elegir cada opción de respuesta de acuerdo con el modelo  $\mathcal{M}_{EV}$  está dada por:

$$P[S_k(t+1)] = \frac{e^{\theta(t)Ev_k}}{\sum e^{\theta(t)Ev_j}} \quad \text{EV-3}$$

En donde  $P[S_k(t+1)]$  es la probabilidad de elegir la opción  $k$  en el ensayo siguiente, y está definida en función de la razón relativa entre el valor esperado de la opción  $k$ ,  $Ev_k$ , y la suma del valor esperado en el ensayo actual de todas las opciones disponibles,  $\sum Ev_j$ . El *parámetro de sensibilidad*,  $\theta(t)$ , controla el impacto de cada valor esperado sobre la probabilidad de elegir cada opción: Cuando  $\theta(t) = 0$  la probabilidad de elegir cada baraja es uniforme, independientemente de los valores esperados actuales. Si  $\theta(t) = 1$ , la probabilidad de elegir cada opción es directamente proporcional a la razón de su valor esperado en relación a la suma de los restantes. Finalmente, si  $\theta(t) \gg 1$  y en la medida que  $\theta(t) \rightarrow \infty$ , el modelo elige la alternativa con mayor valor esperado con certidumbre.

La ecuación (EV-3), también conocida como *regla de elección de maximización suave* o *regla de elección de razón de fuerzas*<sup>3</sup> (Luce, 1959) es común para los tres modelos que evaluamos en este trabajo. Sin embargo, como se mostrará más adelante, cada modelo define  $\theta(t)$  de manera diferente.

---

<sup>3</sup>Los nombres en inglés de la ecuación (EV-3) son *softmax choice rule* y *ratio-of-strength choice rule*.

En el modelo  $\mathcal{M}_{EV}$ ,  $\theta(t)$  está definida como:

$$\theta(t) = (t/10)^c \quad \text{EV-4}$$

$$c \in [-2, 2]$$

En donde  $t$  es el valor del ensayo actual y  $c$  es el *parámetro de consistencia*. Cuando  $c > 0$  la sensibilidad aumenta conforme transcurren los ensayos y la probabilidad de elegir cada baraja depende cada vez más de su valor esperado, mientras que cuando  $c < 0$  la sensibilidad disminuye a medida que se acumulan nuevas observaciones y la elección se vuelve aleatoria, al margen de los valores esperados acumulados.

En resumen, el modelo  $\mathcal{M}_{EV}$  propone que el patrón de elecciones en la TSI depende de tres características psicológicas, representadas por los diferentes parámetros libres del modelo. El parámetro de ponderación de atención ( $w$ ) controla la importancia relativa de las ganancias respecto a las pérdidas, el parámetro de tasa de actualización ( $a$ ) especifica el peso relativo de la experiencia acumulada total respecto a la experiencia anterior inmediata, y el parámetro de consistencia ( $c$ ) pondera la importancia que el conocimiento de cada alternativa tiene sobre la probabilidad de elegirla a medida que la sesión experimental avanza.

### Modelo $\mathcal{M}_{PVL}$

En el modelo  $\mathcal{M}_{PVL}$  la utilidad de la opción  $k$  en el ensayo  $t$  está definida como:

$$u_k(t) = \begin{cases} X(t)^A & \text{si } X(t) \geq 0 \\ -w \cdot |X(t)|^A & \text{si } X(t) < 0 \end{cases} \quad \text{PVL-1}$$

$$A \in [0, 1]$$

$$w \in [0, 5]$$

En donde  $X(t) = W(t) - |L(t)|$  es la ganancia total del ensayo. A diferencia de la función de utilidad del modelo  $\mathcal{M}_{EV}$ , en la ecuación (PVL-1) la utilidad del ensayo no es necesariamente una función lineal de la cantidad de puntos obtenida: La forma específica de la función depende del valor del parámetro  $A$  (*parámetro de forma*): Si  $A \rightarrow 0$  la utilidad toma la forma

de una función escalonada, lo que quiere decir que cualquier resultado positivo, sin importar su magnitud absoluta, tiene el mismo valor de utilidad porque todos tienden a uno; algo similar ocurre cuando el resultado es negativo: Sin importar la cantidad de puntos perdida, la utilidad de estos ensayos tiende a  $-w$ . A medida que  $A \rightarrow 1$  la utilidad del ensayo se vuelve directamente proporcional a la cantidad de puntos obtenida en el ensayo. El otro parámetro libre de la función de utilidad (PVL-1),  $w$ , se conoce como *parámetro de aversión a pérdidas* porque controla el impacto de las pérdidas sobre la utilidad del ensayo. Si  $w > 1$  el peso de las pérdidas es mayor que el de las ganancias. En el caso especial  $w = 1$  las ganancias y las pérdidas tienen la misma importancia relativa, y a medida que  $w \rightarrow 0$  las pérdidas tienden a ser ignoradas.

Una vez que la utilidad del ensayo  $t$  para la opción  $k$  ha sido calculada, el modelo  $\mathcal{M}_{\text{PVL}}$  actualiza el valor esperado de todas las alternativas disponibles en la tarea de acuerdo con:

$$Ev_k(t) = \mathbf{a} \cdot Ev_k(t - 1) + \delta_k(t) \cdot u_k(t) \quad \text{PVL-2}$$

$$\mathbf{a} \in [0, 1]$$

En donde  $Ev_k(t)$  es el valor esperado de la alternativa  $k$  actualizado después del ensayo  $t$ ,  $Ev_k(t - 1)$  es el valor esperado de la misma alternativa hasta el ensayo anterior, y  $u_k(t)$  representa la utilidad obtenida por elegir  $k$  en  $t$ . A diferencia del modelo  $\mathcal{M}_{\text{EV}}$ , el modelo  $\mathcal{M}_{\text{PVL}}$  asume que el valor esperado de todas las alternativas (y no sólo el de la alternativa elegida) se modifica en cada ensayo. De acuerdo con (PVL-2), también conocida como *regla de decaimiento* (Erev y Roth, 1998), el valor esperado de cada alternativa en el ensayo  $t$  es una fracción de su valor esperado hasta el ensayo anterior. El tamaño de dicha fracción está determinado por el parámetro  $a$ , conocido como *parámetro de recencia*: En la medida que  $a \rightarrow 0$  las consecuencias del ensayo actual tienen más peso que la experiencia previa acumulada; cuando  $a \rightarrow 1$  todas las consecuencias pasadas tienen un impacto más fuerte que la consecuencia inmediata anterior. La variable  $\delta_k(t)$  únicamente indica si la alternativa  $k$  fue elegida en el ensayo  $t$ ; si este es el caso, el valor esperado de  $k$  después de  $t$  es una fracción de su valor esperado hasta el ensayo anterior más la utilidad de  $k$  en  $t$ .

Habiendo actualizado el valor esperado de todas las opciones de respuesta después del ensayo  $t$ , el modelo  $\mathcal{M}_{\text{PVL}}$  calcula la probabilidad de elegir cada alternativa en el ensayo siguiente



utilizando la regla de elección de maximización suave:

$$P[S_k(t+1)] = \frac{e^{\theta(t)Ev_k}}{\sum e^{\theta(t)Ev_j}} \quad \text{PVL-3}$$

La ecuación (PVL-3) es idéntica a la ecuación (EV-3). Sin embargo, el modelo  $\mathcal{M}_{\text{PVL}}$  define  $\theta(t)$  como:

$$\theta(t) = 3^c - 1 \quad \text{PVL-4}$$

$$c \in [0, 5]$$

En donde el *parámetro de consistencia*,  $c$ , controla el impacto del valor esperado de cada opción sobre la probabilidad de elegirla: Cuando  $c$  toma valores pequeños la elección se vuelve aleatoria con distribución uniforme sobre todas las alternativas, sin importar sus valores esperados. En el otro extremo, cuando  $c$  toma valores grandes la elección tiende a ser determinista con una preferencia por la alternativa con valor esperado mayor. Es importante resaltar dos diferencias importantes entre las definiciones de  $\theta(t)$  del modelo  $\mathcal{M}_{\text{EV}}$  y del modelo  $\mathcal{M}_{\text{PVL}}$ : 1) Mientras que el modelo  $\mathcal{M}_{\text{EV}}$  asume que  $\theta(t)$  se modifica ensayo a ensayo, el modelo  $\mathcal{M}_{\text{PVL}}$  asume que  $\theta(t)$  permanece constante a lo largo de la sesión, y 2) aunque ambas definiciones de  $\theta(t)$  están en función de un parámetro  $c$ , con interpretación similar en ambos modelos, el rango de valores válidos de  $c$  es diferente en cada modelo.

Recapitulando, el modelo  $\mathcal{M}_{\text{PVL}}$  sugiere que las elecciones de cada participante en la TSI dependen de cuatro características psicológicas que corresponden con los parámetros libres del modelo. El parámetro de forma ( $A$ ) especifica la relación entre pérdidas o ganancias y la utilidad subjetiva de cada ensayo, el parámetro de aversión a pérdidas ( $w$ ) pondera el impacto de las pérdidas experimentadas en cada opción sobre sus valores esperados, el parámetro de recencia ( $a$ ) controla el peso relativo de las experiencias recientes respecto a las pasadas, y el parámetro de consistencia ( $c$ ) representa el impacto que el conocimiento acumulado de cada alternativa tiene sobre la probabilidad de elegirla a lo largo de la tarea.

## Modelo $\mathcal{M}_{\text{PVL}\delta}$

El tercer modelo que evaluamos es un híbrido de los modelos  $\mathcal{M}_{\text{EV}}$  y  $\mathcal{M}_{\text{PVL}}$ . En concreto, el modelo  $\mathcal{M}_{\text{PVL}\delta}$  utiliza la función de utilidad, la regla de elección y la definición de sensibilidad del modelo  $\mathcal{M}_{\text{PVL}}$ , pero importa la regla delta de aprendizaje del modelo  $\mathcal{M}_{\text{EV}}$ . Todos los componentes y parámetros del modelo  $\mathcal{M}_{\text{PVL}\delta}$  tienen la misma función e interpretación que en los modelos originales. Los enlistamos nuevamente buscando mayor claridad.

La función de utilidad del modelo  $\mathcal{M}_{\text{PVL}\delta}$  corresponde con la función de utilidad del modelo  $\mathcal{M}_{\text{PVL}}$ :

$$u_k(t) = \begin{cases} X(t)^{\mathbf{A}} & \text{if } X(t) \geq 0 \\ -\mathbf{w} \cdot |X(t)|^{\mathbf{A}} & \text{if } X(t) < 0 \end{cases} \quad \text{PVLd-1}$$

$$\mathbf{A} \in [0, 1]$$

$$\mathbf{w} \in [0, 5]$$

De acuerdo con el modelo  $\mathcal{M}_{\text{PVL}\delta}$ , los valores esperados de cada alternativa de respuesta se actualizan de acuerdo con la regla delta del modelo  $\mathcal{M}_{\text{EV}}$ <sup>4</sup>:

$$Ev_k(t) = Ev_k(t-1) + \mathbf{a} \cdot (u_k(t) - Ev_k(t-1)) \quad \text{PVLd-2}$$

$$\mathbf{a} \in [0, 1]$$

Al igual que los modelos  $\mathcal{M}_{\text{EV}}$  y  $\mathcal{M}_{\text{PVL}}$ , el modelo  $\mathcal{M}_{\text{PVL}\delta}$  utiliza la regla de elección de maximización suave para calcular la probabilidad de elegir cada opción en función de su valor esperado:

$$P[S_k(t+1)] = \frac{e^{\theta(t)Ev_k}}{\sum e^{\theta(t)Ev_j}} \quad \text{PVLd-3}$$

Finalmente, el componente de sensibilidad  $\theta(t)$  en el modelo  $\mathcal{M}_{\text{PVL}\delta}$  se define como:

$$\theta(t) = 3^c - 1 \quad \text{PVLd-4}$$

$$\mathbf{c} \in [0, 5]$$

---

<sup>4</sup>El modelo  $\mathcal{M}_{\text{PVL}\delta}$  también supone que después de cada ensayo sólo se actualiza el valor esperado de la alternativa elegida, mientras que los valores esperados de las alternativas restantes conservan el valor esperado acumulado hasta el ensayo anterior.

Que conserva el supuesto de sensibilidad invariante entre ensayos del modelo  $\mathcal{M}_{PVL}$ .

Los tres modelos que hemos descrito representan tres explicaciones diferentes sobre por qué los participantes muestran cierto patrón de elecciones en la TSI. Las tres explicaciones suponen que al resolver la tarea tres procesos psicológicos entran en juego: 1) Un proceso encargado de computar un valor de utilidad a partir del resultado de cada ensayo, 2) un proceso que integra la utilidad obtenida en cada ensayo con la experiencia previa de cada alternativa, y 3) un proceso que decide qué alternativa elegir con base al conocimiento acumulado sobre cada alternativa hasta el ensayo actual. Pese a que los tres modelos comparten estos supuestos cada uno propone una explicación distinta sobre cómo funciona cada proceso. El funcionamiento de dichos procesos depende de una o más características psicológicas que corresponden con los parámetros libres de cada modelo y de la forma funcional particular del modelo especificada en las ecuaciones de cada componente.

## Capítulo 4

# Recolección de Datos

### Participantes

Recolectamos datos de 22 estudiantes (13 mujeres) inscritos en una materia sobre aprendizaje y conducta adaptable en la Facultad de Psicología, UNAM. Los participantes tienen una edad media de 19.5 ( $DE=0.79$ ), distribuida en un rango de 18.6 a 20.25 años.

### Aparatos y Materiales

Construimos una versión del problema de Monty Hall con papel ilustración y otros materiales. El aparato está compuesto por tres compartimientos separados por barreras opacas con espacio suficiente para albergar un dulce o algunas monedas en cada uno. El contenido de cada compartimiento puede ser explorado por el participante deslizando la puerta correspondiente hacia arriba. Aparte, el diseño de la caja permite al investigador colocar recompensas en cada compartimiento desde la parte posterior del instrumento sin necesidad de abrir las puertas.

La figura 4-1 presenta una fotografía del aparato experimental.

### Procedimiento

Todos los participantes fueron entrevistados de manera individual en el Laboratorio 25 de la Facultad de Psicología, UNAM. Al inicio de la sesión explicamos a cada uno que el ejercicio formaba parte de un estudio sobre la manera en la que la gente toma decisiones y que en caso de participar podría ganar dinero en efectivo. Después de obtener consentimiento oral describimos

la tarea de la siguiente forma:

*En este juego tu objetivo consiste en obtener la mayor cantidad de puntos posible. En cada ensayo voy a esconder uno de estos tres papeles (etiquetados con 1, 3, o 5) en alguna de estas tres puertas. Después podrás elegir la puerta que quieras. Cuando hayas elegido una puerta yo descubriré el contenido de otra. La puerta que voy a descubrir siempre estará vacía y nunca descubriré la puerta que elegiste al principio. Finalmente deberás decidir entre quedarte con la puerta que habías escogido o cambiar a la otra que permanece cerrada. Si el papel se encuentra en la puerta que elijas al final ganarás la cantidad de puntos que indica el papel; si el papel no se encuentra en la segunda puerta que elijas, los perderás.*

*Comenzarás a jugar con 100 puntos. Dependiendo de tus decisiones y resultados en el juego esa cantidad podrá aumentar o disminuir. Al final del experimento contaremos los puntos que acumules. Cuando todos los participantes hayan completado el ejercicio premiaremos a los mejores con dinero en efectivo. Si quieres cobrar, trata de acumular tantos puntos como puedas.*

Habiendo escuchado la explicación de la tarea cada participante jugó un ensayo de prueba. En este ensayo nos aseguramos de que el participante conocía sus opciones de respuesta en cada etapa pidiéndole que nos explicara qué cosas podía hacer y las consecuencias potenciales de cada acción. Todos los participantes demostraron un entendimiento claro del problema en el ensayo de prueba.

Después de la explicación de la tarea y antes de comenzar con los ensayos definitivos preguntamos a cada participante si alguna vez había escuchado sobre o participado en algún juego similar. Ningún participante reportó conocer el problema.

Cada sesión individual consistió en 100 ensayos del PMH. Cada ensayo incluyó las 4 etapas del problema definidas en la página 2. Antes de comenzar cada participante recibió 100 puntos en forma de cuadros de papel de 1cm x 1cm. A lo largo de los ensayos el experimentador retiraba o aumentaba el número de cuadros de papel correspondiente con el resultado del ensayo.

La ubicación del premio y la cantidad de puntos en juego en cada ensayo fueron elegidas al azar de acuerdo con distribuciones uniformes independientes sobre las tres puertas. En los ensayos en los que el investigador podía decidir entre más de una puerta para mostrar decidimos al

azar de acuerdo con una distribución uniforme sobre ambas alternativas. Todos los participantes jugaron contra la misma secuencia de puertas premiadas y montos en juego.

En cada ensayo registramos la cantidad de puntos en juego, la puerta premiada, la primera puerta elegida, la puerta mostrada y la segunda puerta elegida. Estos datos permiten saber si el participante se quedó o cambió de puerta, si encontró o no la recompensa, y la cantidad de puntos que ganó o perdió.

De los 22 participantes originales 5 no completaron la tarea por fatiga o aburrimiento. No incluimos sus datos en los análisis que reportamos en este trabajo.

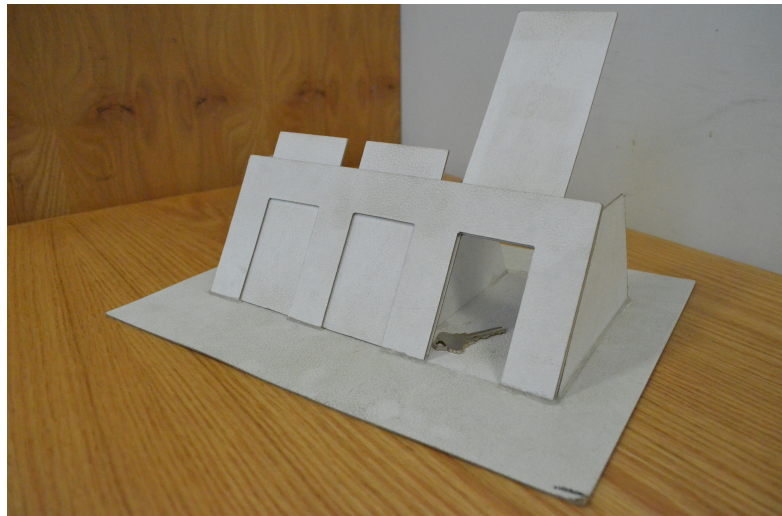


Figura 4-1: Aparato experimental. Cada compartimiento puede ocultar un premio pequeño. Las puertas son opacas y no es posible ver su contenido sin abrirlas.



## Capítulo 5

# Resultados

La variable dependiente de principal interés en este trabajo es la secuencia de “cambios” y de “no cambios” de cada participante. Dado que dichas opciones pagan a tasas distintas, la primera pregunta que nos interesa responder es si los participantes son sensibles a la diferencia en tasas de pago y en consecuencia tienden a elegir la alternativa que paga mejor a medida que la sesión avanza y tienen contacto con ambas opciones.

La figura 5-1 presenta un resumen de las secuencias de “cambios” observadas en nuestro experimento. Cada línea gris une la proporción de “cambios” de cada participante en cada bloque de diez ensayos. La línea negra une los promedios de proporciones de “cambios” entre participantes en cada bloque. Como puede apreciarse en la gráfica, observamos que en promedio la proporción de cambios aumenta aproximadamente de 0.25 al inicio de la sesión hasta 0.5 en los últimos ensayos. Es importante resaltar que esta tendencia poblacional parece el resultado de un proceso de aprendizaje con alta variabilidad individual: Es posible identificar participantes cuya tasa de “cambios” aumenta drásticamente (hasta 1.0 en los últimos dos bloques, por ejemplo) y a participantes que incluso muestran un decremento en la tasa de cambios al final de la sesión.

Tanto la tendencia poblacional como la variabilidad individual son semejantes a las reportadas en varios trabajos previos (Mazur y Kahlbaugh, 2012; Klein et al., 2013; Stagner y Zentall, 2015; Palacios-Huerta, 2003; Friedman, 1998).



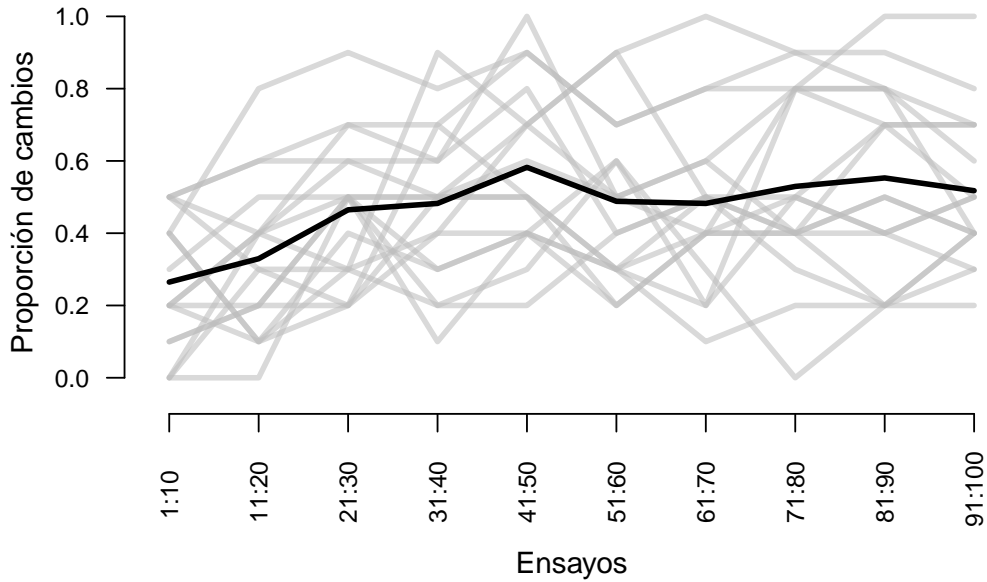


Figura 5-1: Proporción media de cambios en bloques de 10 ensayos. Las líneas grises representan a cada participante. La línea negra une las proporciones medias de cambios en cada bloque de la sesión.

## Modelamiento

En la implementación de los tres modelos que evaluamos en este trabajo asumimos que cada participante tiene valores particulares de los parámetros libres de cada modelo pero que todos los participantes provienen de una distribución poblacional común respecto de cada parámetro. Esta aproximación, conocida como *análisis jerárquico*, permite descomponer la varianza observada en diferentes niveles de análisis y resulta muy útil para modelar diferencias individuales y al mismo tiempo aprender sobre las características generales de la población (Shiffrin et al., 2008; Rouder et al., en prensa). Una forma útil de presentar este tipo de modelos consiste en representar las diferentes variables y los diferentes parámetros del modelo en forma de nodos y en señalar las dependencias condicionales entre parámetros y variables utilizando flechas. La figura 5-2 presenta un ejemplo de este tipo de representación, correspondiente con la versión jerárquica del modelo  $\mathcal{M}_{EV}$ .

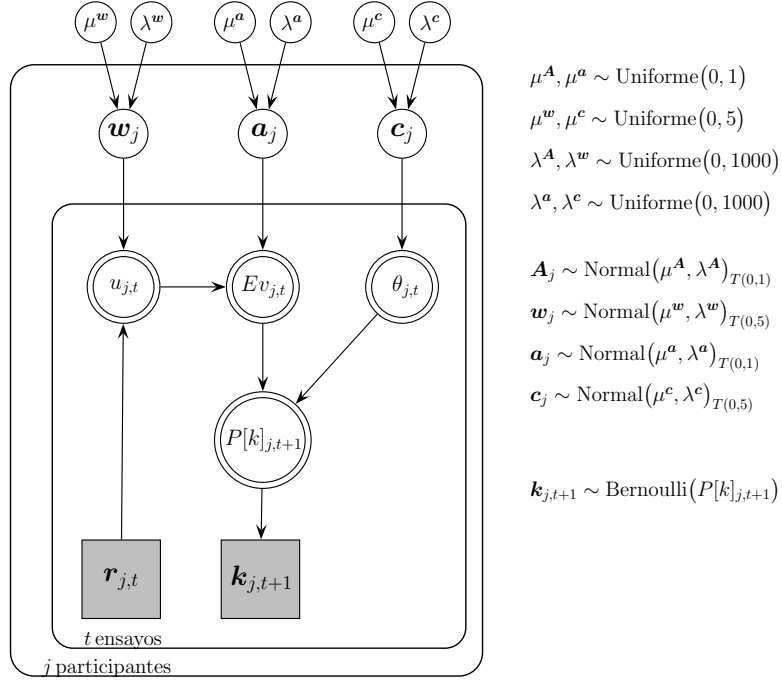


Figura 5-2: Representación gráfica del modelo  $\mathcal{M}_{EV}$  jerárquico.

En la notación gráfica que utilizamos, cada nodo blanco representa una variable no observada y cada nodo gris una variable observada. Los círculos denotan variables continuas y los cuadrados variables discretas. Las flechas señalan dependencias condicionales entre diferentes variables y parámetros, y los platos representan replicaciones independientes del proceso generativo (Lee, 2008).

El proceso generativo del modelo  $\mathcal{M}_{EV}$  comienza en el nodo  $r_{j,t}$ , ubicado en la esquina inferior izquierda del modelo gráfico. Este nodo corresponde con el resultado del ensayo  $t$  del participante  $j$  y por lo tanto se encuentra dentro de los platos correspondientes a ensayos ( $t$ ) y a participantes ( $j$ ). Dado que en nuestra tarea sólo hay seis resultados posibles, todos ellos enteros ( $+5, +3, +1, -1, -3, -5$ ), representamos a  $r_{j,t}$  como una variable discreta. El paso siguiente genera un parámetro  $w_j$  individual de acuerdo con una distribución normal con parámetros  $\mu^w$  y  $\lambda^w$  poblacionales. Los nodos  $\mu^w$  y  $\lambda^w$  se encuentran fuera de ambos platos porque asumimos que son constantes para todos los participantes y para todos los ensayos. En cambio, el nodo  $w_j$  se ubica dentro del plato de participantes pero fuera del plato de ensayos porque asumimos que

cada participante tiene un valor  $\mathbf{w}_j$  particular y que ese valor se mantiene constante a lo largo de todos los ensayos. Tanto el nodo  $\mathbf{w}_j$  individual como los nodos  $\mu^w$  y  $\lambda^w$  poblacionales son blancos porque son nodos desconocidos, en el sentido de que no observamos ni medimos su valor directamente. Sin embargo, como se mostrará más adelante, podemos inferir su valor con base en los valores de los nodos observados, es decir, podemos aprender sobre cuál es el peso relativo de pérdidas contra ganancias de cada individuo y cuál es la tendencia poblacional después de observar las decisiones y resultados de cada participante en cada ensayo de la tarea.

Conociendo el resultado del ensayo  $t$  del participante  $j$  y el valor  $\mathbf{w}_j$  del participante  $j$  el modelo  $\mathcal{M}_{EV}$  calcula la utilidad  $u_{j,t}$  del ensayo  $t$  del participante  $j$ . La relación entre  $u_{j,t}$ ,  $\mathbf{w}_j$  y  $\mathbf{r}_{j,t}$  es determinista porque dado un valor  $\mathbf{w}_j$  y un valor  $\mathbf{r}_{j,t}$  sólo existe un valor  $u_{j,t}$  posible, especificado por la ecuación (EV-1). Este tipo de nodos se representan gráficamente utilizando límites dobles.

En las etapas siguientes, el modelo  $\mathcal{M}_{EV}$  genera un valor  $\mathbf{a}_j$  por cada individuo y actualiza el valor esperado de cada opción en cada ensayo  $Ev_{j,t}$  en función de  $\mathbf{a}_j$  y de  $u_{j,t}$ , de acuerdo con la ecuación (EV-2). Posteriormente, el modelo genera un valor  $\mathbf{c}_j$  individual y un valor  $\theta_{j,t}$  por cada individuo en cada ensayo de acuerdo con (EV-4). Finalmente, la utilidad esperada de cada opción  $Ev_{j,t}$  y el parámetro de sensibilidad  $\theta_{j,t}$  determinan la probabilidad de elegir cada opción en el ensayo siguiente  $P[k]_{j,t+1}$  de acuerdo con la ecuación (EV-3).

La estructura gráfica de los tres modelos es similar, salvo porque los modelos  $\mathcal{M}_{PVL}$  y  $\mathcal{M}_{PVL\delta}$  cuentan con un parámetro adicional  $\mathbf{A}_j$  individual y porque el parámetro de sensibilidad  $\theta$  es constante a lo largo de todos los ensayos aunque variable entre participantes, como se muestra en la figura 5-3.

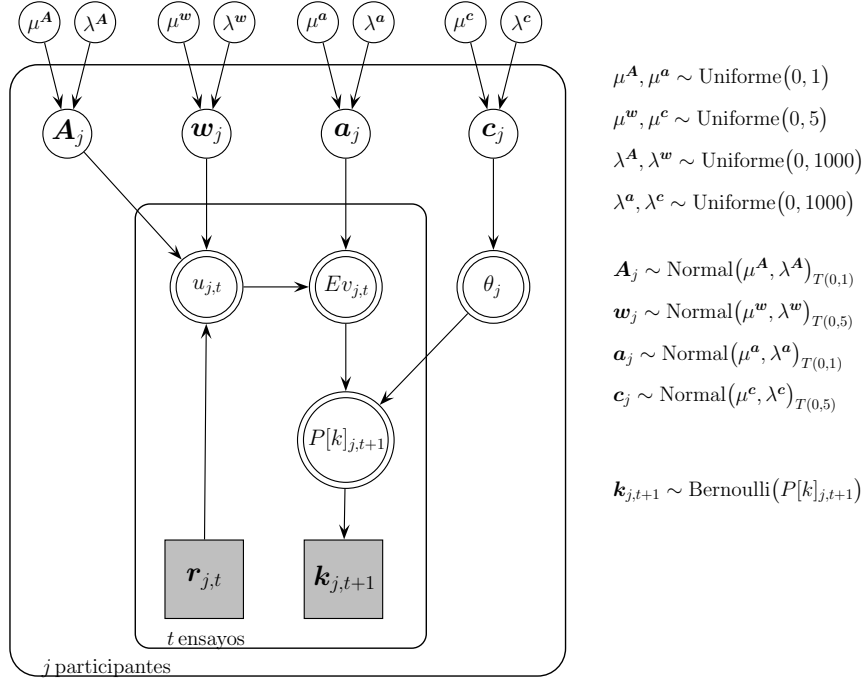


Figura 5-3: Representación gráfica de los modelos  $\mathcal{M}_{\text{PVL}}$  y  $\mathcal{M}_{\text{PVL}\delta}$ .

En los tres modelos el nodo final  $\mathbf{k}_{j,t+1}$  corresponde a las elecciones observadas de cada individuo en cada ensayo. Codificamos los *cambios* como 1 y los *no-cambios* como 0. Por lo tanto, la variable  $\mathbf{k}_{j,t+1}$  se distribuye Bernoulli con parámetro  $P[k]_{j,t+1}$ . Dado que conocemos  $\mathbf{k}_{j,t+1}$ , podemos inferir cuáles son los valores más probables de  $P[k]_{j,t+1}$  utilizando la regla de Bayes:

$$Pr(P[k]_{j,t+1} | \mathbf{k}_{j,t+1}) \propto Pr(\mathbf{k}_{j,t+1} | P[k]_{j,t+1})Pr(P[k]_{j,t+1})$$

En donde  $Pr(P[k]_{j,t+1} | \mathbf{k}_{j,t+1})$  se conoce como *distribución posterior* y responde la pregunta de interés: *¿Qué deberíamos de creer sobre los posibles valores paramétricos dado un conjunto de observaciones?* La distribución posterior depende de la *función de verosimilitud*,  $Pr(\mathbf{k}_{j,t+1} | P[k]_{j,t+1})$ , que especifica qué tan probable es observar el conjunto de datos bajo cada posible valor paramétrico, y de la *distribución prior*,  $Pr(P[k]_{j,t+1})$ , que especifica nuestras creencias sobre los posibles valores paramétricos antes de tomar en cuenta los datos recolectados.

La idea del análisis Bayesiano es que habiendo observado la variable  $\mathbf{k}_{j,t+1}$  podemos aprender sobre la variable  $P[k]_{j,t+1}$ . A su vez, podemos utilizar el conocimiento ganado sobre  $P[k]_{j,t+1}$

para inferir los valores probables de  $Ev_{j,t}$  y de  $\theta_{j,t}$ , y éstos pueden ser utilizados para inferir los valores probables de la utilidad de cada ensayo y de todos los valores paramétricos individuales y poblacionales. En otras palabras, el análisis Bayesiano especifica una distribución posterior sobre cada nodo desconocido. En el caso de los parámetros poblacionales sólo existe una distribución posterior, pero en el caso de los parámetros individuales hay tantas distribuciones posteriores como participantes.

Implementamos el análisis Bayesiano de los tres modelos gráficos en JAGS (Plummer, 2003). JAGS es un software diseñado para aproximar distribuciones posteriores por métodos numéricos utilizando una técnica de muestreo conocida como *cadena Monte Carlo de Markov*. Utilizamos dos cadenas independientes para cada parámetro de cada modelo. El número de iteraciones, de muestreos deshechados, y de muestreos retenidos fue diferente para cada modelo, pero en los tres conservamos el mismo número total de muestreos posteriores (ver cuadro 5-1).

<b>Modelo</b>	<b>Cadenas</b>	<b>Iteraciones</b>	<b>Muestreos deshechados</b>	<b>Muestreos retenidos</b>	<b>Muestreos posteriores totales</b>
$\mathcal{M}_{EV}$	2	70 000	50 000	1 de cada 20	2 000
$\mathcal{M}_{PVL}$	2	50 000	10 000	1 de cada 40	2 000
$\mathcal{M}_{PVL\delta}$	2	40 000	20 000	1 de cada 20	2 000

Tabla 5-1: Detalles de los muestreos posteriores de cada modelo.

Verificamos que ambas cadenas de cada parámetro convergieran en la distribución posterior evaluando el estadístico  $\hat{R}$  (Gelman y Rubin, 1992) y analizándolas visualmente. El estadístico  $\hat{R}$  compara la variabilidad dentro de cada cadena contra la variabilidad entre cadenas. En general, valores de  $\hat{R}$  mayores a 1.05 indican que es dudoso que las cadenas hayan alcanzado convergencia confiable. Ninguno de los parámetros en ninguno de los tres modelos mostró un valor  $\hat{R} > 1.038$ . Visualmente, dos cadenas que convergen en la misma distribución posterior se ven como “orugas gordas y peludas” cuando los valores sampleados son graficados como función de su índice (Lee y Wagenmakers, 2013, p. 80). La figura 5-4 presenta un ejemplo de este tipo de análisis: Ambas cadenas son indistinguibles (salvo por la clave de color) porque los muestreos de ambas se encuentran en el mismo rango de valores, lo que indica que las cadenas convergen y que sus muestreos provienen de la misma distribución posterior.

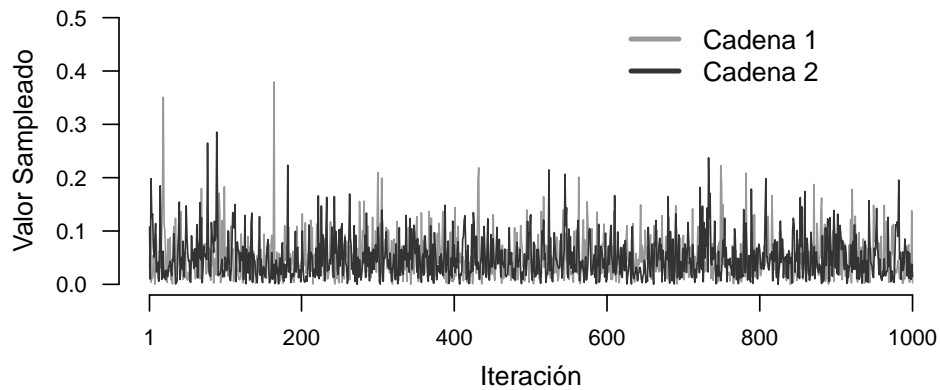


Figura 5-4: Muestréos posteriores separados por cadenas. En la figura aparecen los muestréos correspondientes al parámetro  $\boldsymbol{w}$  del Participante 1 en el modelo  $\mathcal{M}_{EV}$ . Las cadenas son indistinguibles y ambas se mueven en el mismo rango de valores, lo que indica que han convergido en la misma distribución posterior.

El resultado final de JAGS consiste en varios conjuntos de muestréos, cada uno correspondiente con una distribución posterior sobre cada parámetro o nodo desconocido. Podemos analizar directamente dichas distribuciones posteriores utilizando medidas descriptivas estándar. La figura 5-5 es un ejemplo de este análisis. En ella aparece el histograma de todos los muestréos posteriores del parámetro  $\boldsymbol{a}_j$  (*tasa de actualización*) del Participante 2<sup>1</sup>, de acuerdo con el modelo  $\mathcal{M}_{PVL\delta}$ . La línea delgada y punteada señala la media posterior, que computamos como el promedio de todos los muestréos posteriores. La línea sólida y gruesa demarca el intervalo de mayor credibilidad, que comienza en el percentil 2.5 y termina en el percentil 97.5 de los muestréos posteriores. Este intervalo especifica qué tan probable es que el verdadero valor paramétrico se encuentre entre dichos límites. En este ejemplo, deberíamos creer que el verdadero valor de  $\boldsymbol{a}_2$ , de acuerdo con el modelo  $\mathcal{M}_{PVL\delta}$ , se encuentra entre 0.03 y 0.13 con 95% de confianza después de observar la secuencia de decisiones y resultados del Participante 2.

---

<sup>1</sup>Al exponer nuestros resultados en las páginas siguientes tomamos como ejemplo a los participantes 2, 6 y 7. Elegimos a estos participantes porque permiten ilustrar claramente las características de nuestros análisis, pero los resultados y tendencias que reportamos con ellos se sostienen en la mayor parte de nuestra muestra.

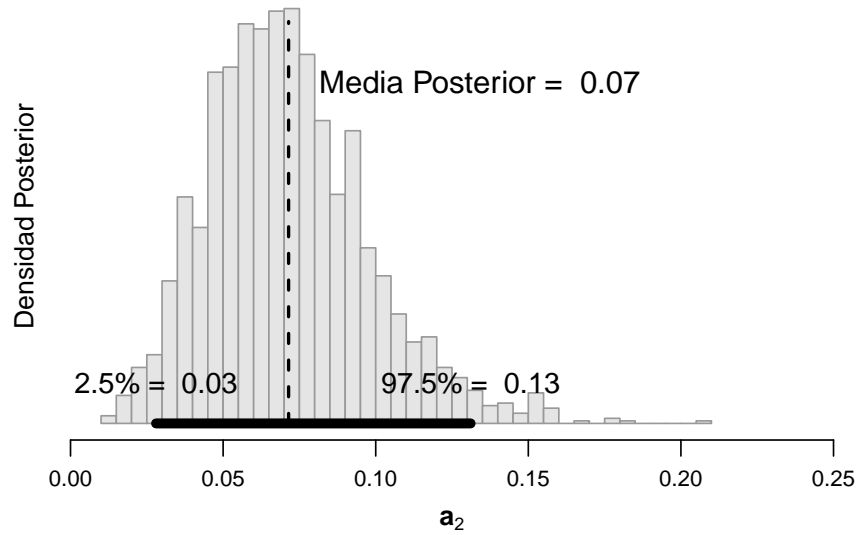


Figura 5-5: Histograma de la distribución posterior del parámetro  $\mathbf{a}_j$  del Participante 2, de acuerdo con el modelo  $\mathcal{M}_{\text{PVL}\delta}$ . La línea punteada y delgada indica la media posterior. La línea sólida y gruesa señala el intervalo de mayor densidad posterior (95 %).

De acuerdo con el modelo  $\mathcal{M}_{\text{PVL}\delta}$ , y en particular con la regla delta de aprendizaje (Ecuación (PVLd-2)), el parámetro  $\mathbf{a}_j$  controla el peso relativo de la experiencia inmediatamente anterior con respecto al de la experiencia pasada en la actualización del valor esperado de cada opción de respuesta. En el caso del Participante 2, el modelo  $\mathcal{M}_{\text{PVL}\delta}$  sugiere que este participante otorga poco peso a las experiencias recientes y que la experiencia pasada tiene mayor peso relativo porque es altamente probable que el valor  $\mathbf{a}_2$  esté por debajo de 0.5.

Esta conclusión es válida sólo si el modelo  $\mathcal{M}_{\text{PVL}\delta}$  representa el proceso de aprendizaje del Participante 2 correctamente, lo cual no necesariamente es el caso. Otros modelos candidatos pueden ofrecer una explicación más adecuada del mismo conjunto de observaciones y en tal caso las conclusiones sobre las características psicológicas del participante deberían basarse en los valores paramétricos inferidos desde dichos modelos.

La pregunta que buscaremos responder en las secciones siguientes es, precisamente, *¿Qué modelo ofrece la mejor explicación de nuestros datos?* Para responderla utilizaremos dos técnicas de evaluación de modelos recientemente desarrolladas por Steingroever et al. (2013b, 2014). La

primera técnica, conocida como *ajuste post hoc absoluto*, mide el ajuste de un modelo al conjunto de datos que pretende explicar. La segunda, el *método de simulación*, evalúa la capacidad del modelo para generar el conjunto de datos. Ambas técnicas utilizan los valores paramétricos inferidos en el análisis Bayesiano previamente detallado.

A continuación mostramos cómo implementamos dichas técnicas. Para anticipar los resultados de ambos análisis, el modelo  $\mathcal{M}_{\text{PVL}}$  muestra el mejor ajuste para la mayoría de los participantes, pero el modelo  $\mathcal{M}_{\text{PVL}\delta}$  ofrece la simulación más adecuada.

### Ajuste Post Hoc Absoluto

Para implementar el método de *ajuste post hoc absoluto* (APHA) seguimos los siguientes pasos (ver Steingroever et al. (2014), Apéndice A).

1. Tomamos un punto al azar de la distribución posterior conjunta de los parámetros individuales de cada modelo. Para el participante  $j$ , este punto representa una combinación de valores paramétricos  $\{w_j, a_j, c_j\}$  en el modelo  $\mathcal{M}_{\text{EV}}$ , y  $\{A_j, w_j, a_j, c_j\}$  en los modelos  $\mathcal{M}_{\text{PVL}}$  y  $\mathcal{M}_{\text{PVL}\delta}$ . Nos aseguramos de que el punto muestreado se encontrara dentro del intervalo de mayor credibilidad de cada distribución marginal.
2. Fijamos el valor esperado inicial de ambas opciones (*cambiar o quedarse*) en cero,  $Ev_{k,0} = 0$ . Por lo tanto, la probabilidad de cambiar en el primer ensayo,  $P[\text{cambiar}]_{j,1} = \frac{1}{2}$  para los tres modelos.
3. Informamos a cada modelo la elección y el resultado del participante  $j$  en el primer ensayo,  $t = 1$ .
4. Utilizando el resultado observado en  $t = 1$  calculamos la utilidad de la opción elegida en dicho ensayo bajo cada modelo.
5. Actualizamos el valor esperado de ambas opciones (en el caso del modelo  $\mathcal{M}_{\text{PVL}}$ ) o sólo de la opción elegida (en el caso de los modelos  $\mathcal{M}_{\text{EV}}$  y  $\mathcal{M}_{\text{PVL}\delta}$ ) en el ensayo  $t = 1$ .
6. Calculamos la probabilidad de elegir cada opción en el ensayo siguiente con cada modelo utilizando los valores esperados actualizados correspondientes y guardamos los valores de



probabilidad calculados.

7. Repetimos los pasos 3-6 para los ensayos  $t = 2$  y hasta  $t = 99$ .
8. Repetimos los pasos 1-7 para cada sujeto 200 veces para conservar la incertidumbre de la distribución posterior conjunta.

En el APHA calculamos la probabilidad de elegir cada opción pero **no generamos** decisiones con ningún modelo. En otras palabras, el APHA computa los valores de probabilidad calculados por cada modelo tomando como base las elecciones y los resultados observados por cada participante y los compara contra la proporción de cambios observada de cada participante. La figura 5-6 presenta un ejemplo gráfico de este análisis. En cada panel de la figura 5-6 la línea negra y gruesa corresponde con las proporciones de cambios observadas del Participante 2 en cada bin de 10 ensayos (ver figura 5-1). Cada una de las 200 líneas grises une las probabilidades calculadas por el método de APHA de un muestreo particular de la distribución posterior conjunta de los parámetros individuales de cada modelo. Finalmente, la línea negra y delgada conecta las probabilidades promedio calculadas por APHA en cada bin de 10 ensayos. El grado de similitud entre la línea negra y gruesa y la negra y delgada es una primera aproximación al nivel de ajuste post hoc absoluto de cada modelo.

Una inspección visual rápida de la figura 5-6 permite apreciar que, en general, los tres modelos ajustan razonablemente bien en la última parte de la sesión, en la que el Participante 2 mostró una proporción de cambios elevada, pero muestran un ajuste deficiente en la primera parte, caracterizada por una proporción baja de cambios observados. Aparte, parece que los modelos  $\mathcal{M}_{EV}$  y  $\mathcal{M}_{PVL\delta}$  ofrecen un ajuste cercano al desempeño observado en los bins 11-20 y 21-30, mientras que el modelo  $\mathcal{M}_{PVL}$  muestra una desviación considerable en dichos ensayos. También resulta evidente que los tres modelos parecen tener la misma incertidumbre posterior en el sentido de que la dispersión de las líneas grises es similar en los tres paneles de la figura 5-6.

Un patrón similar aparece en la figura 5-7, que presenta el APHA para el Participante 6. Nuevamente, parece que los tres modelos ofrecen un ajuste similar al desempeño observado a partir del tercer bin, pero sistemáticamente calculan probabilidades de cambio mayores a las observadas en los primeros 20 ensayos de la tarea.

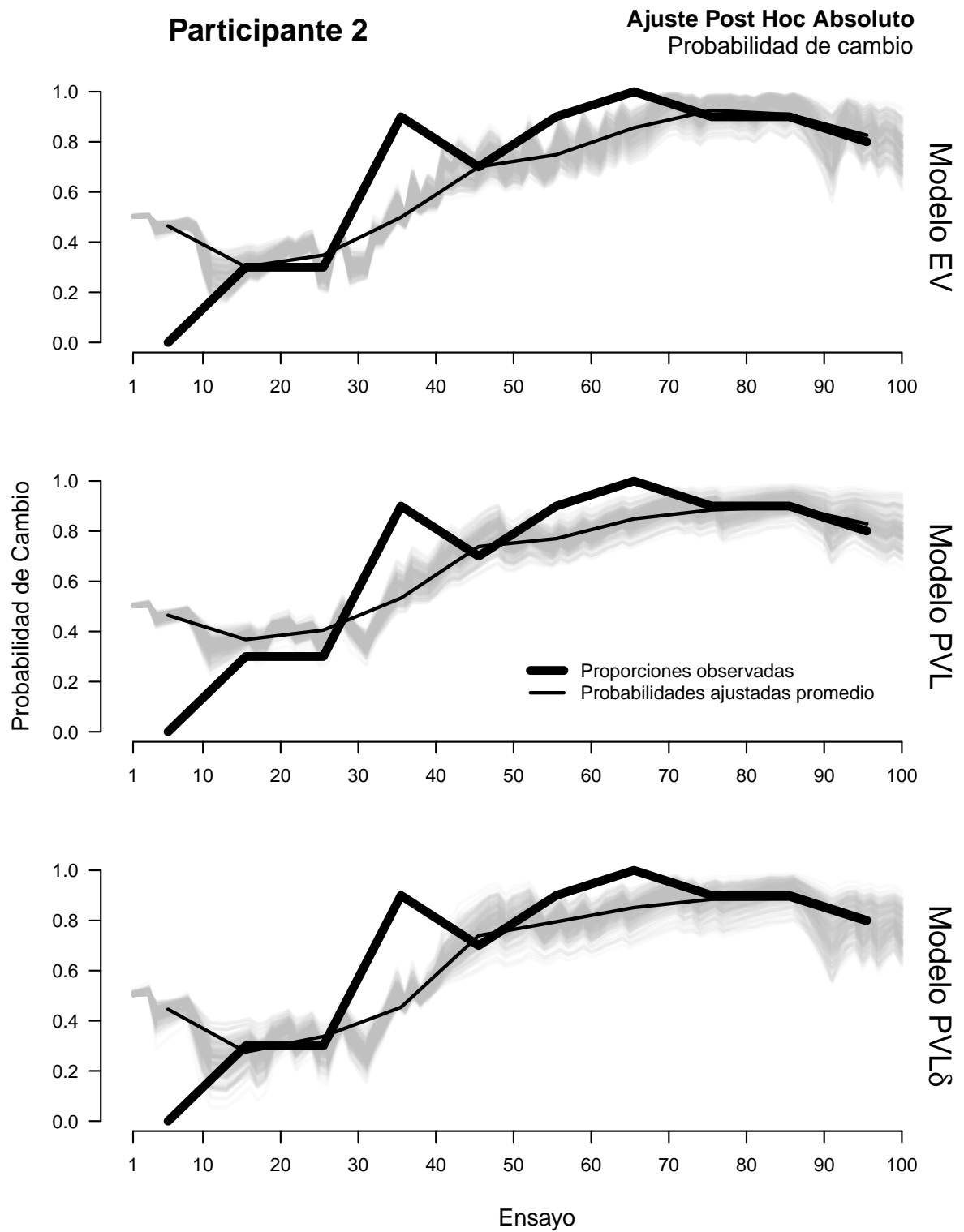


Figura 5-6: Ajuste Post Hoc Absoluto del Participante 2.

### Participante 6

### Ajuste Post Hoc Absoluto Probabilidad de cambio

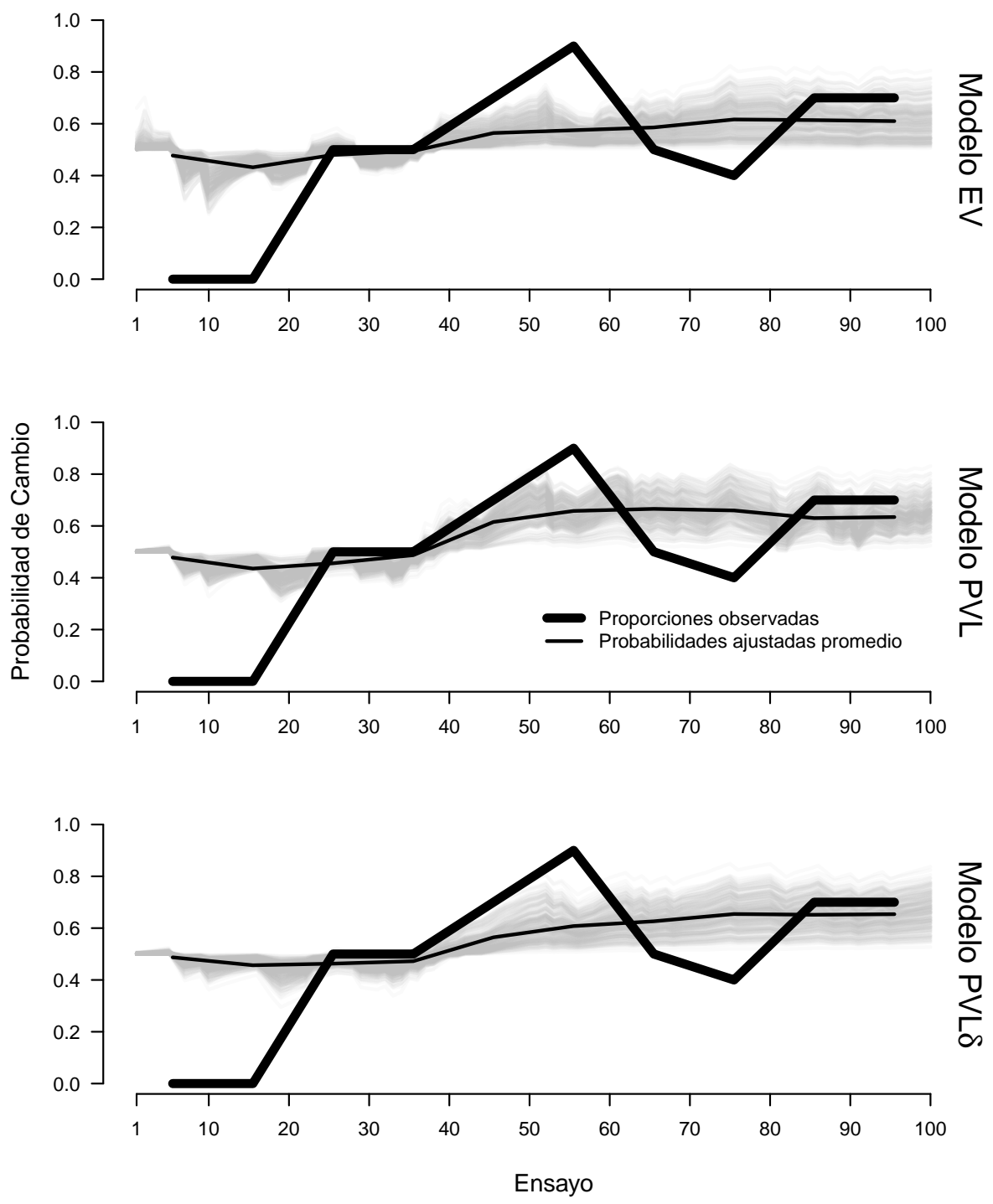


Figura 5-7: Ajuste Post Hoc Absoluto del Participante 6.

Sin embargo, a diferencia del APHA del Participante 2, el modelo  $\mathcal{M}_{EV}$  parece mostrar una incertidumbre posterior menor respecto del Participante 6 comparada con la de los modelos  $\mathcal{M}_{PVL}$  y  $\mathcal{M}_{PVL\delta}$ , particularmente en los bins intermedios.

Con el fin de cuantificar al ajuste post hoc absoluto de cada modelo proponemos una medida que evalúa la distancia entre las probabilidades de cambio ajustadas y las elecciones observadas penalizando la incertidumbre posterior del modelo en cuestión.

Para el participante  $j$  y el modelo  $\mathcal{M}$ , nuestra medida de ajuste está definida como:

$$APHA_{j,\mathcal{M}} = \frac{\sum_t^{100} |E_t^{obs} - P_t^{ajst}| + \sigma_t^{ajst}}{100}$$

Donde  $E_t^{obs}$  corresponde a la elección del participante  $j$  en el ensayo  $t$ ,  $P_t^{ajst}$  es el promedio de las probabilidades de cambio ajustadas por el modelo  $\mathcal{M}$  en el ensayo  $t$ , y  $\sigma_b^{ajst}$  es la desviación estándar de la distribución de las probabilidades ajustadas por el modelo  $\mathcal{M}$  en el ensayo  $t$ . El resultado  $APHA_{j,\mathcal{M}}$  proporciona una medida sobre la distancia entre los datos observados y el ajuste del modelo ensayo a ensayo, penalizando la incertidumbre del modelo a lo largo de la sesión. Si  $APHA_{j,\mathcal{M}} \rightarrow 0$  el promedio de las probabilidades ajustadas es cercano a los cambios observados y la desviación de las probabilidades ajustadas es pequeña, ambas características de un ajuste aceptable.

El cuadro 5-2 presenta las medidas de ajuste de los tres modelos en relación a los datos de cada participante. Para cada participante seleccionamos como mejor modelo al modelo con la medida  $APHA_{j,\mathcal{M}}$  menor. El modelo  $\mathcal{M}_{PVL}$  ajusta mejor que los modelos  $\mathcal{M}_{EV}$  y  $\mathcal{M}_{PVL\delta}$  en 8 de los 17 participantes. Por su parte, el modelo  $\mathcal{M}_{EV}$  ofrece el mejor ajuste en 3 participantes, mientras que el modelo  $\mathcal{M}_{PVL\delta}$  ajusta mejor en los 6 participantes restantes.

Con base en el análisis APHA podríamos concluir que el modelo  $\mathcal{M}_{PVL}$  es el mejor modelo para dar cuenta del desempeño humano en el PMH. Sin embargo, como mostraremos a continuación, el nivel de ajuste entre un modelo y cierto conjunto de datos no necesariamente refleja la capacidad del modelo para *generar* el conjunto de datos. Con el fin de evaluar la capacidad de los modelos para generar el patrón de elecciones observado en el PMH utilizamos una técnica diferente, conocida como el *método de simulación*, que presentamos a continuación.

Participante	$\mathcal{M}_{EV}$	$\mathcal{M}_{PVL}$	$\mathcal{M}_{PVL\delta}$	Mejor modelo
1	0.524	<b>0.507</b>	0.517	$\mathcal{M}_{PVL}$
2	<b>0.371</b>	0.376	0.381	$\mathcal{M}_{EV}$
3	0.522	<b>0.515</b>	0.521	$\mathcal{M}_{PVL}$
4	0.519	0.518	<b>0.516</b>	$\mathcal{M}_{PVL\delta}$
5	0.530	<b>0.524</b>	0.529	$\mathcal{M}_{PVL}$
6	0.516	<b>0.509</b>	0.515	$\mathcal{M}_{PVL}$
7	0.363	<b>0.336</b>	0.348	$\mathcal{M}_{PVL}$
8	0.447	0.448	<b>0.441</b>	$\mathcal{M}_{PVL\delta}$
9	0.514	0.516	<b>0.511</b>	$\mathcal{M}_{PVL\delta}$
10	<b>0.520</b>	0.522	0.528	$\mathcal{M}_{EV}$
11	0.504	<b>0.501</b>	0.502	$\mathcal{M}_{PVL}$
12	0.535	<b>0.494</b>	0.537	$\mathcal{M}_{PVL}$
13	0.514	0.515	<b>0.510</b>	$\mathcal{M}_{PVL\delta}$
14	0.516	0.519	<b>0.514</b>	$\mathcal{M}_{PVL\delta}$
15	<b>0.517</b>	0.525	0.521	$\mathcal{M}_{EV}$
16	0.519	<b>0.516</b>	0.519	$\mathcal{M}_{PVL}$
17	0.519	0.523	<b>0.512</b>	$\mathcal{M}_{PVL\delta}$

Tabla 5-2: Medidas de ajuste  $APHA_{j,\mathcal{M}}$  de cada modelo para cada participante.

## Simulación

El *método de simulación* (SIM) tiene varias similitudes con el APHA, pero también difiere en aspectos importantes. Para implementarlo seguimos los siguientes pasos (ver Steingroever et al. (2014), Apéndice B).

1. Tomamos un punto al azar de la distribución posterior conjunta de los parámetros individuales de cada modelo. Para el participante  $j$ , este punto representa una combinación de valores paramétricos  $\{w_j, a_j, c_j\}$  en el modelo  $\mathcal{M}_{EV}$ , y  $\{A_j, w_j, a_j, c_j\}$  en los modelos  $\mathcal{M}_{PVL}$  y  $\mathcal{M}_{PVL\delta}$ . Nos aseguramos de que el punto muestreado se encontrara dentro del intervalo de mayor credibilidad de cada distribución marginal.

2. Fijamos el valor esperado inicial de ambas opciones (*cambiar* o *quedarse*) en cero,  $Ev_{k,0} = 0$ . Por lo tanto, la probabilidad de cambiar en el primer ensayo,  $P[\text{cambiar}]_{j,1} = \frac{1}{2}$  para los tres modelos.
3. Generamos una elección en el ensayo  $t = 1$  de acuerdo con  $P[\text{cambiar}]_{j,1} = \frac{1}{2}$  con cada modelo y guardamos las elecciones generadas.
4. Generamos el resultado correspondiente a la elección generada en el ensayo  $t = 1$  para cada modelo de acuerdo con los pagos esperados para cada opción de respuesta en el PMH, es decir,  $\Pr(\text{ganar}|\text{cambiar}) = \frac{2}{3}$  y  $\Pr(\text{ganar}|\text{no cambiar}) = \frac{1}{3}$ . La cantidad de puntos en juego en el ensayo  $t = 1$  fue la misma que enfrentaron todos los participantes en dicho ensayo.
5. Utilizando el resultado observado en  $t = 1$  calculamos la utilidad de la opción elegida en dicho ensayo con cada modelo.
6. Actualizamos el valor esperado de ambas opciones (en el caso del modelo  $\mathcal{M}_{\text{PVL}}$ ) o sólo de la opción elegida (en el caso de los modelos  $\mathcal{M}_{\text{EV}}$  y  $\mathcal{M}_{\text{PVL}\delta}$ ) en el ensayo  $t = 1$ .
7. Calculamos la probabilidad de elegir cada opción en el ensayo siguiente con cada modelo utilizando los valores esperados actualizados de cada opción de respuesta.
8. Repetimos los pasos 3-7 para los ensayos  $t = 2$  y hasta  $t = 100$ .
9. Repetimos los pasos 1-8 para cada sujeto 200 veces para conservar la incertidumbre de la distribución posterior conjunta.

A diferencia del APHA, el método SIM **genera** elecciones utilizando los valores paramétricos muestreados de la distribución posterior conjunta y permite evaluar si cada modelo puede comportarse como cada participante en el mismo ambiente.

Examinamos el nivel de acuerdo entre las elecciones generadas por cada modelo y las elecciones observadas de cada participante utilizando la representación gráfica que aparece en la figura 5-8, que consiste en un registro acumulativo de ambos tipos de elecciones.

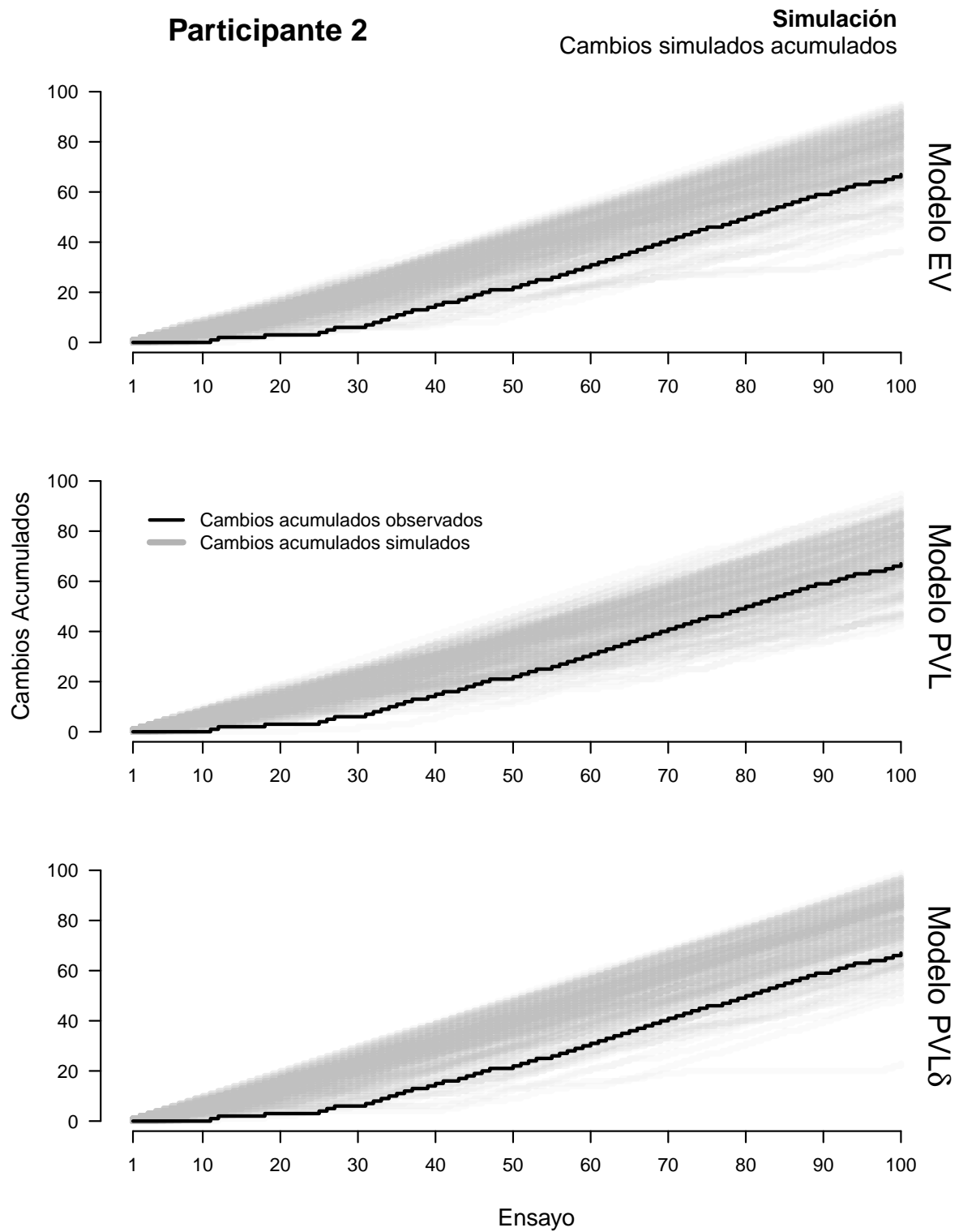


Figura 5-8: Cambios simulados acumulados del Participante 2.

# Participante 6

Simulación  
Cambios simulados acumulados

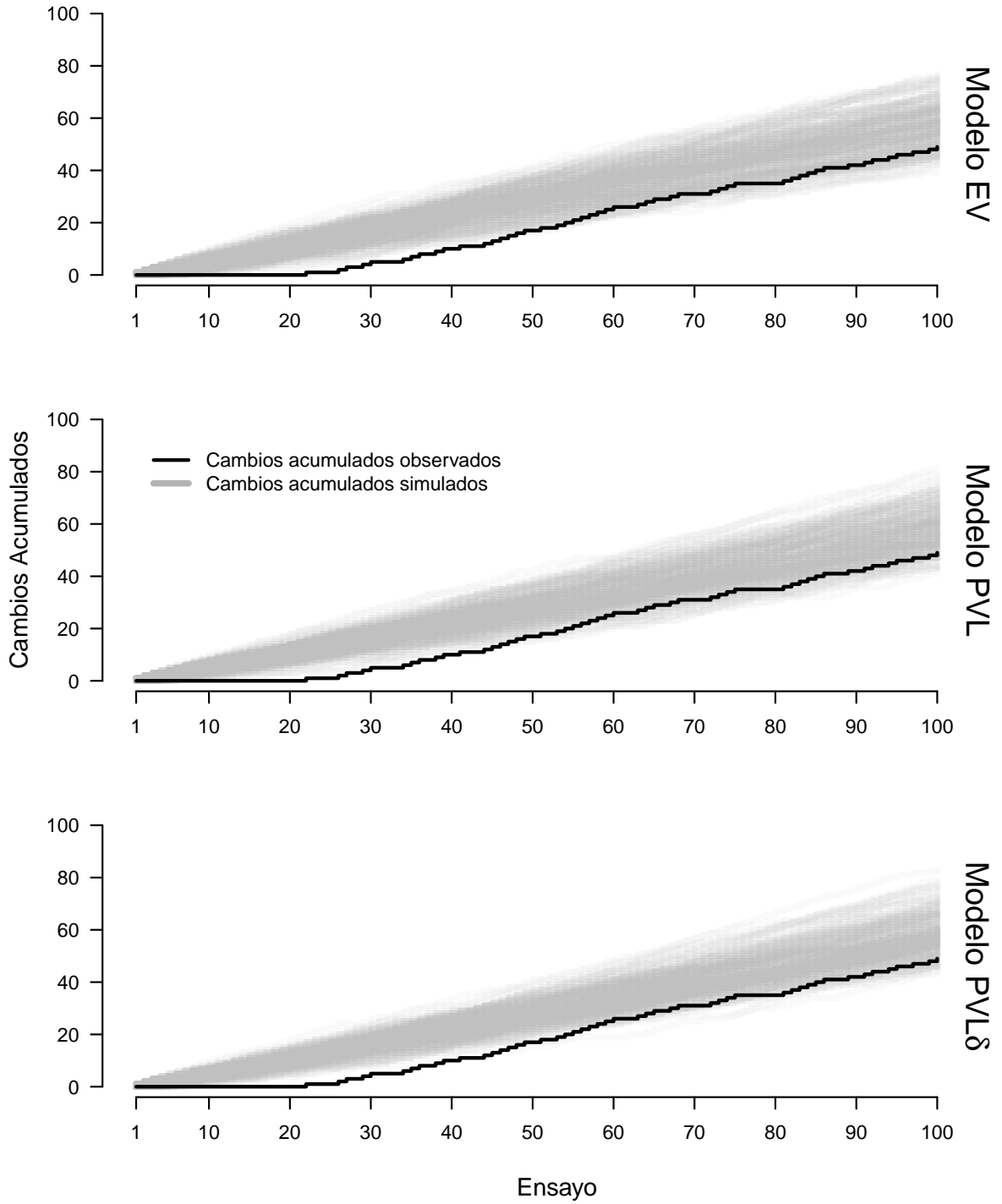


Figura 5-9: Cambios simulados acumulados del Participante 6.



Las líneas grises en cada panel de la figura 5-8 corresponden con los cambios acumulados generados con cada muestreo de la distribución posterior conjunta individual de cada modelo. La línea negra en los tres p aneles presenta los cambios acumulados observados del Participante 2. La regularidad que salta a la vista es que la l inea negra se ubica por debajo de la mayor a de l ineas grises en los tres p aneles, lo cual indica que los tres modelos terminan con un n umero de cambios acumulados mayor que los observados en el Participante 2. Es importante resaltar que la pendiente de las l ineas grises es similar a la de la l inea negra a partir del ensayo 30, pero diferente en los primeros ensayos de la tarea. En otras palabras, parece que los tres modelos deciden cambiar a la misma tasa que el Participante 2 en los  ultimos 70 ensayos, pero que cambian mucho m as seguido que el participante en los primeros 30.

Un patr on similar aparece en el registro acumulativo observado y simulado del Participante 6 (figura 5-9). Nuevamente, el registro acumulativo observado consistentemente termina por debajo de la mayor a de cambios acumulados simulados por los tres modelos. En el caso del Participante 6 la diferencia entre los cambios observados y los simulados en los primeros ensayos es todav a m as evidente: Mientras que el participante no muestra ning un cambio en los primeros 20 ensayos los tres modelos muestran entre 5 y 15 cambios en el mismo per odo de la tarea.

Para cuantificar el nivel de acuerdo entre los cambios simulados por el modelo  $\mathcal{M}$  y los observados en el participante  $j$  proponemos una medida similar a la de APHA. Nuestra medida est a definida como:

$$SIM_{j,\mathcal{M}} = \frac{\sum_t^{100} |C_t^{obs} - C_t^{sim}| + \sigma_t^{sim}}{100}$$

Donde  $C_t^{obs}$  es la cantidad de cambios acumulados del participante  $j$  hasta el ensayo  $t$ ,  $C_t^{sim}$  es el promedio de los cambios simulados acumulados hasta el ensayo  $t$  del modelo  $\mathcal{M}$ , y  $\sigma_t^{sim}$  es la desviaci on est andar de la distribuci on de cambios simulados acumulados del modelo  $\mathcal{M}$  hasta el ensayo  $t$ . La medida  $SIM_{j,\mathcal{M}}$  cuantifica la distancia entre los cambios acumulados del participante y los cambios simulados acumulados del modelo ensayo a ensayo, penalizando a cada modelo por la incertidumbre promedio en la sesi on. Valores de  $SIM_{j,\mathcal{M}}$  cercanos a 0 indican que hay poca distancia entre los cambios observados y los simulados y que hay poca incertidumbre en la distribuci on de cambios simulados acumulados. El cuadro 5-3 presenta las medidas  $SIM_{j,\mathcal{M}}$  de los tres modelos para cada participante. Nuevamente, seleccionamos como

mejor modelo al modelo con el valor  $\mathbf{SIM}_{j, \mathcal{M}}$  menor.

El método SIM sugiere que el modelo  $\mathcal{M}_{\text{PVL}\delta}$  ofrece la mejor simulación en 9 de los 17 participantes, mientras que los modelos  $\mathcal{M}_{\text{EV}}$  y  $\mathcal{M}_{\text{PVL}}$  simulan adecuadamente a 6 y 2 participantes, respectivamente.

Participante	$\mathcal{M}_{\text{EV}}$	$\mathcal{M}_{\text{PVL}}$	$\mathcal{M}_{\text{PVL}\delta}$	Mejor modelo
1	<b>17.77</b>	25.00	22.52	$\mathcal{M}_{\text{EV}}$
2	16.51	<b>14.14</b>	19.18	$\mathcal{M}_{\text{PVL}}$
3	<b>15.07</b>	18.76	17.29	$\mathcal{M}_{\text{EV}}$
4	9.50	<b>8.50</b>	8.85	$\mathcal{M}_{\text{PVL}}$
5	5.56	5.46	<b>5.06</b>	$\mathcal{M}_{\text{PVL}\delta}$
6	15.07	14.41	<b>14.02</b>	$\mathcal{M}_{\text{PVL}\delta}$
7	<b>11.67</b>	12.14	12.05	$\mathcal{M}_{\text{EV}}$
8	8.42	9.35	<b>7.92</b>	$\mathcal{M}_{\text{PVL}\delta}$
9	11.82	13.46	<b>11.39</b>	$\mathcal{M}_{\text{PVL}\delta}$
10	<b>14.37</b>	14.97	14.66	$\mathcal{M}_{\text{EV}}$
11	7.33	7.98	<b>7.14</b>	$\mathcal{M}_{\text{PVL}\delta}$
12	<b>21.44</b>	28.10	24.84	$\mathcal{M}_{\text{EV}}$
13	10.72	12.24	<b>10.63</b>	$\mathcal{M}_{\text{PVL}\delta}$
14	11.33	11.98	<b>10.32</b>	$\mathcal{M}_{\text{PVL}\delta}$
15	16.31	17.56	<b>15.83</b>	$\mathcal{M}_{\text{PVL}\delta}$
16	<b>15.29</b>	15.63	15.61	$\mathcal{M}_{\text{EV}}$
17	16.21	17.87	<b>15.85</b>	$\mathcal{M}_{\text{PVL}\delta}$

Tabla 5-3: Medidas de simulación  $\mathbf{SIM}_{j, \mathcal{M}}$  de cada modelo para cada participante.

Al comparar los resultados del método APHA con los del método SIM es evidente que, en general, un modelo que ajusta adecuadamente el patrón de elecciones de cierto participante no necesariamente es el modelo que mejor simula las decisiones del participante. El cuadro 5-4 resume los resultados de ambos métodos y resalta esta diferencia: El mejor modelo en APHA coincide con el mejor modelo en SIM sólo en 6 de los 17 participantes.

Participante	Mejor modelo	Mejor modelo
	APHA	SIM
1	$\mathcal{M}_{\text{PVL}}$	$\mathcal{M}_{\text{EV}}$
2	$\mathcal{M}_{\text{EV}}$	$\mathcal{M}_{\text{PVL}}$
3	$\mathcal{M}_{\text{PVL}}$	$\mathcal{M}_{\text{EV}}$
4	$\mathcal{M}_{\text{PVL}\delta}$	$\mathcal{M}_{\text{PVL}}$
5	$\mathcal{M}_{\text{PVL}}$	$\mathcal{M}_{\text{PVL}\delta}$
6	$\mathcal{M}_{\text{PVL}}$	$\mathcal{M}_{\text{PVL}\delta}$
7	$\mathcal{M}_{\text{PVL}}$	$\mathcal{M}_{\text{EV}}$
<b>8</b>	$\mathcal{M}_{\text{PVL}\delta}$	$\mathcal{M}_{\text{PVL}\delta}$
<b>9</b>	$\mathcal{M}_{\text{PVL}\delta}$	$\mathcal{M}_{\text{PVL}\delta}$
<b>10</b>	$\mathcal{M}_{\text{EV}}$	$\mathcal{M}_{\text{EV}}$
11	$\mathcal{M}_{\text{PVL}}$	$\mathcal{M}_{\text{PVL}\delta}$
12	$\mathcal{M}_{\text{PVL}}$	$\mathcal{M}_{\text{EV}}$
<b>13</b>	$\mathcal{M}_{\text{PVL}\delta}$	$\mathcal{M}_{\text{PVL}\delta}$
<b>14</b>	$\mathcal{M}_{\text{PVL}\delta}$	$\mathcal{M}_{\text{PVL}\delta}$
15	$\mathcal{M}_{\text{EV}}$	$\mathcal{M}_{\text{PVL}\delta}$
16	$\mathcal{M}_{\text{PVL}}$	$\mathcal{M}_{\text{EV}}$
<b>17</b>	$\mathcal{M}_{\text{PVL}\delta}$	$\mathcal{M}_{\text{PVL}\delta}$

Tabla 5-4: Resumen de los métodos APHA y SIM. Cada renglón presenta el modelo que mejor ajusta y el modelo que mejor simula a cada participante.

Una deficiencia importante, compartida por los tres modelos, es la diferencia entre las elecciones de los participantes y las probabilidades ajustadas y elecciones simuladas en los primeros ensayos de la tarea. De la misma manera que ocurre para los Participantes 2 y 6, los tres modelos consistentemente ajustan probabilidades mayores y simulan más cambios acumulados que los observados en la mayoría de participantes durante los primeros ensayos de la sesión. Este patrón refleja un supuesto común de los tres modelos: En la versión que hemos implementado, los modelos  $\mathcal{M}_{\text{EV}}$ ,  $\mathcal{M}_{\text{PVL}}$  y  $\mathcal{M}_{\text{PVL}\delta}$  asumen que el valor esperado inicial de ambas alternativas es cero,  $Ev_{k,0} = 0$ . Como consecuencia, la probabilidad de *cambiar* (y de *quedarse*) en el primer

ensayo es  $P[k]_{j,1} = \frac{1}{2}$  sin importar los valores paramétricos inferidos del participante  $j$ . En los reportes que preceden a este trabajo dicho supuesto generalmente se utiliza para reflejar la falta de conocimiento sobre la tasa a la que cada baraja paga en la TSI antes de tener contacto con las consecuencias de cada alternativa.

En los análisis que hemos presentado mantuvimos el supuesto  $Ev_{k,0} = 0$  para ambas opciones en los tres modelos buscando conservar la mayor similitud posible en nuestra implementación de los métodos APHA y SIM y las previamente reportadas. Es importante resaltar que al mantener dicho supuesto nuestros resultados sugieren que el modelo  $\mathcal{M}_{PVL\delta}$  ofrece la mejor explicación del patrón de decisiones en el PMH por dos razones: 1) Simula adecuadamente a la mitad y ajusta adecuadamente en un tercio de la muestra de participantes, y 2) es el modelo que muestra mejor consistencia, en el sentido de que es el modelo que mejor ajusta *y* que mejor simula a más participantes (ver cuadro 5-4). Este hallazgo coincide con reportes previos al señalar al modelo  $\mathcal{M}_{PVL\delta}$  como una explicación adecuada sobre el proceso de adaptación humano en ambientes inciertos (Steingroever et al., 2013, 2014).

En las secciones siguientes reproducimos los análisis por APHA y por SIM sobre versiones modificadas de los tres modelos que relajan el supuesto  $Ev_{k,0} = 0$ . En concreto, la modificación consiste en especificar los valores esperados iniciales de ambas opciones,  $Ev_{k,0}$ , como parámetros libres en los tres modelos.

Utilizamos la notación “ $\neq 0$ ” para identificar a los análisis y a los modelos que no asumen  $Ev_{k,0} = 0$ , y que en cambio tratan a los valores esperados iniciales como cantidades desconocidas, susceptibles de ser inferidas.

## **Modelamiento $\neq 0$**

En esta sección repetimos los análisis APHA y SIM sobre versiones extendidas de los modelos  $\mathcal{M}_{EV}$ ,  $\mathcal{M}_{PVL}$  y  $\mathcal{M}_{PVL\delta}$ . Las versiones extendidas de los tres modelos añaden dos parámetros individuales desconocidos que corresponden con los valores esperados iniciales de las opciones *cambiar* y *quedarse*. Esta modificación permite relajar el supuesto de que al principio de la tarea es igualmente probable que el participante cambie o no cambie y en consecuencia tiene el potencial de mejorar el ajuste y la capacidad de simulación de los tres modelos. Para distinguir las versiones extendidas de las versiones originales referiremos a las versiones que no asumen

$Ev_{k,0} = 0$  como  $\mathcal{M}_{EV \neq 0}$ ,  $\mathcal{M}_{PVL \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$ .

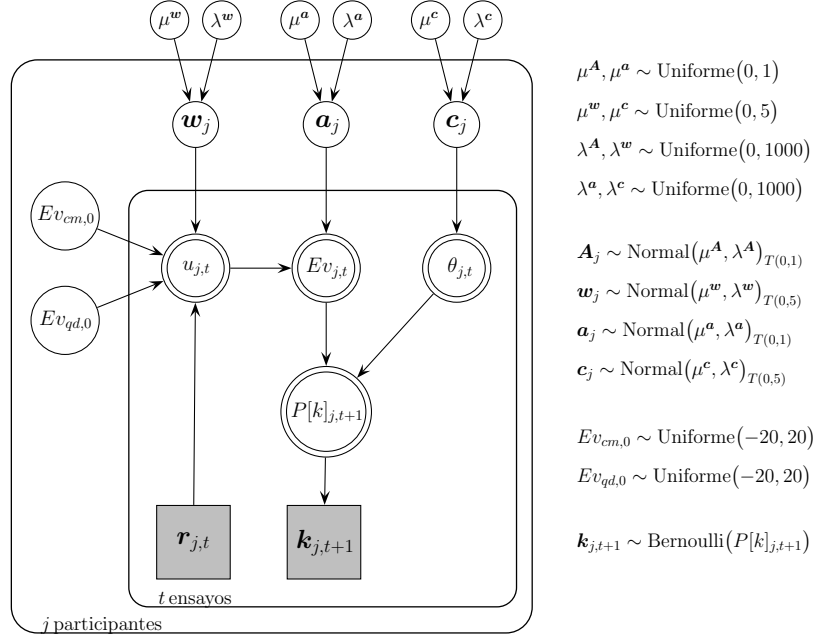


Figura 5-10: Representación gráfica del modelo  $\mathcal{M}_{EV \neq 0}$ .

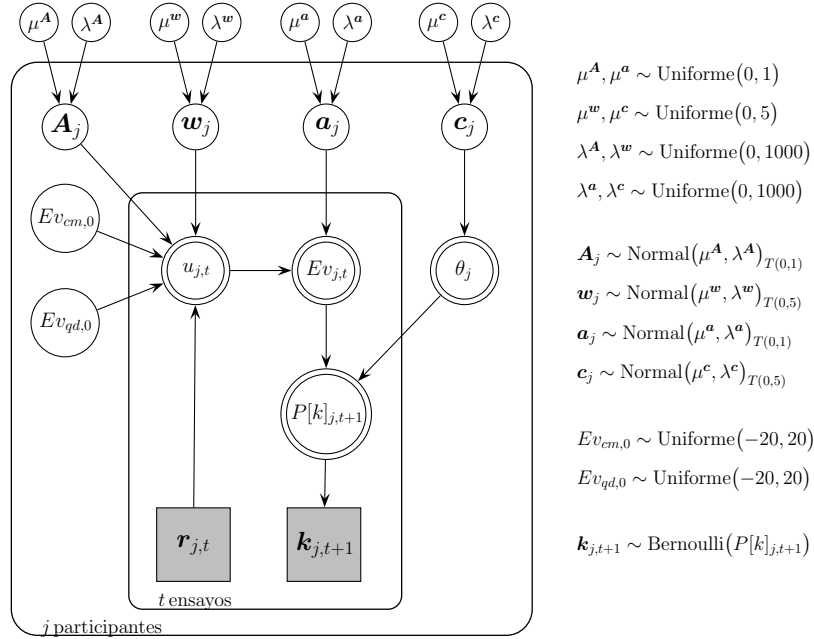


Figura 5-11: Representación gráfica de los modelos  $\mathcal{M}_{PVL \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$ .

Las figuras 5-10 y 5-11 presentan la especificación gráfica del modelo  $\mathcal{M}_{EV \neq 0}$  y de los modelos  $\mathcal{M}_{PVL \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$ , respectivamente. Tomando como base dicha especificación reprodujimos los análisis previamente detallados:

1. Implementamos los tres modelos en JAGS para calcular distribuciones posteriores sobre los parámetros desconocidos.
2. Implementamos el análisis por APHA $\neq 0$  utilizando las distribuciones posteriores inferidas. A diferencia del análisis anterior, en el paso 2 del método APHA $\neq 0$  no fijamos los valores esperados iniciales en cero. En cambio, tomamos un punto al azar de la distribución posterior conjunta  $\{Ev_{cm,0}, Ev_{qd,0}\}$  en cada muestreo aleatorio.
3. Calculamos las medidas  $APHA_{j,\mathcal{M}}^{\neq 0}$  de cada modelo para cada participante.
4. Implementamos el análisis por SIM $\neq 0$  utilizando las distribuciones posteriores inferidas. Nuevamente, tomamos puntos aleatorios de la distribución posterior conjunta  $\{Ev_{cm,0}, Ev_{qd,0}\}$  como valores iniciales en cada repetición del proceso de muestreo.
5. Calculamos las medidas  $SIM_{j,\mathcal{M}}^{\neq 0}$  de cada modelo para cada participante.

### Ajuste Post Hoc Absoluto $\neq 0$

Analizamos el APHA $\neq 0$  de los tres modelos empleando la representación gráfica y la medida de ajuste previamente descritas. Las figuras 5-12 y 5-13 corresponden con el APHA $\neq 0$  de los Participantes 2 y 6<sup>2</sup>. Para ambos participantes dos características del ajuste saltan a la vista. Primero, los tres modelos, pero particularmente los modelos  $\mathcal{M}_{EV \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$ , muestran un alto nivel de incertidumbre sobre las probabilidades ajustadas en los primeros ensayos de la tarea y el valor medio de dichas probabilidades parece diferente a 0.5. Esta regularidad refleja el hecho de que en el análisis APHA $\neq 0$  el valor de probabilidad inicial no es fijo y permite a los modelos ajustar el desempeño observado de ambos participantes durante la primera parte de la sesión. Segundo, aunque los tres modelos parecen aproximar el desempeño observado de cada participante en los últimos ensayos de la tarea, el modelo  $\mathcal{M}_{PVL \neq 0}$  ofrece el ajuste más aceptable

---

<sup>2</sup>Presentamos los resultados de los métodos APHA $\neq 0$  y SIM $\neq 0$  sobre los mismos participantes para facilitar la comparación de las propiedades de ajuste y de simulación de los modelos  $\mathcal{M}_{EV \neq 0}$ ,  $\mathcal{M}_{PVL \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$  en relación a las de los modelos originales.

porque sus predicciones promedio son más cercanas a las proporciones de cambio observadas y porque además es el modelo con menor incertidumbre en la segunda mitad de la sesión. Dichas características se ven reflejadas en las medidas de ajuste  $APHA_{j,\mathcal{M}}^{\neq 0}$ , que comparamos en el cuadro 5-5.

Participante	$\mathcal{M}_{EV \neq 0}$	$\mathcal{M}_{PVL \neq 0}$	$\mathcal{M}_{PVL\delta \neq 0}$	Mejor modelo
1	0.728	<b>0.509</b>	0.563	$\mathcal{M}_{PVL \neq 0}$
2	0.454	<b>0.407</b>	0.432	$\mathcal{M}_{PVL \neq 0}$
3	0.783	<b>0.576</b>	0.629	$\mathcal{M}_{PVL \neq 0}$
4	0.900	<b>0.548</b>	0.613	$\mathcal{M}_{PVL \neq 0}$
5	0.846	<b>0.544</b>	0.610	$\mathcal{M}_{PVL \neq 0}$
6	0.601	<b>0.461</b>	0.598	$\mathcal{M}_{PVL \neq 0}$
7	0.647	<b>0.411</b>	0.509	$\mathcal{M}_{PVL \neq 0}$
8	0.670	<b>0.510</b>	0.582	$\mathcal{M}_{PVL \neq 0}$
9	0.885	<b>0.568</b>	0.654	$\mathcal{M}_{PVL \neq 0}$
10	0.901	<b>0.564</b>	0.631	$\mathcal{M}_{PVL \neq 0}$
11	0.688	<b>0.541</b>	0.598	$\mathcal{M}_{PVL \neq 0}$
12	0.915	<b>0.509</b>	0.599	$\mathcal{M}_{PVL \neq 0}$
13	0.924	<b>0.568</b>	0.622	$\mathcal{M}_{PVL \neq 0}$
14	0.922	<b>0.582</b>	0.667	$\mathcal{M}_{PVL \neq 0}$
15	0.854	<b>0.532</b>	0.613	$\mathcal{M}_{PVL \neq 0}$
16	0.897	<b>0.558</b>	0.616	$\mathcal{M}_{PVL \neq 0}$
17	0.903	<b>0.529</b>	0.574	$\mathcal{M}_{PVL \neq 0}$

Tabla 5-5: Medidas de ajuste  $APHA_{j,\mathcal{M}}^{\neq 0}$  de cada modelo para cada participante.

Las medidas de ajuste revelan que el modelo  $\mathcal{M}_{PVL \neq 0}$  ofrece el mejor ajuste para todos los participantes. Esta regularidad es particularmente llamativa tomando en cuenta el hecho de que agregamos la misma cantidad de parámetros libres a los tres modelos originales.

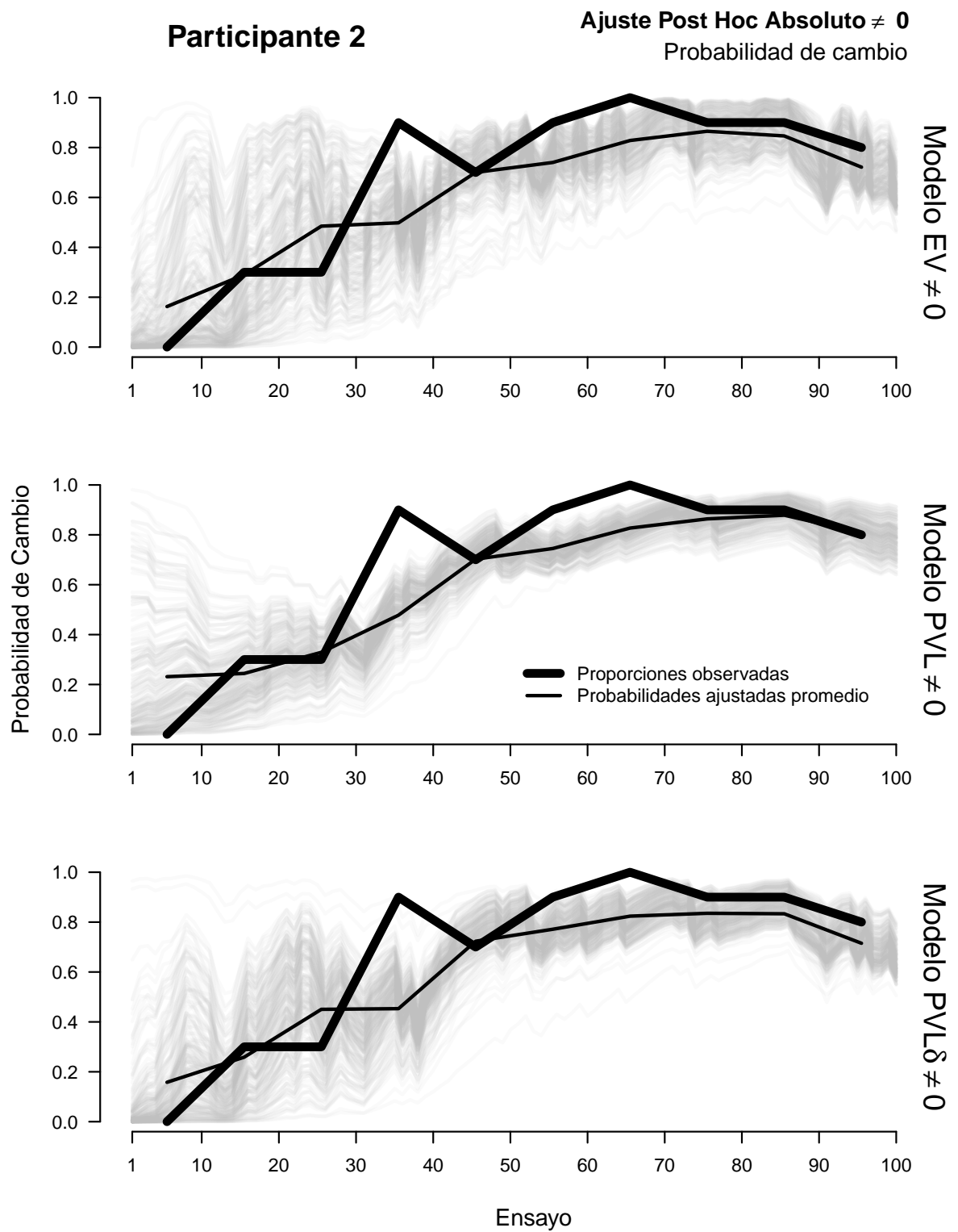


Figura 5-12: Ajuste Post Hoc Absoluto del Participante 2.



# Participante 6

Ajuste Post Hoc Absoluto  $\neq 0$   
Probabilidad de cambio

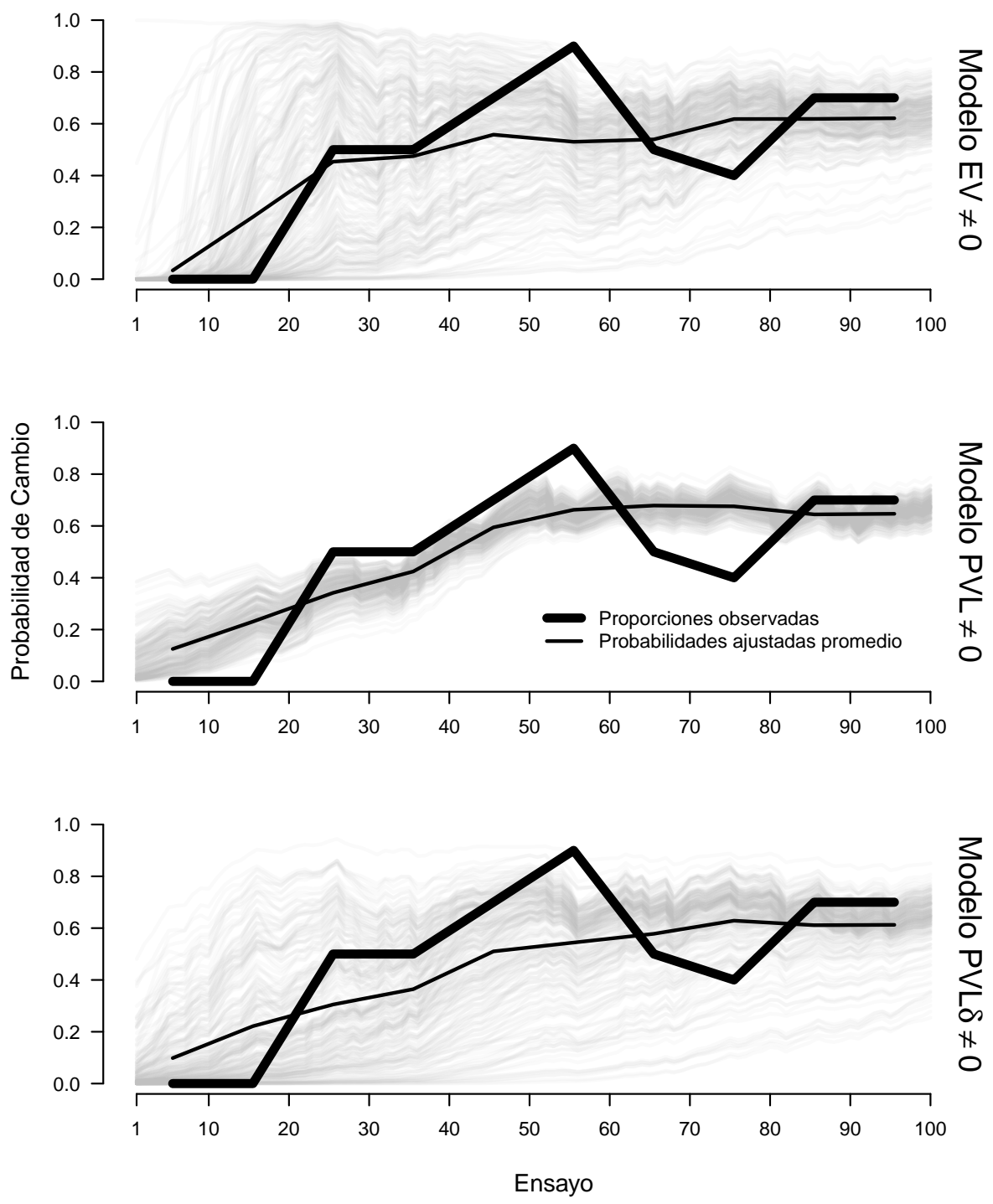


Figura 5-13: Ajuste Post Hoc Absoluto del Participante 6.

## Simulación $\neq 0$

El resultado del análisis  $SIM^{\neq 0}$  para los Participantes 2 y 6 aparece en las figuras 5-14 y 5-15, respectivamente. En ambos casos es posible apreciar a primera vista que la capacidad de simulación de las versiones extendidas  $\mathcal{M}_{EV \neq 0}$ ,  $\mathcal{M}_{PVL \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$  es mejor que la de las versiones originales: A diferencia del análisis anterior, esta vez el registro acumulativo observado generalmente termina en zonas de alta densidad de cambios simulados. Este resultado refleja dos características de las versiones extendidas de los tres modelos: 1) Mientras que las versiones originales tienden a cambiar a una tasa elevada en la primera parte de la sesión, las versiones extendidas comienzan a cambiar después, aproximadamente en los ensayos que registran los primeros cambios observados, y 2) los registros acumulativos de cambios simulados nuevamente muestran una pendiente similar a la del registro acumulativo de cambios observados. Debido a estas características las versiones modificadas  $\mathcal{M}_{EV \neq 0}$ ,  $\mathcal{M}_{PVL \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$  generan elecciones mucho más parecidas a las observadas en cada participante en comparación con las versiones originales.

El cuadro 5-6 evalúa la capacidad de simulación de las versiones modificadas para cada participante utilizando la medida  $SIM_{j, \mathcal{M}}^{\neq 0}$ . Como puede apreciarse en el cuadro 5-6, el modelo  $\mathcal{M}_{PVL \neq 0}$  simula mejor a 9 de los 17 participantes, mientras que los modelos  $\mathcal{M}_{PVL\delta \neq 0}$  y  $\mathcal{M}_{EV \neq 0}$  simulan adecuadamente a 6 y a 2 participantes, respectivamente.

El resultado de las versiones modificadas  $\mathcal{M}_{EV \neq 0}$ ,  $\mathcal{M}_{PVL \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$  en los métodos APHA y SIN se resume en el cuadro 5-7. Es importante resaltar que en los análisis de las versiones modificadas el mejor modelo en APHA $^{\neq 0}$  coincide con el mejor modelo en  $SIM^{\neq 0}$  en 9 de los 17 participantes. El modelo  $\mathcal{M}_{PVL \neq 0}$  resulta exitoso al ajustar a toda la muestra y al simular adecuadamente a la mayoría de participantes, y en consecuencia es el modelo más consistente entre las tres versiones modificadas.

## Participante 2

Simulación  $\neq 0$   
Cambios simulados acumulados

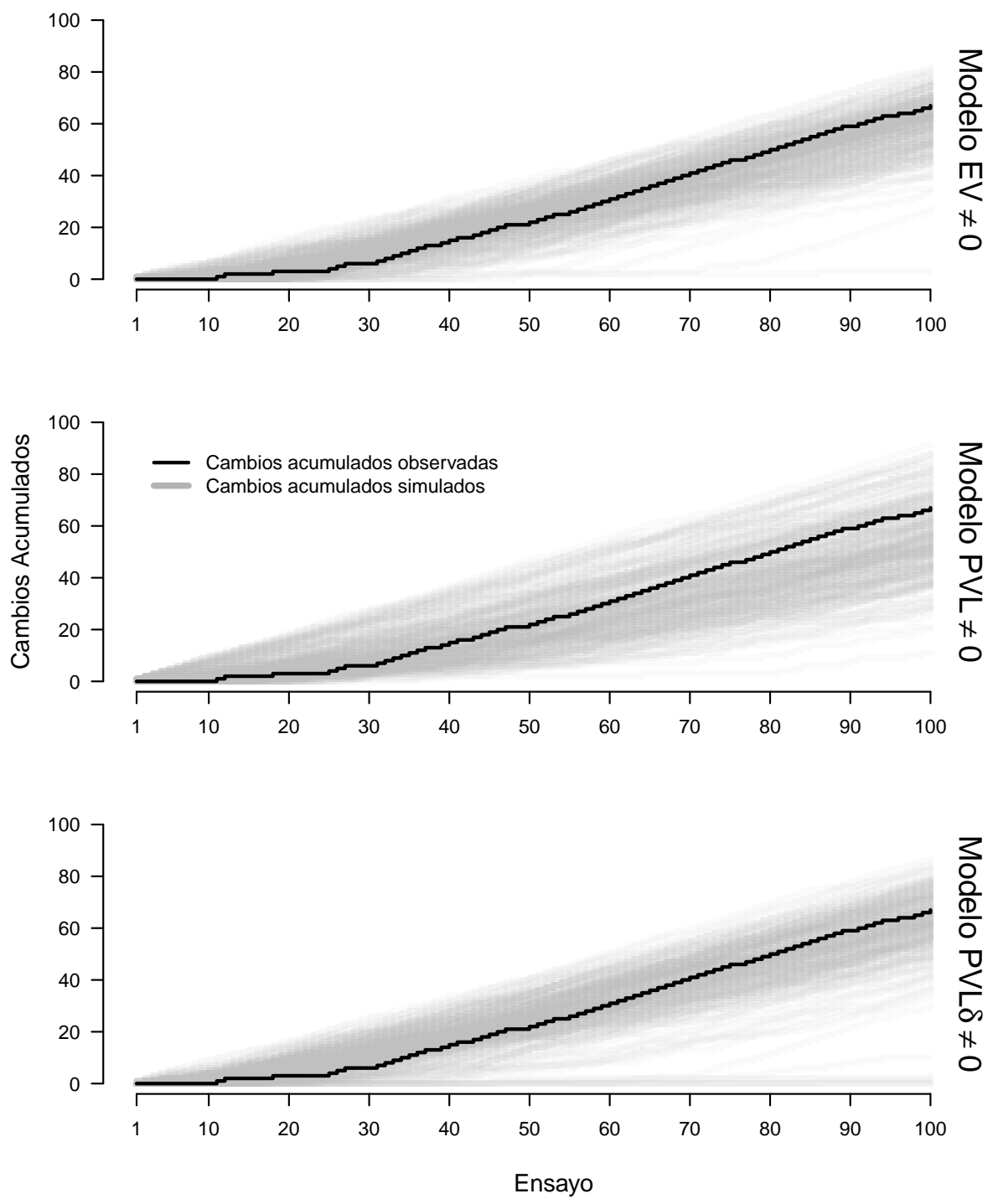


Figura 5-14: Cambios simulados acumulados del Participante 2.

# Participante 6

Simulación  $\neq 0$   
Cambios simulados acumulados

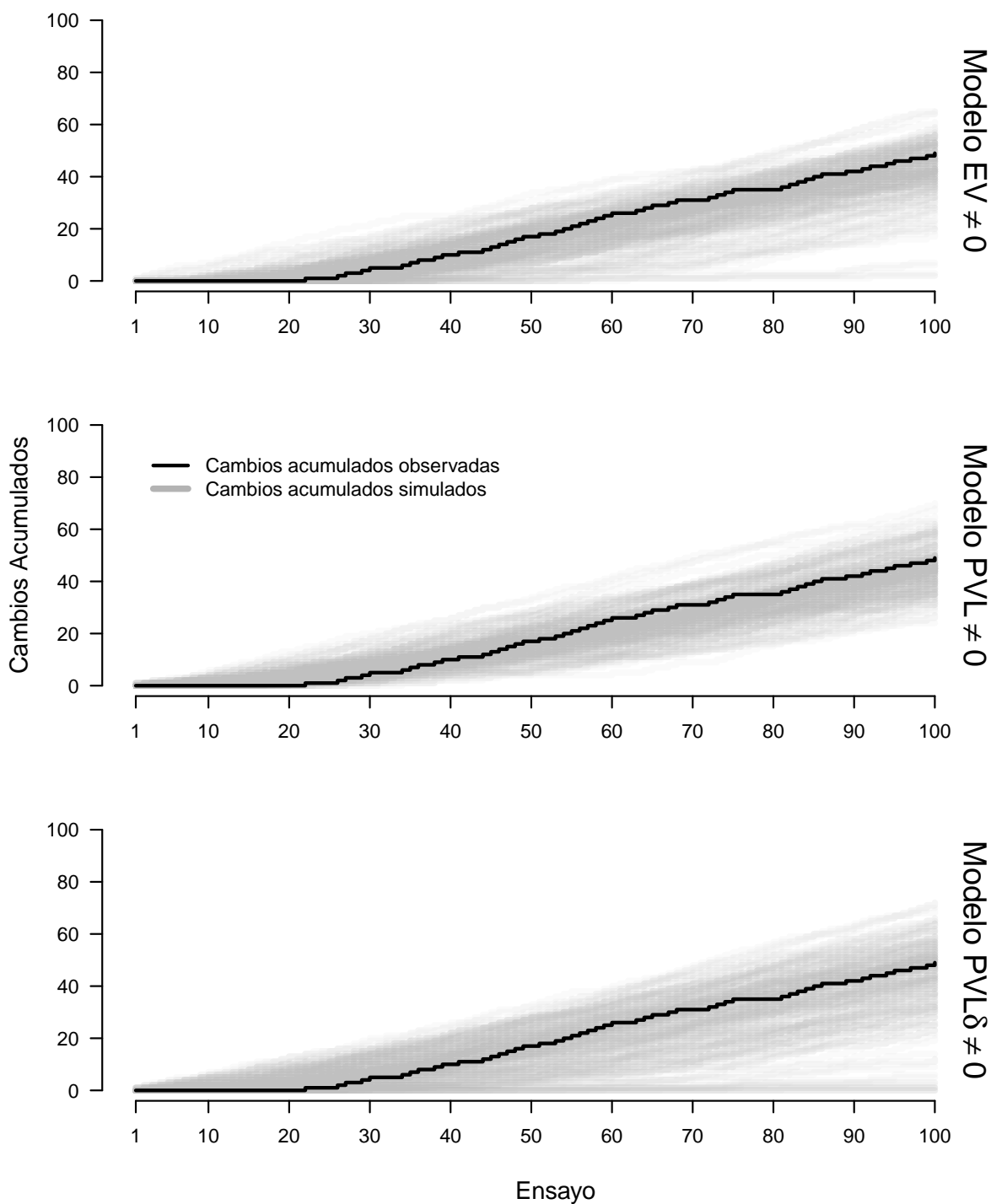


Figura 5-15: Cambios simulados acumulados del Participante 6.

Participante	$\mathcal{M}_{\text{EV} \neq 0}$	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL}\delta \neq 0}$	Mejor modelo
1	20.77	13.35	<b>11.20</b>	$\mathcal{M}_{\text{PVL}\delta \neq 0}$
2	<b>9.18</b>	13.86	11.42	$\mathcal{M}_{\text{EV} \neq 0}$
3	21.51	15.91	<b>12.26</b>	$\mathcal{M}_{\text{PVL}\delta \neq 0}$
4	19.47	<b>7.72</b>	8.81	$\mathcal{M}_{\text{PVL} \neq 0}$
5	13.04	<b>6.30</b>	9.12	$\mathcal{M}_{\text{PVL} \neq 0}$
6	9.40	<b>7.73</b>	11.80	$\mathcal{M}_{\text{PVL} \neq 0}$
7	29.95	<b>20.39</b>	26.04	$\mathcal{M}_{\text{PVL} \neq 0}$
8	28.47	<b>16.83</b>	22.33	$\mathcal{M}_{\text{PVL} \neq 0}$
9	28.55	<b>10.86</b>	13.43	$\mathcal{M}_{\text{PVL} \neq 0}$
10	21.22	<b>17.24</b>	22.50	$\mathcal{M}_{\text{PVL} \neq 0}$
11	<b>10.67</b>	11.65	13.96	$\mathcal{M}_{\text{EV} \neq 0}$
12	31.75	14.70	<b>11.80</b>	$\mathcal{M}_{\text{PVL}\delta \neq 0}$
13	25.84	10.55	<b>10.26</b>	$\mathcal{M}_{\text{PVL}\delta \neq 0}$
14	27.89	<b>11.28</b>	12.69	$\mathcal{M}_{\text{PVL} \neq 0}$
15	20.10	11.20	<b>10.26</b>	$\mathcal{M}_{\text{PVL}\delta \neq 0}$
16	21.18	10.29	<b>10.04</b>	$\mathcal{M}_{\text{PVL}\delta \neq 0}$
17	28.96	<b>8.24</b>	9.85	$\mathcal{M}_{\text{PVL} \neq 0}$

Tabla 5-6: Medidas de simulación  $SIM_{j,\mathcal{M}}^{\neq 0}$  de cada modelo para cada participante.

Participante	Mejor modelo APHA <sup>≠0</sup>	Mejor modelo SIM <sup>≠0</sup>
1	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL}\delta \neq 0}$
2	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{EV} \neq 0}$
3	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL}\delta \neq 0}$
4	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL} \neq 0}$
5	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL} \neq 0}$
6	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL} \neq 0}$
7	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL} \neq 0}$
8	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL} \neq 0}$
9	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL} \neq 0}$
<b>10</b>	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL} \neq 0}$
11	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{EV} \neq 0}$
12	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL}\delta \neq 0}$
13	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL}\delta \neq 0}$
<b>14</b>	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL} \neq 0}$
15	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL}\delta \neq 0}$
16	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL}\delta \neq 0}$
<b>17</b>	$\mathcal{M}_{\text{PVL} \neq 0}$	$\mathcal{M}_{\text{PVL} \neq 0}$

Tabla 5-7: Resumen de los métodos APHA<sup>≠0</sup> y SIM<sup>≠0</sup>. Cada renglón presenta el modelo que mejor ajusta y el modelo que mejor simula a cada participante.



## Capítulo 6

# Discusión

El objetivo de este trabajo fue evaluar la capacidad de tres modelos de aprendizaje por reforzamiento para explicar la secuencia de elecciones humanas en una versión repetida del PMH. Nuestros resultados sugieren que entre las versiones originales el modelo  $\mathcal{M}_{\text{PVL}\delta}$  ofrece el mejor ajuste y la mejor simulación para la mayoría de participantes. Este hallazgo coincide con reportes previos y sugiere que el modelo  $\mathcal{M}_{\text{PVL}\delta}$  es una descripción aceptable del proceso de aprendizaje humano en ambientes inciertos cuando se asume que las expectativas iniciales son iguales para todas las opciones de respuesta. Sin embargo, al analizar la capacidad de ajuste y simulación de las versiones modificadas de los tres modelos descubrimos que el modelo  $\mathcal{M}_{\text{PVL}\neq 0}$  consistentemente ajusta a toda la muestra de participantes y simula adecuadamente a más de la mitad, lo cual constituye un desempeño superior al obtenido por el modelo  $\mathcal{M}_{\text{PVL}\delta}$  original.

Interpretamos estos resultados como evidencia de que el modelo  $\mathcal{M}_{\text{PVL}\neq 0}$  captura aspectos importantes del proceso de aprendizaje humano en ambientes inciertos y resulta particularmente adecuado para modelar situaciones en las que las expectativas iniciales sobre cada opción de respuesta pueden ser diferentes de cero. En las secciones siguientes discutimos las implicaciones de nuestros hallazgos en relación a los procesos psicológicos que operan al enfrentar ambientes inciertos. Estructuramos la sección de discusión en torno a cuatro preguntas centrales:

*¿Cuáles son las expectativas iniciales sobre “cambiar” y “quedarse”?*

*¿Cómo se modifica el valor esperado de ambas alternativas a lo largo de la sesión?*

*¿Qué papel juegan las pérdidas sobre la utilidad de cada respuesta?*

*¿Todos los participantes aprenden igual? ¿En qué se parecen? ¿En qué difieren?*



*¿Cuáles son las expectativas iniciales sobre “cambiar” y “quedarse”?* Las versiones modificadas  $\mathcal{M}_{EV \neq 0}$ ,  $\mathcal{M}_{PVL \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$  relajan el supuesto de que los valores esperados iniciales de las opciones de respuesta “cambiar” y “quedarse” son cero y en cambio tratan a dichos valores como parámetros libres individuales. En consecuencia, es posible estimar su valor con base en los datos observados. La figura 6-1 presenta las distribuciones posteriores sobre los valores esperados iniciales de ambas opciones para cada participante. Cada panel de la figura corresponde con cada modelo. Las líneas claras representan la densidad posterior del valor esperado inicial de “quedarse” de cada participante, mientras que las líneas oscuras corresponden con la opción “cambiar”. Como puede apreciarse en la figura, la mayoría de las líneas grises y claras se ubican sobre el mismo rango de valores según el modelo  $\mathcal{M}_{EV \neq 0}$ , pero aparecen consistentemente separadas de acuerdo con los modelos  $\mathcal{M}_{PVL \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$ . En otras palabras, el modelo  $\mathcal{M}_{EV \neq 0}$  sugiere que deberíamos creer que la mayoría de participantes se comporta como si la expectativa inicial de “cambiar” fuera igual a la de “quedarse”. En contraste, los modelos  $\mathcal{M}_{PVL \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$  sugieren que los participantes se comportan como si la expectativa inicial de “cambiar” fuera mucho menor que la de “quedarse”. Este resultado refleja la robusta tendencia poblacional a “no cambiar” de puerta la primera vez que se enfrenta el PMH, y ofrece la ventaja adicional de que proporciona una medida del tamaño de dicha tendencia.

Es importante resaltar que los parámetros desconocidos que corresponden a los valores iniciales esperados están especificados a nivel individual, de la misma manera que los parámetros principales de todos los modelos. Sin embargo, en las versiones modificadas no especificamos restricciones jerárquicas sobre los valores esperados iniciales de ambas opciones. Una extensión interesante podría consistir, precisamente, en asumir no sólo que los valores iniciales esperados son parámetros desconocidos para cada participante, sino que todos los participantes provienen de la misma población respecto de los valores esperados iniciales. Aparte de restringir la variación individual inferida respecto de los valores esperados iniciales en las versiones modificadas reportadas, esta extensión potencial permitiría predecir el desempeño de un participante nuevo sin asumir un valor esperado fijo para “quedarse” o “cambiar”, sino infiriéndolo a partir del conocimiento extraído de la población analizada en este estudio.

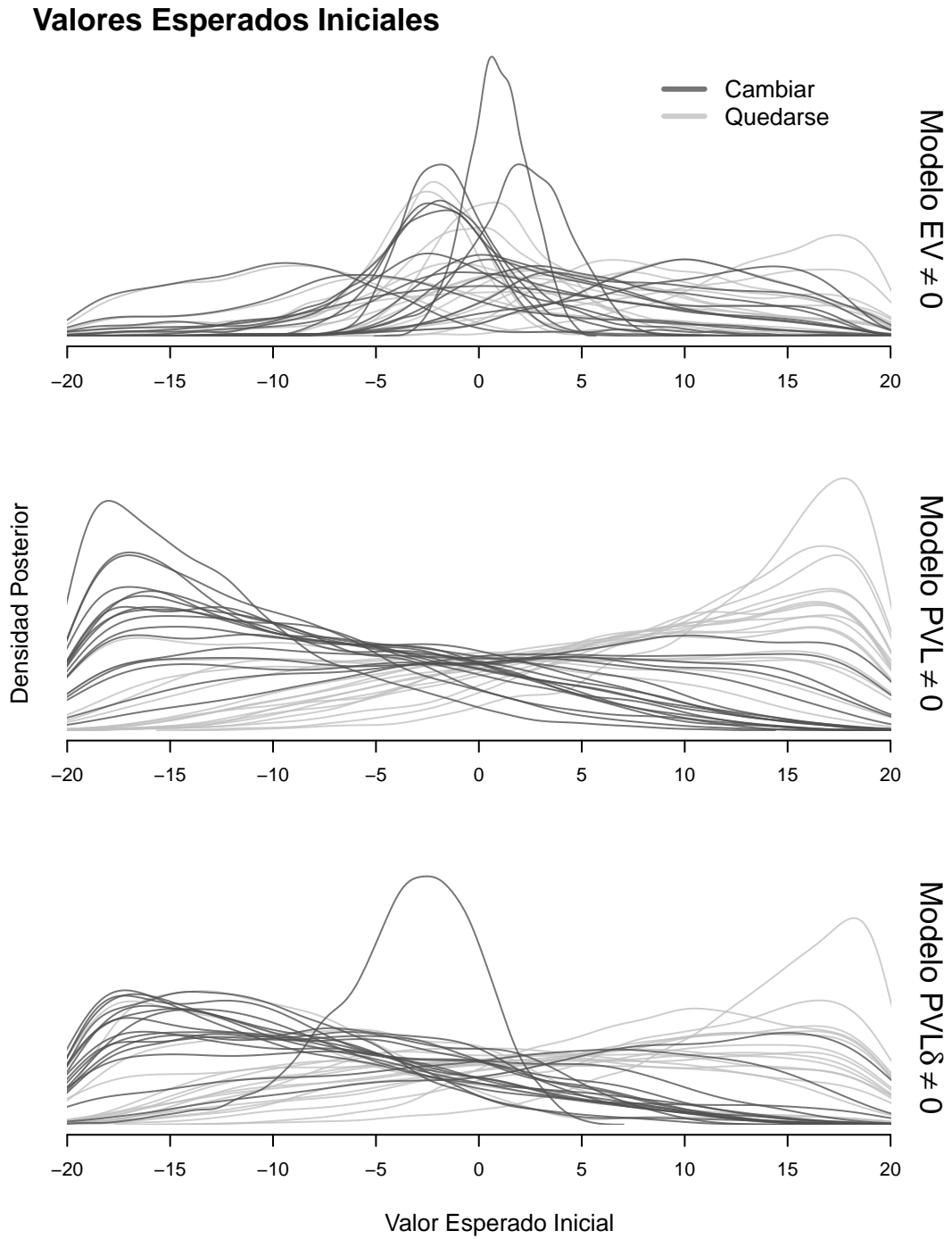


Figura 6-1: Valores esperados iniciales inferidos de acuerdo con los modelos  $\mathcal{M}_{EV \neq 0}$ ,  $\mathcal{M}_{PVL \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$ . En color claro se muestran las distribuciones posteriores del valor esperado inicial de “quedarse” y en color oscuro las de “cambiar”. Cada línea representa un participante.

*¿Cómo se modifica el valor esperado de ambas alternativas a lo largo de la sesión?* Los tres modelos originales y las versiones modificadas suponen que los valores esperados de ambas alternativas se modifican en la medida que los participantes las exploran y experimentan las consecuencias de elegir cada una. Podemos examinar las distribuciones posteriores de los valores esperados según cada modelo para analizar en detalle *cómo* y *cuánto* se modifican los valores esperados iniciales de cada opción a lo largo de la sesión. Un ejemplo de este análisis se resume en la figura 6-2. En el panel superior graficamos las elecciones observadas del Participante 7 en cada ensayo de la sesión. Los puntos de la hilera superior representan los “cambios” y los de la parte inferior los “no cambios”. La clave de color indica los ensayos en los que el participante encontró (negro) o no encontró (blanco) el premio. Como se espera de acuerdo con la probabilidad de ganar en cada opción del PMH, la proporción de puntos negros es mayor en la hilera superior que en la inferior (aproximadamente 2/3 y 1/3, respectivamente). Esta gráfica permite identificar cuatro patrones de elecciones del participante: 1) Al inicio de la sesión el participante explora ambas opciones, en el sentido de que alterna entre ambas alternativas en los primeros 10 ensayos. 2) Después del primer período de exploración el participante elige una alternativa (“cambiar”) a una tasa alta durante los siguientes 40 ensayos, muestreando ocasionalmente la otra alternativa. 3) Aproximadamente a la mitad de la sesión el participante vuelve a explorar ambas alternativas y finalmente 4) el participante explota la alternativa “cambiar” en el último cuarto de la sesión.

En la misma figura, los paneles del segundo renglón presentan los valores esperados posteriores de ambas opciones a lo largo de la sesión, según cada modelo. En cada panel, la línea oscura une las medias posteriores de los valores esperados de “cambiar” en todos los ensayos, mientras que la línea clara une las medias posteriores de los valores esperados de “quedarse”. La zona sombreada oscura demarca el intervalo de mayor credibilidad posterior (95 %) del valor esperado de “cambiar”, y la zona sombreada clara el de “quedarse”. En estos tres paneles una diferencia notable salta a la vista: Mientras que los modelos  $\mathcal{M}_{EV \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$  infieren que los valores esperados de ambas opciones permanecen similares a lo largo de la sesión, el modelo  $\mathcal{M}_{PVL \neq 0}$  sugiere que éstos se vuelven cada vez más diferentes. En otras palabras, el modelo  $\mathcal{M}_{PVL \neq 0}$  parece sensible a la diferencia en tasas de pagos experimentadas por el participante y sugiere que a lo largo de la sesión se vuelve cada vez más fácil distinguirlas.

## Participante 7

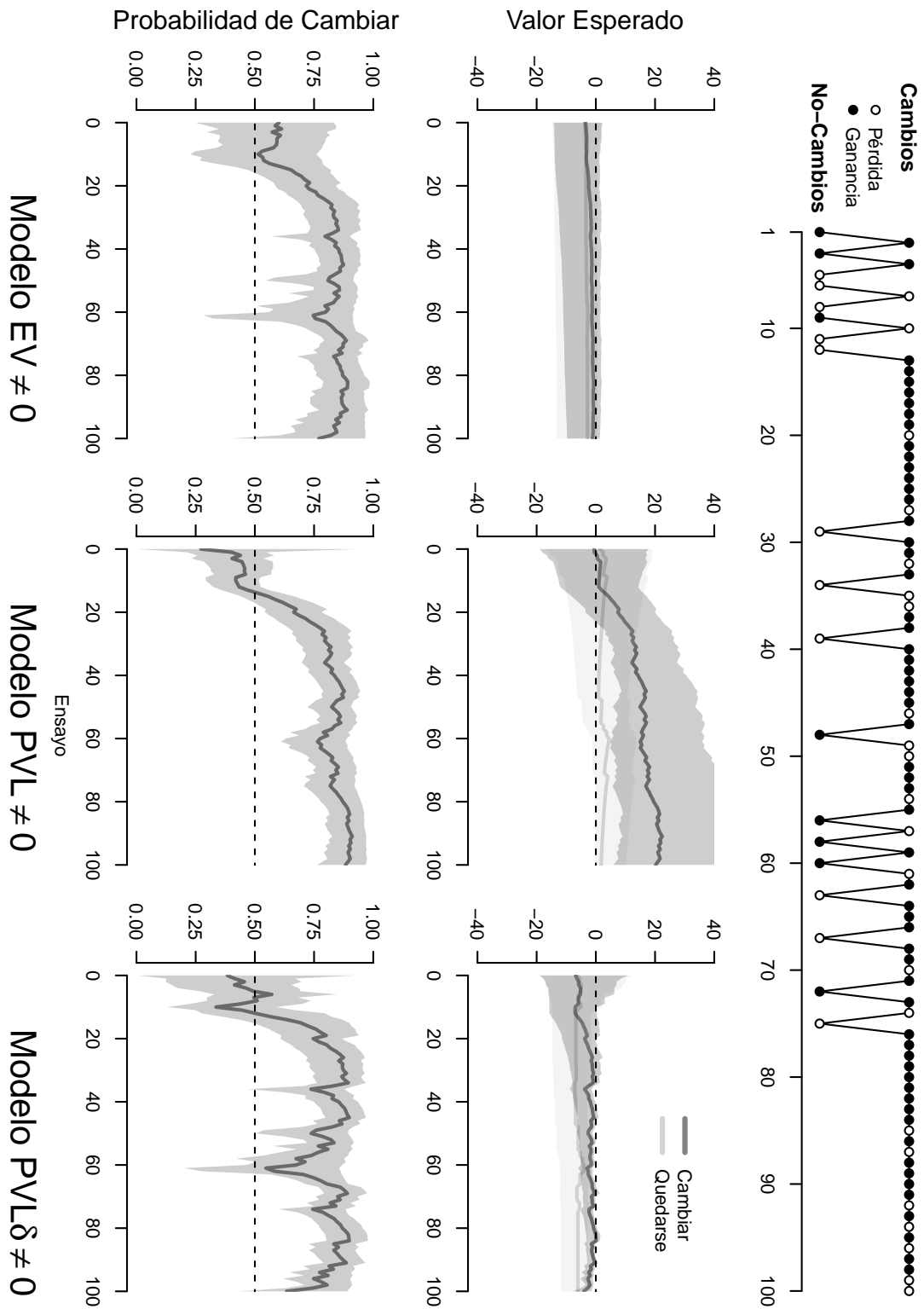


Figura 6-2: Valores esperados y probabilidad de “cambiar” a lo largo de la sesión para el Participante 7.

Como resultado, el modelo  $\mathcal{M}_{\text{PVL} \neq 0}$  concluye que la probabilidad de cambiar a lo largo de la sesión primero crece y después se vuelve estable, a diferencia de los modelos  $\mathcal{M}_{\text{EV} \neq 0}$  y  $\mathcal{M}_{\text{PVL}\delta \neq 0}$ . Como puede apreciarse en las gráficas del último renglón de la figura, los modelos  $\mathcal{M}_{\text{EV} \neq 0}$  y  $\mathcal{M}_{\text{PVL}\delta \neq 0}$  infieren que la probabilidad de cambiar es altamente volátil en cada ensayo y son altamente sensibles a pequeñas fluctuaciones en la tasas de cambios y de pagos observadas. En particular, ambos modelos muestran una alta incertidumbre sobre la probabilidad de cambiar alrededor del ensayo 60 e incluso aceptan valores menores a 0.5 como valores de probabilidad de cambiar creíbles en dichos ensayos. En contraste, aunque la media posterior de la probabilidad de cambiar inferida por el modelo  $\mathcal{M}_{\text{PVL} \neq 0}$  también disminuye a la mitad de la tarea, el decremento no es tan grande y la zona de mayor credibilidad siempre se ubica por encima de 0.5. Esta diferencia sugiere que los modelos  $\mathcal{M}_{\text{EV} \neq 0}$  y  $\mathcal{M}_{\text{PVL}\delta \neq 0}$  concluyen que a la mitad de la sesión el Participante 7 mostró un cambio de preferencias genuino, para después cambiar de nuevo, mientras que el modelo  $\mathcal{M}_{\text{PVL} \neq 0}$  interpreta a los “no-cambios” observados entre los ensayos 55 y 75 como un producto aleatorio de un proceso Bernoulli estable.

*¿Qué papel juegan las pérdidas sobre la utilidad de cada respuesta?* Una de las ventajas del análisis que hemos presentado es que aparte de evaluar el ajuste y la capacidad de simulación de cada modelo podemos aprender sobre los parámetros individuales de cada participante examinando las distribuciones posteriores correspondientes. En el caso del componente de utilidad, por ejemplo, podemos examinar directamente la forma de la función de valor de los modelos  $\mathcal{M}_{\text{PVL} \neq 0}$  y  $\mathcal{M}_{\text{PVL}\delta \neq 0}$ . La figura 6-3 resume dicho análisis tomando como ejemplo al Participante 4. En los paneles laterales graficamos la distribución posterior sobre los parámetros  $w_4$  y  $A_4$ , y en los paneles principales presentamos 100 formas creíbles de la función de valor del participante basadas en muestreos de la distribución posterior conjunta  $\{w_4, A_4\}$  según cada modelo.

Pese a que ambos modelos utilizan el mismo componente de utilidad y la misma función de valor, las curvas de los paneles centrales de la figura 6-3 tienen formas muy diferentes. Por un lado, el modelo  $\mathcal{M}_{\text{PVL} \neq 0}$  sugiere que cualquier pérdida, sin importar su valor absoluto, es percibida como nula, mientras que cualquier ganancia, nuevamente al margen de su valor absoluto, es percibida como uno. En contraste, la función de valor inferida por el modelo  $\mathcal{M}_{\text{PVL}\delta \neq 0}$  indica que tanto pérdidas como ganancias tienden a ser directamente proporcionales a sus valores

absolutos, y además el valor percibido de las pérdidas es aproximadamente cinco veces su valor absoluto.

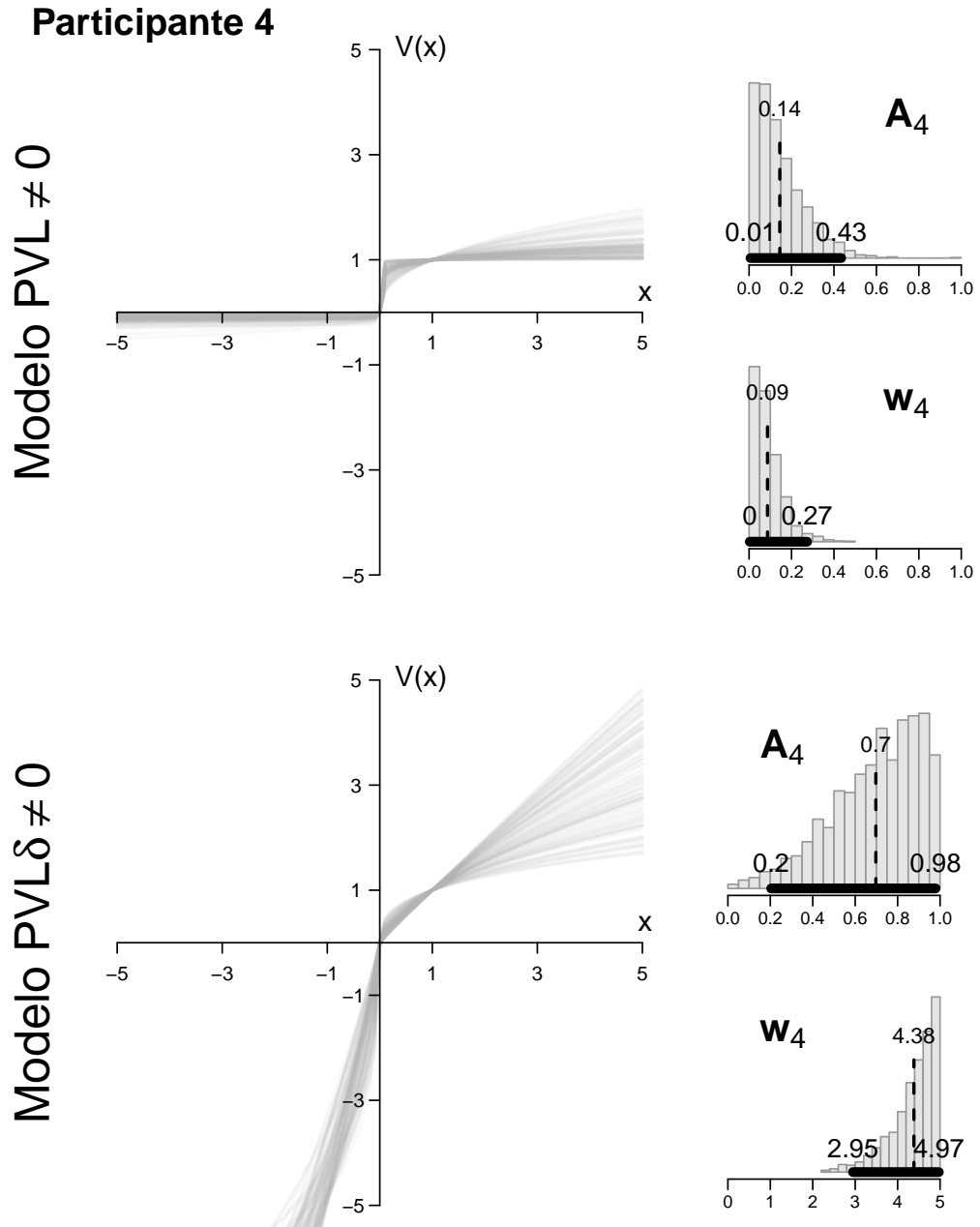


Figura 6-3: Funciones de Valor del Participante 4 de acuerdo con los modelos  $\mathcal{M}_{PVL \neq 0}$  y  $\mathcal{M}_{PVL\delta \neq 0}$ .

Las diferencias de forma en la función de valor de ambos modelos reflejan los valores paramétricos inferidos por cada uno. El modelo  $\mathcal{M}_{\text{PVL} \neq 0}$  concluye que, dados los datos, los valores más probables de los parámetros  $\mathbf{w}_4$  y  $\mathbf{A}_4$  son cercanos a cero. En la función de valor (ecuaciones (PVL-1) y (PVLd-1)) esta combinación de valores corresponde con el caso en el que las pérdidas son ignoradas (parámetro  $\mathbf{w} = 0$ ) y en el que todas las ganancias valen lo mismo (parámetro  $\mathbf{A} = 0$ ). Por su parte, el modelo  $\mathcal{M}_{\text{PVL}\delta \neq 0}$  llega a la conclusión contraria y sugiere que los valores más probables de ambos parámetros son valores altos en sus respectivos rangos. Valores altos de los parámetros  $\mathbf{w}$  y  $\mathbf{A}$  implican que las pérdidas tienen un valor relativo mucho mayor que las ganancias ( $\mathbf{w} = 5$ ) y que el valor subjetivo de la apuesta es una función lineal de su valor absoluto ( $\mathbf{A} = 1$ ).

En tanto que los resultados de los métodos de evaluación descritos previamente sugieren que el modelo  $\mathcal{M}_{\text{PVL} \neq 0}$  ofrece una mejor explicación del patrón de elecciones del Participante 4 en comparación con la del modelo  $\mathcal{M}_{\text{PVL}\delta \neq 0}$ , podemos concluir que la forma de la función de valor inferida por el modelo  $\mathcal{M}_{\text{PVL} \neq 0}$  describe mejor la representación subjetiva de diferentes montos de ganancias y pérdidas en este participante. Aparte, la discrepancia entre las conclusiones de ambos modelos respecto de la forma de la función de valor sugieren nuevas manipulaciones experimentales para evaluar las predicciones de ambos modelos. Por ejemplo, estudios futuros podrían evaluar el desempeño en el PMH repetido con montos de pérdida nulos, pequeños, y muy grandes. Si el desempeño observado difiere entre las diferentes condiciones de monto de pérdida, el modelo  $\mathcal{M}_{\text{PVL}\delta \neq 0}$  ganaría soporte. Por el contrario, si el patrón de “cambios” es similar al margen del monto de pérdida, como sugiere el modelo  $\mathcal{M}_{\text{PVL} \neq 0}$ , dicha evidencia constituiría soporte adicional para este modelo.

*¿Todos los participantes aprenden igual? ¿En qué se parecen? ¿En qué difieren?* Aparte de aprender sobre cada participante, el método de inferencia paramétrica jerárquico que hemos presentado permite describir la población de participantes en términos de la tendencia central y de la variabilidad individual de cada parámetro inferida por cada modelo. Como ejemplo, la figura 6-4 presenta las distribuciones posteriores de cada parámetro individual y la distribución posterior de la media poblacional para cada parámetro y cada modelo. Como puede apreciarse en los paneles superiores de la figura 6-4, prácticamente todas las distribuciones posteriores

individuales son iguales para los parámetros  $\mathbf{w}$  y  $\mathbf{A}$ , correspondientes al componente de utilidad de los tres modelos. Consistente con este resultado, la distribución posterior de la media poblacional tiene una forma muy similar a la de cada distribución posterior individual en estos parámetros. En otras palabras, las extensiones jerárquicas de los modelos  $\mathcal{M}_{\text{EV} \neq 0}$ ,  $\mathcal{M}_{\text{PVL} \neq 0}$  y  $\mathcal{M}_{\text{PVL}\delta \neq 0}$  sugieren que todos los participantes ponderan las pérdidas y ganancias de manera similar (parámetro  $\mathbf{w}$ ) y, en el caso de los modelos  $\mathcal{M}_{\text{PVL} \neq 0}$  y  $\mathcal{M}_{\text{PVL}\delta \neq 0}$ , que la forma de la función de valor es muy parecida entre participantes (parámetro  $\mathbf{A}$ ).

Por el contrario, los participantes parecen ser diferentes tanto en el componente que integra la utilidad de cada ensayo con la experiencia previa en cada alternativa, como en el componente que decide qué alternativa elegir con base en el valor esperado de todas las alternativas disponibles.

En el caso del componente de actualización inferimos una variabilidad individual considerable en el parámetro  $\mathbf{a}$ , que corresponde con la tasa de actualización en los modelos  $\mathcal{M}_{\text{EV}}$  y  $\mathcal{M}_{\text{PVL}\delta \neq 0}$ , y con el parámetro de recencia en el modelo  $\mathcal{M}_{\text{PVL} \neq 0}$ . A diferencia de los parámetros  $\mathbf{w}$  y  $\mathbf{A}$ , las distribuciones posteriores de los parámetros  $\mathbf{a}$  son diferentes entre participantes, aunque se ubican en un rango relativamente estrecho del rango válido de este parámetro. La distribución posterior de la media poblacional del parámetro  $\mathbf{a}$  esta vez se ubica en la zona de mayor densidad posterior entre participantes. En otras palabras, los tres modelos sugieren que, aunque para todos los participantes la experiencia acumulada previa tiene más peso que el ensayo inmediato anterior al actualizar el valor esperado de cada alternativa<sup>1</sup>, algunos participantes le otorgan más peso que otros.

Un resultado similar se observa al analizar las distribuciones posteriores individuales y poblacionales sobre  $\mathbf{c}$  en los tres modelos. En particular, al comparar los parámetros  $\mathbf{c}_j$  inferidos por los modelos  $\mathcal{M}_{\text{PVL} \neq 0}$  y  $\mathcal{M}_{\text{PVL}\delta \neq 0}$  se puede apreciar que la dispersión individual inferida por el primer modelo es menor que la inferida por el segundo, incluso cuando ambos utilizan la misma función de consistencia (ecuaciones (PVL-4) y (PVLd-4)).

---

<sup>1</sup>Las distribuciones posteriores sobre  $\mathbf{a}$  en los tres modelos apuntan a la misma conclusión aunque difieren en rangos porque la interpretación de  $\mathbf{a}$  tiene el sentido contrario en el modelo  $\mathcal{M}_{\text{PVL} \neq 0}$ , que utiliza la *regla de actualización por decaimiento*, respecto de los modelos  $\mathcal{M}_{\text{EV} \neq 0}$  y  $\mathcal{M}_{\text{PVL}\delta \neq 0}$ , que comparten la *regla delta* (ver ecuaciones (EV-2), (PVL-2) y (PVLd-2)).



## Distribuciones Posteriores Individuales y Poblacionales

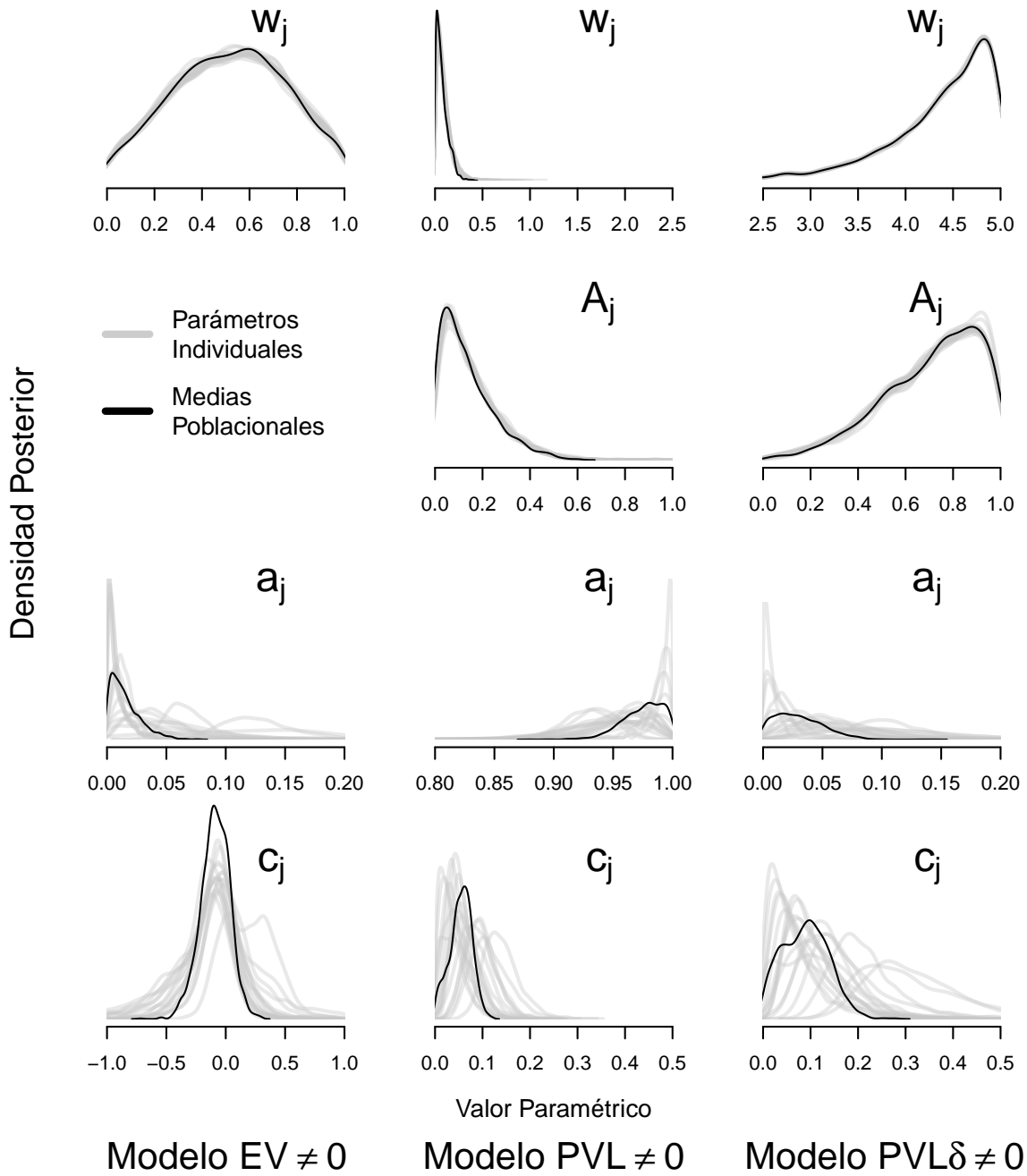


Figura 6-4: Distribuciones posteriores de los parámetros individuales y de la media poblacional de cada parámetro en cada modelo.

Los resultados anteriores suponen que todos los participantes provienen de una sola población respecto de cada parámetro. Este supuesto puede no ser el caso. Por ejemplo, suponiendo que en la población humana hay algunos individuos con un parámetro de recencia alto, otros con un parámetro de recencia bajo, y otros con parámetros de recencia en diferentes rangos intermedios, la mejor manera de describir dicha población y cada participante en ella consiste en asumir que cada individuo puede provenir de alguna de diferentes subpoblaciones o grupos, cada uno caracterizado por cierta distribución de valores paramétricos. El problema de inferencia en este ejemplo hipotético consiste en especificar no sólo cuál es el valor paramétrico individual, sino de qué grupo proviene cada individuo, cuántos grupos hay en la población, y cuál es el valor paramétrico medio en cada grupo. Desafortunadamente, al asumir un número de poblaciones fijo, las extensiones jerárquicas que hemos implementado en este trabajo no pueden responder dichas preguntas.

En años recientes se ha desarrollado un conjunto de técnicas de inferencia que permiten solucionar este problema. Dichas técnicas, conocidas como *métodos Bayesianos no paramétricos*, asumen que cada individuo en una población puede provenir de diferentes subgrupos o componentes, y permiten agrupar a individuos que se parecen entre sí en cierto atributo (Gershman y Blei, 2012). Esta aproximación puede resultar productiva para identificar a diferentes grupos de participantes en términos de su desempeño en el PMH y de los diferentes valores paramétricos en los modelos que hemos evaluado (o en otros), y potencialmente puede ofrecer nuevas explicaciones para entender por qué ciertos individuos parecen tener facilidad para resolver el PMH mientras que otros no.



## Capítulo 7

# Conclusiones

Los resultados de este trabajo sugieren que una versión modificada del modelo  $\mathcal{M}_{\text{PVL}}$  (Ahn et al., 2008, 2011), el modelo  $\mathcal{M}_{\text{PVL} \neq 0}$ , ofrece el mejor ajuste y muestra la mejor capacidad de simulación de un conjunto de observaciones recolectadas en el Problema de Monty Hall repetido.

Después de ver nuestros datos, dicho modelo concluye que en el PMH las pérdidas tienden a ser ignoradas y cualquier monto de ganancia tiene el mismo valor subjetivo. Ambas conclusiones sugieren manipulaciones experimentales adicionales para evaluar las predicciones de este y de los diferentes modelos evaluados en este trabajo.

Las técnicas de evaluación de modelos que hemos utilizado permiten no sólo evaluar el desempeño absoluto de cada modelo, sino también computar distribuciones posteriores sobre cada parámetro en cada modelo para cada individuo. Al analizar la dispersión de los valores paramétricos individuales descubrimos que nuestra muestra de estudiantes es similar en algunos componentes pero muy variable en otros. La variabilidad individual inferida sugiere que los individuos en la muestra pueden agruparse en diferentes subpoblaciones respecto de algunos parámetros pero en una sola población respecto de otros.

La similitud entre nuestros resultados y los previamente reportados utilizando la Tarea de Selección de Iowa (Steingroever et al., 2014, 2013a, 2013b) sugieren que el PMH puede verse como un problema de bandido al menos con respecto a las opciones “cambiar” o “quedarse” (Wetzels et al., 2010), y abre la puerta para explorar la capacidad explicativa de diferentes modelos no evaluados en este trabajo pero extensamente utilizados para modelar problemas de este tipo (Lee et al. 2009, 2011; Zeigenfuss y Lee, 2009; Zhang, Lee y Munro, 2009).



## Capítulo 8

# Referencias

- Ahn, W.-Y., Busemeyer, J. R., Wagenmakers, E.-J. y Stout, J. C. (2008). Comparison of decision learning models using the generalization criterion method. *Cognitive Science*, *32*, 1376-1402.
- Ahn, W.-Y., Krawitz, A., Kim, W., Busemeyer, J. R. y Brown, J. W. (2011). A model-based fMRI analysis with hierarchical Bayesian parameter estimation. *Journal of Neuroscience Psychology and Economics*, *4*, 95-110.
- Bechara, A., Damasio, A. R., Damasio, H. y Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*, 7-15.
- Busemeyer, J. R. y Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment: Decomposing performance on the Bechara gambling task. *Psychological Assessment*, *14*, 253-262.
- De Neys, W. y Verschueren, N. (2006). Working memory capacity and a notorious brain teaser: The case of the Monty Hall dilemma. *Experimental Psychology*, *53*, 123-131.
- Erev, I. y Roth, A. E. (1998). Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, *88*, 848-881.
- Fridberg, D. J., Queller, S., Ahn, W.-Y., Kim, W., Bishara, A. J. Busemeyer, J. R., Porrino, L. y Stout, J. C. (2010). Cognitive mechanisms underlying risky decision-making in chronic

- cannabis users. *Journal of Mathematical Psychology*, 54, 28-38.
- Friedman, D. (1998). Monty Hall's three doors: Construction and deconstruction of a choice anomaly. *The American Economic Review*, 88, 933-946.
- Gelman, A., y Rubin, D. B. (1992). Inference from interative simulation using multiple sequences. *Statistical Science*, 7, 457-472.
- Gershman, S. J., y Blei, D. M. (2012). A tutorial on Bayesian nonparametric models. *Journal of Mathematical Psychology*, 56, 1-12.
- Gnedin, A. V. (2012). The mondee gills game. *The Mathematical Intelligencer*, 34, 34-41.
- Granberg, D. (1999a). Cross-cultural comparison of responses to the Monty Hall dilemma. *Social Behavior and Personality*, 27, 431-437.
- Granberg, D. (1999b). A new version of the Monty Hall dilemma with unequal probabilities. *Behavioural Processes*, 48, 25-34.
- Granberg, D. y Brown, T. A. (1995). The Monty Hall dilemma. *Personality and Social Psychology Bulletin*, 21, 711-723.
- Granberg, D. y Dorr, N. (1998). Further exploration of two-stage decision making in the Monty Hall dilemma. *American Journal of Psychology*, 111, 561:579.
- Herbranson, W. T. y Schroeder, J. (2010). Are birds smarter than mathematicians? Pigeons (*Columbia livia*) perform optimally on a version of the Monty Hall dilemma. *Journal of Comparative Psychology*, 124, 1-13.
- Klein, E. D., Evans, T. A., Schultz, N. B. y Beran, M. J. (2013). Learning how to make a deal: Human (*Homo sapiens*) and monkey (*Macaca mulatta*) performance when repeatedly faced with the Monty Hall dilemma. *Journal of Comparative Psychology*, 127, 103-108.
- Lee, M. D. (2008). Three case studies in the Bayesian analysis of cognitive models. *Psychonomic Bulletin & Review*, 15, 1-15.
- Lee, M. D., y Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.

- Lee, M. D., Zhang, S., Munro, M. N. y Steyvers, M. (2009). Using heuristic decision models to understand human and optimal decision-making on bandit problems. En A. Howes, D. Peebles, y R. Cooper (Eds.), 9th International Conference on Cognitive Modeling - ICCM2009, Manchester, UK.
- Lee, M. D., Zhang, S., Munro, M. N. y Steyvers, M. (2011). Psychological models of human and optimal performance in bandit problems. *Cognitive Systems Research*, 12, 164-174.
- Luce, R. D. (1959). *Individual Choice Behavior*. New York: Wiley.
- Mazur, J. E. y Kahlbaugh, P. E. (2012). Choice behavior of pigeons (*Columbia livia*), college students, and preschool children (*Homo sapiens*) in the Monty Hall dilemma. *Journal of Comparative Psychology*, 126, 407-420.
- Palacios-Huerta, I. (2003). Learning to open Monty Hall's doors. *Experimental Economics*, 6, 235-251.
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, 20-22.
- Rosenhouse, J. (2009). *The Monty Hall Problem. The remarkable story of math's most contentious brain teaser*. Oxford University Press.
- Rouder, J. N., Morey, R. D. y Pratte, M. S. (en prensa). Bayesian hierarchical models. *New Handbook of Mathematical Psychology, Volume 1: Measurement and Methodology*.
- Selvin, S. (1975a). A problem in probability (letter to the editor). *The American Statistician*, 29, 67.
- Selvin, S. (1975b). On the Monty Hall problem (letter to the editor). *The American Statistician*, 29, 134.
- Shiffrin, R. M., Lee, M. D., Kim, W. y Wagenmakers, E.-J. (2008). A survey of model evaluation approaches with a tutorial on hierarchical Bayesian methods. *Cognitive Science*, 32, 1248-1284.



- Stagner, J. P. y Zentall, T. R. (2015). Further investigation of the Monty Hall dilemma in pigeons and rats. *Behavioural Processes*, *112*, 14-21.
- Steingroever, H., Wetzels, R. y Wagenmakers, E.-J. (2013a). A comparison of reinforcement learning models for the Iowa gambling task using parameter space partitioning. *The Journal of Problem Solving*, *5* (2), 1-32.
- Steingroever, H., Wetzels, R. y Wagenmakers, E.-J. (2013b). Validating the PVL-delta model for the Iowa gambling task. *Frontiers in Psychology*, *4*, Article 898.
- Steingroever, H., Wetzels, R. y Wagenmakers, E.-J. (2014). Absolute performance of reinforcement-learning models for the Iowa gambling task. *Decision*, *1*, 161-183.
- Tversky, A. y Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297-323.
- vos Savant, M. (1990a). "Ask Marilyn". *Parade Magazine*, September 9, p. 15.
- vos Savant, M. (1990b). "Ask Marilyn". *Parade Magazine*, December 2, p. 25.
- vos Savant, M. (1991). "Ask Marilyn". *Parade Magazine*, February 17, p. 12.
- vos Savant, M. (1996). *The Power of Logical Thinking*. St. Martin's Press.
- Wetzels, R., Vandekerckhove, J., Tuerlinckx, F. y Wagenmakers, E.-J. (2010). Bayesian parameter estimation in the expectancy valence model of the Iowa gambling task. *Journal of Mathematical Psychology*, *54*, 14-27.
- Zeigenfuse, M. D., y Lee, M. D. (2009). Bayesian nonparametric modeling of individual differences: A case study using decision-making on bandit problems. En N. Taatgen, H. van Rijn, J. Nerbonne, y L. Shonmaker (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, pp. 1412-1415. Austin, TX: Cognitive Science Society.
- Zhang, S., Lee, M. D. y Munro, M. N. (2009). Human and optimal exploration and exploitation in bandit problems. En A. Howes, D. Peebles y R. Cooper (Eds.), *9th International Conference on Cognitive Modeling - ICCM2009*, Manchester, UK.