Banco de México

Working Papers

N° 2024-03

Utilidad Esperada Descontada Aleatoria

Jose Apesteguia
ICREA, Universitat Pompeu Fabra and BSE

Miguel A. Ballester
University of Oxford

Ángelo Gutiérrez-Daza
Banco de México

Junio de 2024

La serie de Documentos de Investigación del Banco de México divulga resultados preliminares de trabajos de investigación económica realizados en el Banco de México con la finalidad de propiciar el intercambio y debate de ideas. El contenido de los Documentos de Investigación, así como las conclusiones que de ellos se derivan, son responsabilidad exclusiva de los autores y no reflejan necesariamente las del Banco de México.

The Working Papers series of Banco de México disseminates preliminary results of economic research conducted at Banco de México in order to promote the exchange and debate of ideas. The views and conclusions presented in the Working Papers are exclusively the responsibility of the authors and do not necessarily reflect those of Banco de México.

Documento de Investigación
2024-03

Working Paper
2024-03

Utilidad Esperada Descontada Aleatoria*

Jose Apesteguia†
ICREA, Universitat Pompeu Fabra and BSE

Miguel A. Ballester‡
University of Oxford

Ángelo Gutiérrez-Daza§
Banco de México

Resumen: Este artículo introduce el modelo de utilidad esperada descontada aleatoria (RDEU, por sus siglas en inglés), el cual se desarrolla para modelar y estimar preferencias de riesgo y tiempo heterogéneas. El modelo RDEU vincula de forma explícita la heterogeneidad entre las preferencias de los individuos y la heterogeneidad en sus decisiones. Se demuestra que el modelo posee propiedades sólidas de estática comparativa, se discute su identificación econométrica y se demuestra su conveniencia computacional. Finalmente, se utilizan dos bases de datos provenientes de diferentes estudios experimentales para ilustrar las ventajas del modelo RDEU frente a alternativas comúnmente utilizadas en la literatura a la hora de estimar la heterogeneidad en las preferencias entre individuos.
Palabras Clave: Heterogeneidad, Preferencias sobre Riesgo, Preferencias sobre Tiempo, Estática Comparativa, Modelos de Utilidad Aleatoria
JEL Classification: C01, D01

Abstract: This paper introduces the random discounted expected utility (RDEU) model, which we have developed as a means to deal with heterogeneous risk and time preferences. The RDEU model provides an explicit linkage between preference and choice heterogeneity. We prove it has solid comparative statics, discuss its identification, and demonstrate its computational convenience. Finally, we use two distinct experimental datasets to illustrate the advantages of the RDEU model over common alternatives for estimating heterogeneity in preferences across individuals.
Keywords: Heterogeneity, Risk Preferences, Time Preferences, Comparative Statics, Random Utility Models

*Este artículo circuló inicialmente como «Modelos de Utilidad Aleatoria para el Tratamiento Conjunto de Preferencias sobre Riesgo y Tiempo». Agradecemos a Daniel Benjamin, Holger Gerhardt, Jesse Shapiro y a un dictaminador anónimo por sus comentarios, y a Rafael Suchy por su ayuda en la investigación. Agradecemos también el apoyo del el Programa Severo Ochoa para Centros de Excelencia en I+D, la Beca Doctoral Severo Ochoa, el Ministerio de Ciencia e Innovación y el Balliol College.

† Department of Economics and Business. Correo: jose.apesteguia@upf.edu
‡ Department of Economics. Correo: miguel.ballester@economics.ox.ac.uk
§ Dirección General de Investigación Económica. Correo: angelo.gutierrez@banxico.org.mx

1. Introducción

Las situaciones económicas que implican simultáneamente riesgo y tiempo impregnan la mayoría de las esferas de la vida cotidiana, y la heterogeneidad en el comportamiento individual es la regla. En este artículo, se desarrolla un modelo para el tratamiento de preferencias heterogéneas por riesgo y tiempo, se establecen las probabilidades de elección predichas por el modelo, y se muestra que tiene una estática comparativa intuitiva. También se muestra que el modelo es fácil de implementar en la práctica, y que explica notablemente bien la heterogeneidad en las elecciones observadas en datos de dos de los diseños experimentales más influyentes en la literatura. Así, se proporciona un nuevo modelo que es computacionalmente eficiente y cuenta fundamentos teóricos sólidos, lo que lo hace idóneo para el análisis de preferencias heterogéneas por riesgo y tiempo.

Nuestro modelo es estocástico y se basa en una distribución de probabilidades sobre una determinada colección de funciones de utilidad. Esto nos permite establecer un vínculo directo entre la heterogeneidad no-observada en las preferencias y la heterogeneidad observada en las elecciones. En el trabajo, se adopta la familia de utilidades más estándar para el tratamiento del riesgo y el tiempo: las utilidades esperadas descontadas. De lo anterior, proviene el nombre de *modelo de utilidad esperada descontada aleatoria*, o RDEU, por sus siglas en inglés. El modelo se estudia bajo dos de los mecanismos de elicitación de preferencias por riesgo y tiempo en diseños experimentales más utilizados en la literatura: las listas dobles de precios múltiples (DMPL, por sus siglas en inglés) y los presupuestos convexos (CB, por sus siglas en inglés). El fuerte contraste entre ambos mecanismos, uno con elecciones binarias y otro con un espacio de elección continuo, nos permite mostrar la flexibilidad del modelo.

La adopción de modelos de utilidad aleatoria (RUM, por sus siglas en inglés) con múltiples parámetros de preferencia en aplicaciones empíricas ha sido lenta, entre otras razones, por su complejidad computacional. El cálculo de las probabilidades de elección en estos modelos implica la integración numérica de múltiples variables, lo que dificulta su análisis y estimación. En el caso de la utilidad esperada descontada, esto exige integrar la distribución conjunta de dos variables: el factor de descuento y la curvatura de la función de utilidad monetaria. En este trabajo se demuestra que la

anterior "maldición de la dimensión" no afecta al modelo RDEU: dada una curvatura de la función monetaria, siempre existe una estructura ordenada que vincula el descuento y las elecciones. Así, las probabilidades condicionales de elección para cualquier curvatura pueden calcularse directamente y luego agregarse con facilidad, haciendo que el modelo sea teórica y empíricamente conveniente.

Utilizando este enfoque, basado en la probabilidad condicional de observar una elección, se analizan las propiedades de estática comparativa estocástica del modelo RDEU. En particular, se analizan los cambios en la tendencia central y la dispersión de la distribución de probabilidad de los dos componentes principales de la utilidad esperada descontada: la curvatura de la función de utilidad y la tasa de descuento. Aunque el tratamiento teórico de la estática comparativa en modelos RUM con más de un parámetro es complicado, los resultados son coherentes con la comprensión común. En primer lugar, se observa que un cambio en la distribución de probabilidades hacia un mayor descuento sólo tiene efecto en problemas de tiempo, en los que desplaza la elección hacia opciones más tempranas. En segundo lugar, un cambio en la distribución de probabilidades hacia curvaturas mayores tiene un efecto en todos los tipos de problemas: genera cambios de elección (i) hacia opciones más seguras en listas de precios múltiples que implican riesgo, (ii) hacia opciones más tempranas en listas de precios múltiples que implican tiempo, y (iii) hacia perfiles de consumo más suaves en conjuntos convexos. Además, una mayor dispersión en cualquier variable de la distribución de probabilidad conduce a una mayor estocasticidad en la elección. Estos resultados son fundamentales para proporcionar a la literatura económica un marco bien fundamentado para la correcta interpretación y estimación de las variables de interés, es decir, el descuento y la curvatura.

Todos los resultados anteriores son para representaciones generales de utilidad esperada descontada y distribuciones de probabilidad no restringidas. En el trabajo se prosigue con una discusión de las implicaciones de estos resultados bajo parametrizaciones estándar. Siguiendo la práctica habitual en la literatura, se consideran funciones de utilidad monetaria de aversión relativa al riesgo constante (CRRA, por sus siglas en inglés), que están determinadas por un parámetro r que describe la curvatura.[1] Por lo tanto, cada utilidad se caracteriza por un par de parámetros

[1]La literatura contempla una variedad de formulaciones de esta función de utilidad, algunas de las cuales tienen implicaciones perversas cuando se introducen consideraciones de tiempo. La

(r, δ), donde δ captura la tasa de descuento. Para dar una mayor intuición de las propiedades del modelo, su identificación y su aplicación empírica, también se considera el caso especial estándar en el que r y δ siguen una distribución normal bivariada. Esta estructura paramétrica adicional acentúa aún más la conveniencia del modelo: el cálculo de la probabilidad de elección requiere de la evaluación de las distribuciones condicional y marginal de una normal bivariada, las cuales son a su vez distribuciones normales. Por lo anterior, el supuesto paramétrico reduce la dimensionalidad del problema numérico, haciendo rutinario el cálculo de las probabilidades de elección. Además, los cambios y las dispersiones de la distribución de probabilidad son el resultado de variaciones en los dos primeros momentos de la distribución, lo que facilita la identificación econométrica en el caso paramétrico.

Tras establecer los fundamentos teóricos del modelo, se ilustran sus ventajas empíricas con un ejercicio de estimación estructural utilizando datos de dos importantes ejemplos del tipo de mecanismos de elicitación considerados: Andersen, Harrison, Lau y Rutstrom (2008) (en adelante AHLR) y Andreoni y Sprenger (2012b) (en adelante AS). Se comparan las estimaciones agregadas y a nivel individual del modelo RDEU con las obtenidas utilizando las estrategias empíricas empleadas en los respectivos trabajos y en la literatura posterior.

La literatura ha utilizado a menudo RUM con choques i.i.d. aditivos para analizar diseños experimentales DMPL, donde los problemas de decisión son binarios, implican solo consideraciones de riesgo o tiempo, y una lotería con una única fecha define las alternativas. Teóricamente, la implementación de este enfoque con representaciones estándar de la utilidad esperada descontada puede conducir a predicciones paradójicas y propiedades de estática comparativa perversas (Wilcox, 2011; Apesteguia y Ballester, 2018), dificultando así una comprensión profunda de las preferencias de riesgo y tiempo. Empíricamente, se muestra que el modelo RDEU ofrece un mejor ajuste global a los datos que estas alternativas. También se comporta mejor que las recientes implementaciones de RUM i.i.d. aditivos utilizando la corrección de Wilcox (2011). Además, existen diferencias importantes entre los modelos a nivel individual: el modelo RDEU ofrece estimaciones razonables de las

sección 2.1 analiza las condiciones necesarias para el uso adecuado de una familia CRRA en un entorno que implica tanto riesgo como tiempo.

preferencias de riesgo y tiempo, que están altamente correlacionadas con las estimaciones comúnmente obtenidas a partir del cambio de decisión dentro de las tareas de riesgo y tiempo en los diseños DMPL.[2] Por el contrario, las estimaciones obtenidas a través de modelos RUM i.i.d. aditivos están débilmente correlacionadas con las estimaciones semiparamétricas (SPE, por sus siglas en inglés) y toman valores inverosímiles para un subconjunto específico de individuos que se identifican en el trabajo.

Por otra parte, en los diseños experimentales CB, los menús son continuos, implican consideraciones de riesgo y tiempo, y cada alternativa concede un par de loterías fechadas. Para este diseño experimental, la literatura ha utilizado mínimos cuadrados no lineales, asumiendo una utilidad esperada descontada única, para estimar las preferencias de riesgo y tiempo en un conjunto de datos. Para ello, se suele introducir aleatoriedad perturbando la condición de primer orden de un problema de maximización de la utilidad. La aleatoriedad introducida en este enfoque carece de fundamento teórico y económico, ya que no relaciona explícitamente la heterogeneidad de la elección con la heterogeneidad de las preferencias. Además, este enfoque no es adecuado para comprender la gran heterogeneidad de las elecciones observada en los datos. Otro enfoque en el análisis de los conjuntos de datos CB utiliza, como en el caso de los DMPL, modelos RUM con choques i.i.d. aditivos. Empíricamente, esta extensión multinomial tiende a estimar funciones de utilidad convexas para poder explicar la gran proporción de soluciones de esquina observada en los datos experimentales CB, a costa de no poder explicar la gran fracción de elecciones observadas en el rango intermedio de los conjuntos de presupuestos. En contraste con estos enfoques, el modelo RDEU explica empíricamente la prevalencia observada de las elecciones de esquina e interiores, a la vez que proporciona estimaciones plausibles del descuento y la curvatura de la función de utilidad.

En conclusión, los resultados teóricos y las aplicaciones empíricas en este trabajo ilustran la utilidad del modelo RDEU como un marco robusto y unificador para estimar las preferencias de tiempo y de riesgo con datos experimentales, teniendo en cuenta la gran heterogeneidad de las elecciones entre individuos. Este ejercicio está

[2]Véase, por ejemplo, Holt y Laury (2002).

relacionado con la literatura metodológica reciente sobre estimaciones de preferencias en una variedad de entornos (véase, por ejemplo, DellaVigna, 2018; Cattaneo et al., 2020; Dardanoni et al., 2020; Aguiar y Kashaev, 2021; Barseghyan et al., 2021). Este trabajo se diferencia de la literatura citada en que se centra en las preferencias de riesgo y tiempo y establece las propiedades de estática comparativa del modelo.

2. Utilidad Esperada Descontada Aleatoria

Una lotería es una colección finita de premios/pagos monetarios y las probabilidades asociadas a estos. Es decir, es un vector de la forma $l = [p_1, \ldots, p_N; x_1, \ldots x_N]$, con $p_n \geq 0$, $\sum_{n=1}^{N} p_n = 1$ y $x_n \geq 0$. Una lotería fechada (l, t) está formada por una lotería y un momento en el tiempo $t \geq 0$, en el que se concede el premio resultante.[3]

La utilidad esperada descontada (DEU, por sus siglas en inglés) es el modelo determinista de comportamiento más utilizado para estudiar las preferencias por riesgo y tiempo. Este estudio considera una familia $\{u_r\}_{r\in\mathbb{R}}$ de funciones de utilidad continuas y estrictamente crecientes sobre el dinero, que se normalizan para satisfacer $u_r(\omega) = 0$ en un nivel de riqueza de referencia $\omega > 0$. Se imponen tres supuestos básicos sobre la familia de utilidades monetarias. En primer lugar, debe incluir la utilidad monetaria lineal que se denota como $r = 0$. En segundo lugar, la familia está estrictamente ordenada por concavidad, es decir, $r < r'$ significa que $u_{r'}$ es estrictamente más cóncava que u_r.[4] En tercer lugar, la convexidad y la concavidad son ilimitadas cuando r tiende a $-\infty$ y $+\infty$, respectivamente. Muchas familias satisfacen estos requisitos básicos, incluida la ampliamente utilizada CRRA o la familia de utilidades con aversión al riesgo absoluto constante (CARA, por sus siglas en inglés). El factor de descuento del individuo se denota por $e^{-\delta}$ con $\delta \in \mathbb{R}$.[5]. Dados los parámetros $(r, \delta) \in \mathbb{R}^2$, la evaluación DEU de una secuencia de loterías

[3]Cuando sea necesario, los premios/pagos y el número de resultados de las loterías fechadas $j = 1, \ldots, J$ se denotarán como x_n^j, p_n^j y N_j. Se asume también, como es habitual en la literatura, que los premios monetarios adjudicados se consumen al momento de ser recibidos.

[4]Formalmente, u es estrictamente más cóncava que u' si existe una función g creciente y estrictamente cóncava tal que $u'(x) = g(u(x))$ para cada x. Como resultado, las utilidades con $r > 0$ (respectivamente, $r < 0$) representan la aversión al riesgo (respectivamente, amor al riesgo).

[5]Por lo tanto, $\delta > 0$ (respectivamente, $\delta < 0$) representa la tasa de impaciencia o aversión al retraso (respectivamente, amor al retraso). El factor de descuento se escribe de esta forma por comodidad, pues permite utilizar una normal bivariada simple en la estimación paramétrica. Obsérvese que, como alternativa, se podría escribir simplemente $\frac{1}{1+d} = e^{-\delta} \in \mathbb{R}_{++}$, con d representando el factor de descuento en los reales positivos.

fechadas $(l^1, t^1), \ldots, (l^J, t^J)$ es:

$$DEU_{r,\delta}((l^1, t^1), \ldots, (l^J, t^J)) = \sum_{j=1}^{J} e^{-\delta t^j} \sum_{n=1}^{N_j} p_n^j u_r(\omega + x_n^j). \tag{2.1}$$

Ahora se puede definir el modelo estocástico que se analiza en este trabajo, denominado como el *modelo de utilidad esperada descontada aleatoria* (RDEU, por sus siglas en inglés). Sea f una densidad medible con soporte completo sobre $\mathbb{R}^2$, la cual captura la prevalencia de cada posible preferencia DEU. En el momento de realizar una elección en un problema de decisión, se genera una realización de los parámetros (r, δ), la cual tiene probabilidad $f(r, \delta)$, y se selecciona la alternativa que maximiza la $DEU_{r,\delta}$ en el problema de decisión.[6]

A continuación se describen formalmente los problemas de decisión asociados a los dos diseños experimentales analizados en este trabajo. Uno de los diseños más destacados en la literatura experimental implica el uso de las llamadas listas de precios múltiples dobles (DMPL), como en Andersen et al. (2008), donde los problemas de decisión son binarios, implican sólo consideraciones de riesgo o de tiempo, y las alternativas se definen mediante una única lotería fechada.[7] En un problema de decisión bajo riesgo, cada una de las dos alternativas corresponde a una sola lotería contingente de dos estados con premios concedidos en el presente. Es decir, dado $x_1^1 > x_1^0 > x_2^0 > x_2^1$ y $p \in (0, 1)$, el menú de riesgo asociado es $A_{\mathcal{R}} = \{0_{\mathcal{R}}, 1_{\mathcal{R}}\}$ donde $0_{\mathcal{R}} = ([p, 1-p; x_1^0, x_2^0], 0)$ y $1_{\mathcal{R}} = ([p, 1-p; x_1^1, x_2^1], 0)$.[8] En un problema de decisión temporal, cada una de las dos alternativas está compuesta por una única lotería fechada degenerada. Es decir, dado $t^0 < t^1$ y $x^0 < x^1$, el menú temporal asociado es $A_{\mathcal{T}} = \{0_{\mathcal{T}}, 1_{\mathcal{T}}\}$ donde $0_{\mathcal{T}} = ([1; x^0], t^0)$ y $1_{\mathcal{T}} = ([1; x^1], t^1)$. Dada la naturaleza binaria de los menús de riesgo y tiempo, las probabilidades de elección del modelo RDEU en cualquier problema de decisión están determinadas por la probabilidad de elección de una de las dos alternativas del menú, digamos $0_{\mathcal{R}}$ en

[6]Dado que se supone que f es medible, las indiferencias entre alternativas máximas no son importantes y se obviarán en el documento.

[7]Véase también Burks et al. (2009), Dohmen et al. (2010), Tanaka et al. (2010), Benjamin et al. (2013), Falk et al. (2018) o Jagelka (2021). En el Apéndice B se estudia una versión híbrida en la que actúan simultáneamente consideraciones de riesgo y de tiempo. (véase, Ahlbrecht y Weber (1997), Coble y Lusk (2010), Baucells y Heukamp (2012) y Cheung (2015)).

[8]A veces, se considera $p \in \{0, 1\}$. Estos casos son triviales, ya que una de las dos loterías está dominada y, por lo tanto, el modelo RDEU predice una probabilidad de elección igual a cero.

un menú de riesgo y $0_{\mathcal{T}}$ en un menú de tiempo. Sea $\Gamma(0_{\mathcal{R}}, A_{\mathcal{R}}) \subseteq \mathbb{R}^2$ (respectivamente, $\Gamma(0_{\mathcal{T}}, A_{\mathcal{T}}) \subseteq \mathbb{R}^2$) la colección de todas las combinaciones de parámetros (r, δ) para las que el maximizador de $DEU_{r,\delta}$ dentro del menú $A_{\mathcal{R}}$ (respectivamente $A_{\mathcal{T}}$) es $0_{\mathcal{R}}$ (respectivamente $0_{\mathcal{T}}$). Las f-medidas de estos conjuntos describen las probabilidades de elección:

$$\mathcal{P}_f(0_{\mathcal{R}}, A_{\mathcal{R}}) = \int_{\Gamma(0_{\mathcal{R}}, A_{\mathcal{R}})} f(r, \delta) \mathrm{d}(r, \delta),$$

$$\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) = \int_{\Gamma(0_{\mathcal{T}}, A_{\mathcal{T}})} f(r, \delta) \mathrm{d}(r, \delta).$$

En un diseño alternativo, propuesto Andreoni y Sprenger (2012a,b), los sujetos se enfrentan a los denominados menús convexos.[9] Estos menús son continuos, implican consideraciones de riesgo y tiempo, y cada alternativa en ellos concede un par de loterías fechadas. Formalmente, dados $x^0 \leq x^1$, $t^0 < t^1$ y $p^0, p^1 \in (0, 1]$, el menú convexo asociado es $A_{\mathcal{C}} = [0, 1]$ donde la alternativa $a \in A_{\mathcal{C}}$ se define por la secuencia de dos loterías fechadas $([p^0, 1 - p^0; (1 - a)x^0, 0], t^0), ([p^1, 1 - p^1; ax^1, 0], t^1)$. Dada la naturaleza continua de los menús convexos, las probabilidades de elección del modelo RDEU están determinadas por la probabilidad de elección acumulada de seleccionar alternativas por debajo de cualquier valor $a \in [0, 1]$. Sea $\Gamma([0, a], A_{\mathcal{C}}) \subseteq \mathbb{R}^2$ la colección de todas las combinaciones de parámetros (r, δ) para las que el maximizador de $DEU_{r,\delta}$ dentro del menú $A_{\mathcal{C}}$ es una alternativa por debajo de a. Las f-medidas de estos conjuntos describen las probabilidades de elección:

$$\mathcal{P}_f([0, a], A_{\mathcal{C}}) = \int_{\Gamma([0,a], A_{\mathcal{C}})} f(r, \delta) \mathrm{d}(r, \delta).$$

2.1. **Versión Paramétrica.** Todos los resultados teóricos a continuación aplican a representaciones generales de utilidad esperada descontada y distribuciones de probabilidad sin restricciones adicionales a las discutidas hasta ahora. Sin embargo, añadir restricciones paramétricas adicionales será útil para ilustrar las principales intuiciones de los resultados teórico, así como para su el uso del modelo RDEU como herramienta práctica en un ejercicio de estimación. Por lo tanto, se ilustrará

[9] Los menús convexos se utilizan ampliamente para el estudio de una variedad de preferencias económicas. Véase, por ejemplo, Choi et al. (2007), Fisman et al. (2007), Augenblick et al. (2015), Carvalho et al. (2016), Alan y Ertac (2018) y Kim et al. (2018).

la intuición de cada resultado teórico utilizando funciones de utilidad monetaria CRRA.

Dado el parámetro $r \in \mathbb{R}$, la utilidad CRRA de un premio adicional $x \geq 0$ es:

$$u_r^{crra}(\omega + x) = \begin{cases} \frac{(\omega+x)^{1-r} - \omega^{1-r}}{1-r} & \text{siempre que } r \neq 1; \\ \log(\omega + x) - \log \omega & \text{en caso contrario.} \end{cases}$$

Comentamos brevemente las razones detrás de la parametrización elegida para la familia CRRA, ya que la literatura contempla muchas formulaciones diferentes y no todas son apropiadas cuando intervienen tanto el riesgo como el tiempo. En primer lugar, al igual que en el caso del estudio de las preferencias por riesgo, se puede suponer que el parámetro r pertenece a $\mathbb{R}$, pero para ello es necesario el supuesto de riqueza base $\omega > 0$. De lo contrario, las utilidades monetarias $\{u_r^{crra}\}_{r \geq 1}$ no estarían bien definidas para pagos nulos.

En segundo lugar, se tiene que garantizar que todas las utilidades monetarias son estrictamente crecientes y, por lo tanto, la potencia $(\omega + x)^{1-r}$ debe volver a escalarse con la constante $\frac{1}{1-r}$. En tercer lugar, como este escalamiento crea utilidades negativas siempre que $r > 1$, la adición de la constante $-\frac{\omega^{1-r}}{1-r}$ garantiza que las utilidades sean siempre positivas y hace que $u_r^{crra}(\omega) = 0$, lo que facilita el análisis de loterías con premios nulos e implica la propiedad de continuidad estándar $\lim_{r \to 1} u_r^{crra}(\omega + x) = u_1^{crra}(\omega + x)$.[10]

Cuando la dicusión se enfoque en el caso paramétrico, también se impondrán algunas restricciones a la distribución de probabilidad f. Los métodos computacionales discutidos en el Apéndice C permiten la estimación eficiente del modelo para cualquier distribución caracterizada por un vector finito de parámetros $\theta \in \Theta$.[11] En el análisis principal de este trabajo, nos enfocamos en el caso en que (r, δ) sigue una distribución normal bivariada, de modo que $\theta \equiv (\mu_r, \sigma_r, \mu_\delta, \sigma_\delta, \rho)$, donde μ_z y σ_z son la media y la desviación típica correspondientes del parámetro $z \in \{r, \delta\}$, y ρ es el coeficiente de correlación entre r y δ. Este supuesto proporciona una referencia natural para comparar el modelo y los resultados empíricos de las siguientes secciones con otros modelos de la literatura. Como se muestra más adelante, también permite

[10] El Apéndice E incluye una discusión sobre el papel de la riqueza ω con utilidades CRRA.
[11] El Apéndice D presenta un ejemplo con distribuciones normal y beta truncada.

obtener expresiones sencillas para las probabilidades de elección en el modelo, mismas que se explotarán para proporcionar condiciones que permitan la identificación de θ en cada escenario experimental.

3. Listas Dobles De Precios Múltiples: Teoría

3.1. **Menús de Riesgo.** Dado que en estos problemas de decisión toda la acción tiene lugar en el presente, el parámetro de descuento δ no desempeña ningún papel. Además, el tipo de loterías en juego siempre crea una estructura intuitiva y ordenada de elecciones para el parámetro r. A continuación, se demuestra que, para cada menú de riesgo $A_{\mathcal{R}}$, existe una constante real $K(A_{\mathcal{R}})$ tal que la alternativa $0_{\mathcal{R}}$ se elige si y sólo si $r \geq K(A_{\mathcal{R}})$. Por lo tanto, la probabilidad de elección de la alternativa $0_{\mathcal{R}}$ es la f-medida del conjunto rectangular $\Gamma(0_{\mathcal{R}}, A_{\mathcal{R}}) = \{(r, \delta) : r \geq K(A_{\mathcal{R}})\}$, que puede calcularse convenientemente utilizando la FDA marginal de r, denotada como F^r.

Además, las propiedades de estática comparativa relacionadas a los cambios en la tendencia central y la dispersión del parámetro r son intuitivas y se deducen inmediatamente a partir de la estructura impuesta: cuando la masa de la distribución marginal de r se desplaza hacia valores mayores, se garantiza que la probabilidad de elección de la alternativa más segura aumenta de forma estricta. Cuando la masa de la distribución marginal de r se aleja de su mediana, la probabilidad de elección de ambas alternativas se aproxima estrictamente a la mitad, es decir, el comportamiento se vuelve estrictamente más estocástico.[12]

Para formalizar estas ideas, necesitamos definir las nociones estándar de dominio y expansión utilizando FDAs. Formalmente, sean F y G dos FDAs sobre la variable aleatoria z, con dominio en un intervalo abierto. Denote la mediana de la distribución F como $\text{med}(F)$. Entonces, decimos que: (i) F domina a G si $F(z) < G(z)$ se cumple para todos los valores de z y (ii) F expande a G si $\text{med}(F) = \text{med}(G)$, $F(z) > G(z)$ siempre que $z < \text{med}(F)$ y $F(z) < G(z)$ siempre que $z > \text{med}(F)$.[13]

[12]Existe una excepción obvia a este principio cuando la estocasticidad de elección ya es máxima, eligiéndose ambas alternativas con la misma probabilidad 1/2. Esto ocurre cuando la mediana de F^r coincide con el umbral separador $K(A_{\mathcal{R}})$.

[13]La prueba de cada proposición se encuentra disponible en el Apéndice A.

Proposición 1. *Para cada par de RDEUs, f y g, y cada menú $A_{\mathcal{R}}$:*

(1) $\mathcal{P}_f(0_{\mathcal{R}}, A_{\mathcal{R}}) = 1 - F^r(K(A_{\mathcal{R}}))$.
(2) *Si F^r domina a G^r, entonces* $\mathcal{P}_f(0_{\mathcal{R}}, A_{\mathcal{R}}) > \mathcal{P}_g(0_{\mathcal{R}}, A_{\mathcal{R}})$.
(3) *Si F^r expande a G^r con $K(A_{\mathcal{R}}) \neq med(F^r)$, entonces* $|\mathcal{P}_f(0_{\mathcal{R}}, A_{\mathcal{R}}) - \frac{1}{2}| < |\mathcal{P}_g(0_{\mathcal{R}}, A_{\mathcal{R}}) - \frac{1}{2}|$.

3.2. **Menús de Tiempo.** Cuando el tiempo está en juego, comprender el comportamiento del modelo es algo más complicado porque el parámetro de descuento δ, por sí solo, aporta información incompleta sobre las preferencias de los individuos.

Ejemplo 1. *Sean dos individuos DEU-CRRA con $\omega = 100$ y parámetros de preferencia $(r_1, \delta_1) = (0.95, 0.094)$ y $(r_2, \delta_2) = (0, 0.105)$. Aunque $\delta_1 < \delta_2$ (lo que es equivalente a $e^{-\delta_1} = 0.91 > 0.9 = e^{-\delta_2}$) puede sugerir que el individuo 1 es más paciente, se puede ver de inmediato que es efectivamente el único que prefiere* $([1; 71.5], 0)$ *a* $([1; 80], 1)$.

La consideración conjunta de ambos parámetros es entonces necesaria para comprender plenamente las predicciones de la DEU y, en consecuencia, las probabilidades de elección de la RDEU requieren, en principio, la evaluación de una integral doble. Afortunadamente, el análisis vuelve a ser sencillo tras condicionar el parámetro r, ya que esto siempre genera una estructura intuitiva y ordenada de elecciones sobre el parámetro de descuento δ. A continaución se demuestra que, para cualquier menú temporal $A_{\mathcal{T}}$ y cualquier valor de r, existe una constante dependiente del menú $K(A_{\mathcal{T}}|r) \in \mathbb{R}_+$ tal que la alternativa anterior $0_{\mathcal{T}}$ se selecciona si y sólo si $\delta \geq K(A_{\mathcal{T}}|r)$, es decir, $\Gamma(0_{\mathcal{T}}, A_{\mathcal{T}}) = \{(r, \delta) : \delta \geq K(A_{\mathcal{T}}|r)\}$. Como resultado, la probabilidad de elección de la alternativa $0_{\mathcal{T}}$ puede expresarse convenientemente mediante las probabilidades de elección condicionales al parámetro r. En resumen, se pueden evaluar las FDA condicionales del parámetro δ en el parámetro r, denotadas como $F_{\delta|r}$, en el umbral correspondiente $K(A_{\mathcal{T}}|r)$, para luego agregar los valores de r utilizando su densidad marginal, la cual se denota como f^r. La Proposición 2 se basa en esta estructura ordenada, mostrando que los umbrales $\{K(A_{\mathcal{T}}|r)\}_{r\in\mathbb{R}}$ son estrictamente decrecientes en r y constituyen una biyección de $\mathbb{R}$ a $\mathbb{R}_{++}$, por

lo que puede invertirse.[14] De este modo, la estática comparativa de los cambios es inmediata, ya que si se mantiene constante la distribución marginal de r (o δ) y se desplazan hacia arriba las distribuciones condicionales de δ (o r), se garantiza un aumento de la probabilidad de elección de la alternativa anterior $0_{\mathcal{T}}$. En segundo lugar, con respecto a las dispersiones, también se puede demostrar que mantener constante la distribución marginal de un parámetro, una expansión de las distribuciones condicionales del otro siempre crea más estocasticidad.[15]

Proposición 2. *Para cada par de RDEUs, f y g, y cada menú $A_{\mathcal{T}}$:*

(1) $\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) = 1 - \int_r F_{\delta|r} K(A_{\mathcal{T}}|r) f^r(r) \mathrm{d}r = 1 - \int_{\delta>0} F_{r|\delta} K(A_{\mathcal{T}}|\delta) f^{\delta}(\delta) \mathrm{d}\delta$.

(2) (a) *Si $F^r = G^r$, y para todo r $F_{\delta|r}$ domina a $G_{\delta|r}$, entonces $\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) \geq \mathcal{P}_g(0_{\mathcal{T}}, A_{\mathcal{T}})$.*

(b) *Si $F^{\delta} = G^{\delta}$, y para todo δ $F_{r|\delta}$ domina a $G_{r|\delta}$, entonces $\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) \geq \mathcal{P}_g(0_{\mathcal{T}}, A_{\mathcal{T}})$.*

(3) (a) *Si $F^r = G^r$, y para todo r $F_{\delta|r}$ expande a $G_{\delta|r}$ con $K(A_{\mathcal{T}}|r) \neq med(F_{\delta|r})$, entonces*
$|\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) - \frac{1}{2}| < |\mathcal{P}_g(0_{\mathcal{T}}, A_{\mathcal{T}}) - \frac{1}{2}|$.

(b) *Si $F^{\delta} = G^{\delta}$, y para todo δ $F_{r|\delta}$ expande a $G_{r|\delta}$ con $K(A_{\mathcal{T}}|\delta) \neq med(F_{r|\delta})$, entonces*
$|\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) - \frac{1}{2}| < |\mathcal{P}_g(0_{\mathcal{T}}, A_{\mathcal{T}}) - \frac{1}{2}|$.

3.3. **Implicaciones para la Versión Paramétrica.** Los resultados generales de las Proposiciones 1 y 2 tienen las siguientes implicaciones cuando se restringe el análisis al caso paramétrico con utilidades CRRA y una distribución normal bivariada. En el caso de CRRA, los umbrales descritos en la Proposición 1 corresponden al único valor de r que resuelve la ecuación $\frac{1-p}{p} = \frac{(\omega+x_1^1)^{1-r} - (\omega+x_1^0)^{1-r}}{(\omega+x_2^0)^{1-r} - (\omega+x_2^1)^{1-r}}$. En la normal bivariante, la distribución marginal del parámetro r sigue una distribución normal con parámetros μ_r y σ_r. La parte 1 de la proposición afirma que el análisis de las probabilidades de elección en el modelo RDEU se simplifica a un ejercicio computacionalmente sencillo. Además, los desplazamientos y las expansiones de F^r son el

[14] En otras palabras, condicionar en δ también resulta en elecciones ordenadas sobre el parámetro r. Siempre que $\delta \leq 0$, se elige $1_{\mathcal{T}}$. Siempre que $\delta > 0$, existe una constante dependiente del menú $K(A_{\mathcal{T}}|\delta) \in \mathbb{R}$ tal que $0_{\mathcal{T}}$ se elige si y sólo si $r \geq K(A_{\mathcal{T}}|\delta)$.

[15] Como en el caso del riesgo, la mediana de cada distribución condicional $F_{\delta|r}$ debe ser diferente del umbral correspondiente $K(A_{\mathcal{T}}|r)$ cuando se consideran expansiones de δ, con una expresión análoga para el caso de r.

resultado de un aumento de μ_r y σ_r, respectivamente. Por lo tanto, las partes 2 y 3 muestran que la estática comparativa del modelo tiene una intuición sencilla: un aumento de la mediana del parámetro r crea siempre una mayor probabilidad de elección de la alternativa más segura, mientras que un aumento de la varianza del parámetro r genera más estocasticidad de elección.

Del mismo modo, se puede leer la Proposición 2 desde el punto de vista paramétrico. Con CRRA, el umbral puede escribirse como $K(A_{\mathcal{T}}|r) = \frac{1}{t^1-t^0} \log \left[\frac{(\omega+x^1)^{1-r}-\omega^{1-r}}{(\omega+x^0)^{1-r}-\omega^{1-r}} \right]$. Con la normal bivariada, todos las condicionales $F_{\delta|r}$ también son normales, con media $\mu_\delta + \frac{\sigma_\delta}{\sigma_r}\rho(r-\mu_r)$ y desviación estándar $\sqrt{1-\rho^2} \cdot \sigma_\delta$. Esto, combinado con la normalidad ya mencionada de f^r, hace que el cálculo de las probabilidades de elección sea un ejercicio sencillo. Además, considerando $z, z' \in \{r, \delta\}$ con $z \neq z'$, un aumento de μ_z no afecta al marginal $F^{z'}$, al tiempo que genera un cambio dominante en todos los condicionales $F_{z|z'}$. Por lo tanto, la parte 2 afirma que, al aumentar la media de δ o la media de r, se genera una mayor probabilidad de elección de la alternativa temprana. En tercer lugar, el aumento de σ_z no afecta al $F^{z'}$ marginal y, bajo la corrección adecuada de la covarianza, genera la expansión de todos los condicionales $F_{z|z'}$. Por lo tanto, la parte 3 afirma que un aumento de la varianza de r o de δ, con la corrección adecuada de la covarianza, producirá más estocasticidad de elección.

Las Proposiciones 1 y 2 sientan las bases para la identificación no paramétrica del modelo. Ahora se analizará la identificación de los parámetros θ, bajo el supuesto de que (r, δ) sigue una distribución normal bivariada. Considere un conjunto de datos DMPL $\mathcal{O}$ compuesto por un conjunto de observaciones de la elección de un sujeto, o grupo de sujetos, cuando se le presenta un menú de riesgo y/o de tiempo. Supongamos que el conjunto de datos tiene M de estos menús, denominados A_m para $m = 1, \ldots, M$. La siguiente proposición muestra que dos menús de riesgo y tres menús de tiempo, con propiedades comúnmente encontradas en la mayoría de datos experimentales existentes, son suficientes para identificar θ.

Proposición 3. *Suponga que el conjunto de datos $\mathcal{O}$ contiene:*

(a) *Dos menús de riesgo $\{A_{\mathcal{R},a}, A_{\mathcal{R},b}\}$ tales que $K(A_{\mathcal{R},a}) \neq K(A_{\mathcal{R},b})$.*

(b) *Tres menús de tiempo* $\{A_{\mathcal{T},c}, A_{\mathcal{T},d}, A_{\mathcal{T},e}\}$ *que difieren en una de las siguientes tres dimensiones:*

(i) el retraso $t_m^1 - t_m^0$*, (ii) el pago presente* x_m^0*, (iii) el pago futuro* x_m^1*.*

Entonces, θ *está identificado.*

Intuitivamente, la prueba de la Proposición 3 muestra que se puede utilizar la variación en los umbrales de indiferencia $K(A_{\mathcal{R}})$ entre menús de riesgo para identificar los parámetros (μ_r, σ_r) que caracterizan la distribución marginal de r. Condicionando a (μ_r, σ_r), se puede utilizar la variación en el retraso entre menús de tiempo que ofrecen un mismo premio para recuperar $(\mu_\delta, \sigma_\delta, \rho)$. Alternativamente, se puede utilizar la variación en la tasa de retorno implícita entre los menús de tiempo (es decir, la variación en (x^0, x^1)) para sustituir la variación en los retrasos.

Bajo condiciones de regularidad estándar, la identificación de θ implica la consistencia de los estimadores de máxima verosimilitud de este vector de parámetros. Esta propiedad garantiza que un analista puede recuperar el valor poblacional de θ con una muestra de observaciones suficientemente grande. Sin embargo, el comportamiento de estos estimadores con muestras experimentales pequeñas puede generar preocupación. El siguiente resultado muestra que los parámetros verdaderos pueden inferirse con tan solo cinco menús, mitigando así esta preocupación.

Proposición 4. *Para cualquier* $\theta \in \Theta$*, existen cinco menús DMPL que permiten inferir su valor exactamente.*

En contextos experimentales, los menús de riesgo y tiempo suelen adaptarse para incluir variaciones que permitan a los investigadores inferir un conjunto de valores de aversión al riesgo y descuento bajo el supuesto de que los sujetos tienen preferencias deterministas y maximizan su utilidad esperada descontada. Como se explica en la siguiente sección, los investigadores utilizan los cambios en las elecciones entre menús con diferentes umbrales de indiferencia para estimar un intervalo que contenga el valor puntual del coeficiente de aversión al riesgo de un sujeto (suponiendo una función de utilidad CRRA) o su tasa de descuento, dado un valor del coeficiente de aversión al riesgo. La Proposición 3 muestra que el mismo tipo de variación permite a los investigadores inferir los parámetros de la representación RDEU bajo supuestos paramétricos. Más aún, la Proposición 4 muestra que los investigadores

también pueden adaptar los menús DMPL de un diseño experimental para maximizar su capacidad de inferir un conjunto de valores de parámetros. Concluimos esta discusión usando un ejemplo.[16]

Ejemplo 2. *Sea* $\omega \to 0$ *y* $(\mu_r, \sigma_r, \mu_\delta, \sigma_\delta, \rho) = (0.7, 0.7, 0.05, 0.02, -0.5)$. *Consideremos primero los menús de riesgo. La probabilidad de elección de* $([0.5, 0.5; 50, 40], 0)$ *frente a* $([0.5, 0.5; 68, 25], 0)$ *es aproximadamente* $1 - \Phi(0) = 0.5$ *y, dado que el umbral de este problema es* 0.7*, tenemos* $\mu_r = 0.7$. *La probabilidad de elección de* $([0.5, 0.5; 50, 40], 0)$ *frente a* $([0.5, 0.5; 95, 25], 0)$ *es* $1 - \Phi(1) = 0.16$ *y, dado que el umbral de este problema es* 1.4*, se tiene* $\sigma_r = 1.4 - 0.7 = 0.7$. *Considérese ahora los menús de tiempo. Como se argumenta en la prueba de la Proposición 4, cuando* $\omega \to 0$*, el mapa umbral se convierte en el mapa lineal por tramos* $\min\{0, K(A_\mathcal{T})(1-r)\}$*, el cual puede aproximarse mediante el mapa lineal* $K(A_\mathcal{T})(1-r)$ *que pasa por el punto* $(1, 0)$ *y tiene pendiente* $-K(A_\mathcal{T})$. *Luego, considérese la probabilidad de elección de* $([1; 59], 0)$ *frente a* $([1; 70], 1)$*, que es aproximadamente* 0.5. *Dado que la constante de este menú es aproximadamente* $\frac{0.5}{3}$*, se tiene* $\mu_\delta = \frac{0.5}{3}(1 - 0.7) = 0.05$. *Ahora, considérese la probabilidad de elección de* $([1; 70 - \epsilon], 0)$*, con* ϵ *pequeño, frente a* $([1; 70], 1)$*, que es igual a* 0.99. *Esto corresponde a dos y media desviaciones estándar de la distribución normal, y dado que la constante en este caso es* 0*, se deduce que* $\frac{\mu_\delta}{\sigma_\delta} = \Phi^{-1}(0.99) = 2.5$*, y por lo tanto* $\sigma_\delta = 0.02$. *Finalmente, el menú de tiempo que involucra* $([1; 68], 0)$ *y* $([1; 70], 1)$ *tiene una constante de* 0.029*, que es igual al cociente de desviaciones estándar. Por lo tanto, dado que la probabilidad de elección de la opción más temprana es aproximadamente* 0.98 *(lo que corresponde a dos desviaciones estándar de la normal), debe cumplirse que* $\rho = \frac{1}{2}\left[\frac{\frac{-0.3}{0.7}+2.5}{2}\right]^2 - 1$*, lo que resulta aproximadamente en* $\rho = -0.5$.

4. Listas De Precios Múltiples Dobles: Ilustración Empírica

Esta sección ilustra la aplicación empírica del modelo RDEU a los conjuntos de datos DMPL. A contunación, se estima la versión paramétrica del modelo y se comparan sus resultados con los obtenidos empleando otros modelos estructurales

[16]Siguiendo la convención estándar, dejamos que $\Phi(\cdot)$ y $\phi(\cdot)$ denoten la FDA y la FPD de la distribución normal estándar.

utilizados previamente en la literatura. Para ello, se utilizan datos de AHLR. En este estudio experimental, los autores presentaron cuatro tareas de riesgo a una muestra de 253 individuos representativa de la población danesa. Cada tarea constaba de hasta diez menús de riesgo. Los premios monetarios de las alternativas segura y arriesgada, $(x_1^0, x_2^0, x_1^1, x_2^1)$, variaban entre cada tarea. Todos los menús compartían los mismos premios dentro de una tarea determinada, pero diferían en las probabilidades de recompensa (p^0, p^1). Un ejemplo de este tipo de tarea se muestra en el cuadro 7. El experimento también presentaba a cada individuo seis tareas de tiempo, cada una compuesta de hasta 10 menús de tiempo que compartían el mismo premio adelantado x^0. Todos los menús de una tarea temporal dada compartían también el mismo retraso de pago k, pero variaban en el valor del premio retrasado x^1. El retraso k y las fechas de pago (t^0, t^1) cambiaban entre tareas. El cuadro 8 muestra un ejemplo de este tipo de tarea. Siguiendo a estos autores, también se supone que el valor de la riqueza diaria media integrada ω es común a todos los individuos e igual a 118 coronas danesas (DKK) en 2003, lo que equivale aproximadamente a 30 USD al tipo de cambio de ese año.

Nuestro objetivo es estimar la aversión al riesgo y la tasa de descuento tanto a nivel poblacional como individual. Por este motivo, se restringe el análisis a una submuestra del conjunto de datos original que cumple las siguientes restricciones: en primer lugar, se descartan todas las observaciones correspondientes a cuatro menús de riesgo y seis menús de tiempo que contienen loterías dominadas.[17] En segundo lugar, se eliminó de la muestra a los individuos que muestran indiferencia entre las dos alternativas en algunas tareas. Por último, nos centramos en los individuos cuya elección cambia de la lotería segura a la arriesgada en al menos una de las cuatro tareas de riesgo, y también cambia de la lotería temprana a la tardía en al menos una de las seis tareas de tiempo. En otras palabras, eliminamos de la muestra a los individuos que hicieron la misma elección en todas las tareas de riesgo o de tiempo.[18] Estas restricciones dejan una muestra para la estimación compuesta por

[17]Esto se hace para fines de exposición. Extender el modelo añadiendo una probabilidad de error para incluir menús con alternativas dominadas es un procedimiento sencillo. Véase, por ejemplo, Apesteguia y Ballester (2018) y Jagelka (2021).

[18]Esta restricción no es necesaria para calcular estimaciones a nivel poblacional. Sin embargo, es necesaria para obtener estimaciones comparables entre individuos y modelos. La razón es que, en todos los modelos considerados, se requiere variación en las elecciones para identificar puntualmente los parámetros asociados con el coeficiente de aversión al riesgo y la tasa de descuento

202 individuos, cada uno de los cuales se enfrenta a un máximo de 36 menús de riesgo y un máximo de 54 menús de tiempo.

Antes de discutir la metodología para estimar los modelos estructurales, se presentarán los resultados de un estimador semiparamétrico basado en un procedimiento de elicitación utilizado con frecuencia en la literatura. Estas estimaciones constituyen una referencia útil y permiten hacerse una primera idea del grado de heterogeneidad de las preferencias entre y dentro de los individuos del conjunto de datos.

4.1. **Estimación Semiparamétrica.** La literatura suele utilizar las listas de precios múltiples para producir estimaciones de intervalos que contengan a r y δ. Para una tarea de riesgo dada, se pueden identificar los dos menús adyacentes en los que la decisión de un sujeto cambia de la lotería segura a la arriesgada. El umbral de indiferencia $K(A_{\mathcal{R}})$ de estos menús proporciona el límite inferior y superior del intervalo de valores de r consistentes con este cambio. El punto medio de este intervalo se utiliza a menudo como una estimación puntual de r. Utilizando este procedimiento en el conjunto de datos AHLR se obtienen cuatro estimaciones para cada sujeto, que se pueden interpretar como realizaciones de la distribución individual de r. De este modo, es posible estimar μ_r y σ_r para cada individuo a partir de la media y la desviación estándar de las realizaciones obtenidas. También se pueden calcular estimaciones poblacionales de estos parámetros agrupando todas las realizaciones individuales.

Es posible seguir un procedimiento similar para obtener realizaciones de δ a partir de los umbrales de indiferencia $K(A_{\mathcal{T}}|r)$ de los menús adyacentes en una tarea temporal en la que la elección del individuo cambia de la lotería anticipada a la diferida, condicionado a un valor de r específico. Repetimos este procedimiento a lo largo de las seis tareas de tiempo utilizando cada una de las cuatro realizaciones de r obtenidas para este individuo, obteniendo 24 realizaciones de δ para cada individuo. A partir de estos resultados, se calculan estimaciones individuales y poblacionales de μ_δ y σ_δ como antes. También se calculan las estimaciones de ρ a partir de la correlación muestral de estas realizaciones. Denominamos estas estimaciones como *estimaciones semiparamétricas* (SPE) de θ, ya que se obtienen

de un individuo. Si las elecciones son las mismas en todos los menús, solo se puede identificar un intervalo para estos parámetros.

bajo supuestos paramétricos de la función de utilidad, pero no especifican ninguna distribución concreta para (r, δ).

La última columna del cuadro 1 muestra los parámetros estimados utilizando el procedimiento anterior y agrupando todas las realizaciones individuales. Cabe destacar tres resultados. En primer lugar, el coeficiente medio de aversión al riesgo es de 0.715, lo que coincide con las estimaciones de la literatura experimental y con las estimaciones estructurales recogidas en AHLR. Del mismo modo, la tasa de descuento media (anual) es del 13.8%, superior a la estimada por AHLR, pero todavía en el extremo inferior de las estimaciones obtenidas a partir de datos experimentales.

En segundo lugar, existe un alto grado de heterogeneidad en las preferencias. La desviación estándar de todas las realizaciones de r y δ en la muestra es de 0.833 y 0.165, respectivamente. En particular, una gran parte de esta variación corresponde a la heterogeneidad en las decisiones de un mismo sujeto. Para comprobarlo, se calculan los estadísticos resumidos de las estimaciones de (μ_r, μ_δ) a nivel individual y se proporcionan en el cuadro 2.[19] La desviación estándar de μ_r entre individuos es de 0.608, lo que implica que casi la mitad de la varianza de r proviene de la variación entre sujetos. Del mismo modo, la desviación estándar de μ_δ entre individuos es de 0.111, lo que indica que alrededor del 44% de la varianza de δ procede de la variación entre sujetos.

Por último, existe una correlación grande y negativa entre r y δ. El coeficiente de correlación estimado es de -0.761, que es muy similar a la correlación entre las estimaciones individuales (μ_r, μ_δ), que es -0.807. Como se discute en la Proposición 2, valores más altos de r y valores más bajos de δ generan una mayor probabilidad de elección de la alternativa más temprana en los menús de tiempo. Por tanto, la correlación negativa entre r y δ es coherente con el comportamiento observado en los menús de tiempo en la muestra.

4.2. **Estimación Estructural.** Ahora abordamos la estimación estructural de θ usando el modelo paramétrico RDEU y, a efectos de comparación, otras dos alternativas estructurales.

[19] El cuadro 4 en el Apéndice reporta los estadísticos resumidos correspondientes a las estimaciones individuales de $(\sigma_r, \sigma_\delta)$.

El conjunto de datos de este estudio contiene una colección de menús $\{A_m\}_{m=1}^{M}$ y un conjunto de N observaciones para $i = 1, \ldots, I$ individuos, que se denotan como $\mathcal{O}$. La observación (i, m) registra la elección del individuo i en el menú m como una función indicadora $Y_{i,m}$ que toma el valor de cero cuando el individuo elige la lotería temprana/segura del menú (denotada como 0_m), y toma un valor de uno en caso contrario. Para calcular las estimaciones poblacionales, se sigue la literatura y se asume que las preferencias admiten un agente representativo, de modo que $\mathcal{P}_\theta(0_m, A_m)$ es independiente de i. Bajo este supuesto, se puede escribir la función de log-verosimilitud de los datos, condicionada al vector de parámetros θ, como:

$$\log \mathcal{L}(\theta|\mathcal{O}) = \frac{1}{N} \sum_{i,m} \Big[(1 - Y_{i,m}) \log \mathcal{P}_\theta(0_m, A_m) + Y_{i,m} \log(1 - \mathcal{P}_\theta(0_m, A_m)) \Big].$$

El estimador de máxima verosimilitud de θ se calcula maximizando numéricamente la función de log-verosimilitud anterior. Este estimador es consistente y asintóticamente normal bajo condiciones de regularidad estándar, siempre que θ esté identificado. Calculamos los errores estándar robustos de las estimaciones agrupadandolos a nivel individual. También se estiman los parámetros de preferencia por sujeto de forma similar utilizando la submuestra de $\mathcal{O}$ correspondiente a cada individuo.

Solo queda por especificar $\mathcal{P}_\theta(0_m, A_m)$. Las Proposiciones 1 y 2 proporcionan las probabilidades de elección del modelo RDEU.[20] Finalmente, obsérvese que el conjunto de datos AHLR satisface las condiciones expuestas en la Proposición 3. De ello se deduce que el modelo RDEU está identificado y que nuestras estimaciones son consistentes y asintóticamente eficientes.

También se consideran dos modelos alternativos utilizados anteriormente en la literatura. El primer modelo supone que existe una única preferencia subyacente (es decir, $r = \mu_r$ y $\delta = \mu_\delta$) sujeta a ruido i.i.d. aditivo, donde las elecciones están dadas por la siguiente regla:

$$\mathcal{P}_\theta(0_m, A_m) = \frac{DEU_{r,\delta}(0_m)^{\frac{1}{\sigma}}}{DEU_{r,\delta}(0_m)^{\frac{1}{\sigma}} + DEU_{r,\delta}(1_m)^{\frac{1}{\sigma}}}, \tag{4.1}$$

donde $DEU_{r,\delta}(0_m)$ y $DEU_{r,\delta}(1_m)$ denotan, respectivamente, la utilidad esperada descontada de la lotería temprana/segura y la tardía/arriesgada en A_m, como se

[20] El Apéndice C describe el método numérico utilizado para evaluar estas probabilidades de manera eficiente aprovechando el supuesto de que r y δ siguen una distribución normal bivariante.

define en la ecuación (2.1). Esta regla de probabilidad se basa en Luce (1959) y fue introducida en la estimación de las preferencias de riesgo por Holt y Laury (2002). Corresponde a la especificación utilizada en AHLR para calcular estimaciones poblacionales de aversión al riesgo y descuento con sus datos. Siguiendo a estos autores, se especifica $u_r(x) = \frac{(x+\omega)^{1-r}}{1-r}$ y se permite que el parámetro de ruido σ difiera entre las tareas de riesgo y de tiempo, de modo que el modelo se caracteriza por cuatro parámetros: $(\mu_r, \sigma_r, \mu_\delta, \sigma_\delta)$. El modelo se etiqueta como LUCE.

Es importante destacar que este modelo tiene varios problemas teóricos que complican la interpretación de sus estimaciones. En primer lugar, como se muestra en Apesteguia y Ballester (2018), la presencia de choques i.i.d. aditivos hace que $\mathcal{P}_\theta(0_m, A_m)$ sea una función no monótona de (r, δ).[21] En consecuencia, el modelo puede no estar identificado ya que diferentes valores de sus parámetros pueden racionalizar la misma probabilidad observada. En segundo lugar, la forma funcional que suele utilizarse en las valoraciones monetarias genera los problemas discutidos en la Sección 2.1. En particular, las valoraciones son negativas cuando $r > 1$, lo que a su vez genera más problemas en la expresión de la potencia de la ecuación (4.1), involucrando números imaginarios y conduciendo a menores probabilidades de elección de alternativas mejores.[22]

El segundo modelo también supone preferencias deterministas, pero considera en su lugar la siguiente especificación de la probabilidad de elegir la lotería temprana/segura y la tardía/arriesgada en el menú A_m:

$$\mathcal{P}_\theta(0_m, A_m) = \Phi\left(\frac{DEU_{r,\delta}(0_m) - DEU_{r,\delta}(1_m)}{\nu_m \sigma}\right), \tag{4.2}$$

donde ν_m es una constante de normalización específica del menú y σ es un parámetro de ruido que toma valores distintos en las tareas de riesgo y de tiempo. Este modelo se basa en la especificación de "error contextual" propuesta por Wilcox (2011) y aplicada empíricamente por Andersen et al. (2014) y Harrison et al. (2020). Siguiendo a estos autores, suponemos que $\nu_m = 1$ para los menús en las tareas de tiempo

[21] Nótese que, como se discute en Apesteguia y Ballester (2018), las no-monotonicidades son causadas por la no-linealidad de las representaciones de utilidad; el uso estándar de modelos logit mixtos no comparte estos problemas ya que, típicamente, se asume una utilidad latente que es lineal en los parámetros de interés.

[22] Algunos de estos problemas teóricos también se aplican al siguiente modelo alternativo y a el RUM i.i.d. aditivo utilizado en el análisis empírico de los entornos CB. En lo que sigue, nos enfocaremos en las comparaciones empíricas con RDEU.

y establecemos ν_m igual a la diferencia entre la utilidad máxima y la utilidad mínima de los premios en A_m.[23] El modelo se caracteriza, por lo tanto, por los parámetros $(\mu_r, \sigma_r, \mu_\delta, \sigma_\delta)$, y lo denominamos WILCOX.[24]

4.3. **Estimaciones Poblacionales.** Para comenzar, el cuadro 1 presenta los parámetros estimados a nivel de población. La segunda, tercera y cuarta columnas muestran las estimaciones de los modelos RDEU, LUCE y WILCOX, respectivamente.

El coeficiente medio de aversión al riesgo μ_r estimado según el modelo RDEU es de 0.781, ligeramente superior al obtenido a partir de las SPE. Por otra parte, el valor estimado de σ_r es de 0.895, muy similar al obtenido con las estimaciones semiparamétricas. La tasa de descuento media estimada es del 12.5%, también cercana si bien ligeramente inferior al 13.8% de los SPE. La σ_δ estimada es ligeramente inferior pero cercana en magnitud a la desviación típica de los SPE. Por último, la correlación estimada de -0.958 es de gran magnitud y negativa, en consonancia con los resultados de las SPE y la estática comparativa del modelo RDEU analizado en la sección anterior. La estrecha relación entre las estimaciones semiparamétricas y las estimaciones paramétricas del modelo RDEU ilustra la correspondencia intuitiva y estrecha de los parámetros del modelo RDEU con la variación de las elecciones y los menús en el conjunto de datos.

Comparando los resultados del modelo RDEU con los de LUCE y WILCOX, se puede observar que las estimaciones de población de μ_r y μ_δ son muy similares en todos los modelos y, en el caso de μ_r, no se puede rechazar la hipótesis de que son estadísticamente iguales. Sin embargo, los valores estimados de σ_r y σ_δ son bastante diferentes en los distintos modelos. En el caso del modelo RDEU, ambos parámetros tienen una correspondencia directa con la varianza de r y δ de los SPE. Por otro lado, los modelos LUCE y WILCOX tratan estos parámetros como ruido relacionado

[23] A diferencia de estos autores, usamos la función de distribución acumulada de una variable aleatoria normal, en lugar de una distribución logística, para mapear el índice latente a las probabilidades. Esta diferencia no es relevante para los resultados.

[24] La literatura también ha considerado el uso de modelos con coeficientes aleatorios, o modelos logit mixtos, para estimar estructuralmente las preferencias de riesgo y de tiempo (véase, por ejemplo, Andersen et al. (2008) y Andersen et al. (2014)). Estos modelos son muy flexibles y permiten dos niveles de variación: uno a nivel individual y otro a nivel poblacional. A nivel individual, son similares al modelo RUM i.i.d. aditivo descrito anteriormente. A nivel poblacional, permiten variabilidad en r y δ entre individuos. Sin embargo, comparten algunos de los problemas teóricos de los dos modelos anteriores, por lo que no los consideramos aquí.

con la volatilidad de los choques a la utilidad. Por esta razón, su correspondencia con los datos es menos directa.

Si se comparan las log-verosimilitudes de los modelos estimados, se observa que el modelo RDEU se ajusta ligeramente mejor a los datos que los otros dos modelos debido a su mayor capacidad para explicar las elecciones en los menús de tiempo. Esto no es sorprendente, ya que el modelo RDEU permite la correlación entre r y δ, proporcionando un parámetro adicional para ajustarse a los datos. No obstante, las diferencias de ajuste son pequeñas, y los valores estimados de la aversión media al riesgo μ_r y del descuento μ_δ son similares en los tres modelos estructurales.

4.4. **Estimaciones Individuales.** Ahora se analizarán las estimaciones a nivel individual. El cuadro 2 presenta un resumen de varios estadísticos descriptivos de los valores estimados de μ_r y μ_δ para cada individuo según el modelo estructural correspondiente. Las tres últimas filas de este cuadro muestran el coeficiente de correlación de Pearson, el coeficiente de correlación de rangos de Kendall y el coeficiente de correlación de rangos de Spearman entre las estimaciones individuales de cada modelo estructural y de los SPE correspondientes.

Los momentos de las estimaciones individuales son muy similares en los modelos RDEU y SPE. En particular, la media y la desviación típica de μ_r y μ_δ entre los individuos se aproximan notablemente a sus correspondientes estimaciones poblacionales. Además, se observa que las tres medidas de correlación son positivas y muy grandes, lo que proporciona más pruebas de la estrecha relación entre las estimaciones semiparamétricas y RDEU, tanto cualitativa como cuantitativamente.

En cambio, la media y la desviación típica de las estimaciones individuales obtenidas con los modelos LUCE y WILCOX difieren sustancialmente de sus homólogos del modelo poblacional y los SPE. La media del coeficiente de aversión al riesgo en ambos modelos es negativa, y la desviación típica es un orden de magnitud mayor que la estimación poblacional. Del mismo modo, la media y la desviación típica de las estimaciones individuales de μ_δ presentan valores grandes e inverosímiles en los dos modelos alternativos. La correlación de Pearson con SPE en ambos casos es cercana a cero, aunque las medidas de correlación de rango son mayores en comparación. Esto sugiere que los resultados desconcertantes se deben a una proporción de individuos para los que los modelos estiman valores inverosímiles de μ_r y μ_δ. Si

observamos los cuantiles de la distribución de μ_r y μ_δ bajo LUCE y WILCOX, se puede notar que la presencia de una gran masa de valores atípicos en las colas de la distribución da lugar a los valores inesperados para la media y la desviación típica.

Para comprender las diferencias de rendimiento entre los modelos estructurales, la figura 1 presenta gráficas de dispersión de los valores estimados de μ_r y μ_δ frente al valor correspondiente obtenido utilizando estimaciones semiparamétricas. Este último proporciona una buena referencia de los valores del coeficiente de aversión al riesgo y de la tasa de descuento que se esperaría de cada individuo, dadas sus elecciones en las distintas tareas.

La primera fila muestra las gráficas correspondientes al modelo RDEU. Se puede ver que cada punto de ambas gráficas de dispersión está cerca de la línea de 45 grados, lo que confirma la estrecha relación entre las estimaciones de RDEU y SPE observada en los estadísticos resumidos del cuadro 2. El cuadro 4 y la figura 2, en el apéndice, muestran que esta relación estrecha también se mantiene para las estimaciones individuales de σ_r y σ_δ.

La segunda fila de la figura 1 muestra los diagramas de dispersión correspondientes al modelo LUCE. Vale la pena destacar tres observaciones. En primer lugar, existe un marcado límite superior en las estimaciones de μ_r obtenidas con este modelo. Este límite se deriva de los problemas mencionados anteriormente cuando $r > 1$. Esto no es un problema en el caso de las estimaciones poblacionales, dado que, para este conjunto de datos en particular, el coeficiente medio de aversión al riesgo entre los individuos está por debajo de este umbral, como sugieren las SPE. Sin embargo, sí es una restricción problemática para las estimaciones individuales. Según el modelo de SPE, alrededor del 36% de las elecciones en los menús de riesgo son coherentes con $r > 1$, y un tercio de los sujetos de la muestra tienen $\mu_r > 1$. En segundo lugar, la estimación de μ_δ sigue, en promedio, las correspondientes de SPE. Para valores bajos de μ_δ, la estimación LUCE es superior a las correspondientes de SPE. Esto es una consecuencia del límite superior de μ_r: supongamos que un sujeto con $r > 1$ elige la lotería retrasada sobre la temprana en muchos menús de tiempo. Este comportamiento es consistente con tener valores grandes de r, valores bajos de δ, o ambos. Sin embargo, el límite superior hace que el modelo subestime el coeficiente de aversión al riesgo de este individuo. En consecuencia, tiene que sobreestimar el δ

de este sujeto para racionalizar sus elecciones en los menús de tiempo. Por último, obsérvese que hay varios individuos para los que el modelo LUCE estima valores extremadamente bajos de r. Del mismo modo, el modelo estima tasas de descuento inverosímiles, en magnitud o signo, para muchos sujetos de la muestra.

A primera vista, este comportamiento está desconectado de sus elecciones, ya que parece no estar correlacionado con el de SPE. Para entender el origen de este comportamiento errático, se establecen dos grupos de sujetos distintos en cada diagrama de dispersión. En la columna izquierda de la figura 1, el primer grupo (representado en círculos azules) corresponde a 103 sujetos que cambiaron de la lotería segura a la arriesgada en las cuatro tareas de riesgo. Los 99 sujetos restantes (representados en naranja) componen el segundo grupo. Estos sujetos no cambiaron en al menos una de las cuatro tareas de riesgo. La columna derecha de la figura 1 muestra, en círculos azules, a los 112 sujetos que cambiaron de la lotería temprana a la lotería retrasada en las seis tareas de tiempo. Por último, los 90 sujetos restantes que no cambiaron en al menos una de las seis tareas de tiempo se muestran en morado. Se observa un patrón claro: los sujetos para los que el modelo LUCE estima valores inverosímiles de μ_r y μ_δ suelen ser sujetos que no cambiaron en al menos una de las tareas. Estos sujetos muestran preferencias relativamente extremas, junto con cierto grado de estocasticidad en la elección, que el modelo LUCE no puede captar. Vale la pena destacar que estos sujetos no son simples valores atípicos, ya que representan casi la mitad de la muestra en este diseño experimental.

Los resultados correspondientes para el modelo WILCOX se muestran en la tercera fila de la figura 1. En comparación con el modelo LUCE, el modelo capta mejor la heterogeneidad de μ_r y μ_δ reflejada en las estimaciones de SPE. Sin embargo, también proporciona valores inverosímilmente elevados de μ_r para una gran parte de la muestra. El diagrama de dispersión muestra que estas estimaciones grandes suelen obtenerse para los sujetos que no cambiaron en al menos una de las tareas de riesgo. En consecuencia, el modelo también estima valores de μ_δ cercanos a cero para muchos sujetos con valores de SPE tanto pequeños como grandes. Como ya se ha comentado, con valores grandes de r, es más probable que el individuo elija loterías tardías. Supongamos que este individuo elige la lotería temprana en muchos de los menús de tiempo del conjunto de datos. En ese caso, el modelo necesita compensar

el gran valor de r con valores extremadamente bajos de δ para racionalizar sus elecciones. Por último, el modelo también estima valores grandes e inverosímiles de μ_δ para los sujetos que no cambiaron en al menos una de las tareas de tiempo.

Los resultados sugieren que dos enfoques alternativos para estimar estructuralmente las preferencias de riesgo y de tiempo (los modelos LUCE y WILCOX) no son adecuados para captar la gran heterogeneidad en las preferencias observada entre y dentro de los sujetos en los diseños experimentales DMPL. Por el contrario, el modelo RDEU ofrece un marco flexible con sólidos fundamentos teóricos, claras restricciones para su identificación y una conexión intuitiva con las elecciones de DMPL tanto a nivel poblacional como individual.

5. Menús Convexos: Teoría

Aunque los menús convexos pueden parecer más complicados, el análisis puede simplificarse de forma análoga a los DMPL al condicionar de nuevo el parámetro r. Esto crea una estructura intuitiva y ordenada de elecciones para el parámetro δ. Habiendo fijado r, cada $a \in [0, 1)$ tiene un umbral asociado $K(a, A_\mathcal{C}|r) \in (0, 1]$, tal que la elección está por debajo de a si y sólo si el valor del parámetro δ está por arriba del umbral. Es decir, $\Gamma(a, A_\mathcal{C}) = \{(r, \delta) : \delta \geq K(a, A_\mathcal{C}|r)\}$. En el caso de utilidades monetarias convexas, $r \leq 0$, el umbral es único, independiente de a, ya que sólo las soluciones de esquina tienen probabilidad no nula. En el caso de utilidades monetarias estrictamente cóncavas, $r > 0$, el umbral $K(a, A_\mathcal{C}|r)$ corresponde al valor único de δ para el que se cumple la condición de primer orden para la alternativa a, es decir, al valor de δ para el que la derivada de $DEU_{r,\delta}(a)$ respecto de a es igual a cero. El cálculo de las probabilidades de elección resulta nuevamente de la consideración ponderada de todas las distribuciones condicionales $F_{\delta|r}$.

La estática comparativa de los cambios en el parámetro δ es la contraparte, en un espacio continuo, del caso de $A_\mathcal{T}$. Para entender el caso de los cambios en r, se muestra ahora que siempre que $r > 0$, el mapa $\{K(a, A_\mathcal{C}|r)\}_{r\in\mathbb{R}}$ es estrictamente creciente en r si y sólo si $a > \bar{e} = \frac{x^0}{x^0+x^1}$ y estrictamente decreciente cuando $a < \bar{e}$. El valor $\bar{e}$ no es casual, ya que describe la asignación que iguala los dos premios, y por tanto las dos riquezas, entre los períodos t^0 y t^1. Por lo tanto, es posible demostrar que, fijando la distribución marginal de δ y desplazando hacia arriba las

distribuciones condicionales de r, se genera una mayor probabilidad de elección para cualquier vecindad de $\bar{e}$, es decir, las elecciones se vuelven más equilibradas. Este ejercicio de estática comparativa refleja perfectamente el papel de $r > 0$ en $A_{\mathcal{C}}$ como sustitución intertemporal.

La estática comparativa de la dispersión de los parámetros es similar al caso de $A_{\mathcal{T}}$, una vez se tiene en cuenta la continuidad de la variable de elección. En el caso binario de $A_{\mathcal{T}}$, la compensación entre premios anteriores y futuros implica necesariamente la elección de la alternativa $0_{\mathcal{T}}$ frente a la alternativa $1_{\mathcal{T}}$. En el caso continuo actual, esta compensación tiene la alternativa $\bar{e}$ como valor crítico. Las alternativas por debajo (o por arriba) de $\bar{e}$ asignan un premio potencial mayor al período anterior (o al posterior). Ahora se demostrará que, manteniendo constante la distribución marginal de un parámetro, una expansión de las distribuciones condicionales del otro parámetro siempre hace que la probabilidad de elección acumulada $\mathcal{P}_f([0,\bar{e}], A_{\mathcal{C}})$ se acerque a 1/2. Es decir, las probabilidades de elección por debajo y por encima de $\bar{e}$ se acercan, lo que implica que el comportamiento es ahora más estocástico.

Proposición 5. *Para cada par de RDEUs, f y g, y cada menú $A_{\mathcal{C}}$:*

(1) $\mathcal{P}_f([0,a], A_{\mathcal{C}}) = 1 - \int_r F_{\delta|r}(K(a, A_{\mathcal{C}}|r)) f^r(r)\mathrm{d}r$.

(2) (a) *Si $F^r = G^r$, y para todo r $G_{\delta|r}$ domina a $F_{\delta|r}$, $\mathcal{P}_f([0,a], A_{\mathcal{C}}) \geq \mathcal{P}_g([0,a], A_{\mathcal{C}})$ para cada $a \in [0,1)$.*

(b) *Si $F^\delta = G^\delta$, y para todo δ $F_{r|\delta}$ domina a $G_{r|\delta}$, $\mathcal{P}_f([0,\overline{a}], A_{\mathcal{C}}) - \mathcal{P}_f([0,\underline{a}], A_{\mathcal{C}}) \geq \mathcal{P}_g([0,\overline{a}], A_{\mathcal{C}}) - \mathcal{P}_g([0,\underline{a}], A_{\mathcal{C}})$ para todo $0 < \underline{a} < \bar{e} < \overline{a} < 1$.*

(3) (a) *Si $F^r = G^r$, y para todo r $F_{\delta|r}$ expande $G_{\delta|r}$ con $K(\bar{e}, A_{\mathcal{C}}|r) \neq$ med $F(\delta|r)$, $|\mathcal{P}_f([0,\bar{e}], A_{\mathcal{C}}) - \frac{1}{2}| < |\mathcal{P}_g([0,\bar{e}], A_{\mathcal{C}}) - \frac{1}{2}|$.*

(b) *Si $F^\delta = G^\delta$, y para todo δ $F_{r|\delta}$ expande $G_{r|\delta}$ con $K(\bar{e}, A_{\mathcal{C}}|\delta) \neq$ med $F(r|\delta)$, $|\mathcal{P}_f([0,\bar{e}], A_{\mathcal{C}}) - \frac{1}{2}| < |\mathcal{P}_g([0,\bar{e}], A_{\mathcal{C}}) - \frac{1}{2}|$.*

5.1. **Implicaciones para la Versión Paramétrica.** Las implicaciones de la proposición 5 para el caso paramétrico de la CRRA y la normal bivariada están en línea con la discusión general. Con utilidades convexas, el único umbral relevante para δ separa la elección de $a = 0$ y $a = 1$ y corresponde a $K(A_{\mathcal{C}}|r) = \frac{1}{t^1 - t^0} \log \frac{p^1}{p^0} + \frac{1}{t^1 - t^0} \log \left[\frac{(\omega + x^1)^{1-r} - \omega^{1-r}}{(\omega + x^0)^{1-r} - \omega^{1-r}} \right]$. En la parte cóncava, el umbral para δ que determina una

elección por debajo de a puede obtenerse a partir de la condición de primer orden, y corresponde a $K(A_{\mathcal{C}}|r) = \frac{1}{t^1-t^0}\log\frac{p^1}{p^0} + \frac{1}{t^1-t^0}\log\frac{x^1}{x^0} + \frac{1}{t^1-t^0}r\log\left[\frac{(1-a)x^0+\omega}{ax^1+\omega}\right]$.

Hay tres términos en esta expresión: el primer término es el mismo que el primer término del caso convexo y depende de las probabilidades; el segundo término es el límite cuando $r \to 0$ del segundo término del caso convexo, y el tercer término es exclusivo del caso cóncavo. Este último muestra que las soluciones son una función lineal de r. Además, las soluciones son en general interiores y, cuando ω tiende a 0, siempre son interiores. Dado que las probabilidades de elección se construyen de nuevo sobre la base de las probabilidades condicionales que se distribuyen normalmente, el cálculo es rutinario. Ceteris paribus, un aumento de la mediana de δ genera mayores probabilidades de elección para las alternativas que asignan más recursos al período anterior. Dada la naturaleza convexa del menú, el aumento de la mediana de r tiene sobre todo un efecto suavizador, igualando los premios en los dos períodos. Como en el caso anterior, aumentar la varianza de δ o de r conduce a una mayor estocasticidad en la elección.

Como en el caso de las DMPL, la Proposición 5 sienta las bases para la identificación del modelo. Se estudiará ahora la identificación de θ bajo supuestos paramétricos. Considérese un conjunto de datos provenientes de un diseño experimental que implementa un mecanísmo de presupuesto convexo, denotado como $\mathcal{O}$. Este esta compuesto por un conjunto de observaciones de las fichas asignadas por un individuo, o grupo de individuos, cuando se les presenta un conjunto de menús convexos $A_{\mathcal{C},m}$ indexados por $m = 1, \ldots, M$. El siguiente resultado muestra que la variación en el retraso de los pagos, y la variación en la probabilidad de pago o en la tasa de retorno implícita en los menús convexos, son suficientes para identificar θ.

Proposición 6. *Supongamos que el conjunto de datos $\mathcal{O}$ contiene cinco menús convexos con pagos relativamente grandes, tales que $\omega/x^1 \to 0$ y $\omega/x^0 \to 0$, los cuales satisfacen las siguientes condiciones:*

(a) *Dos de los menús $\{A_{\mathcal{C},a}, A_{\mathcal{C},b}\}$ son tales que (i) $t_a^1 - t_a^0 = t_b^1 - t_b^0$ y (ii) $p_a^1/p_a^0 \neq p_b^1/p_b^0$ o $x_a^1/x_a^0 \neq x_b^1/x_b^0$.*

(b) *Los otros tres menús $\{A_{\mathcal{C},c}, A_{\mathcal{C},d}, A_{\mathcal{C},e}\}$ difieren en una de las tres dimensiones:*

(i) *el retraso* $t_m^1 - t_m^0$*, (ii) el premio actual* x_m^0*, (iii) el premio futuro* x_m^1*.*

Entonces, θ *está identificado.*

Intuitivamente, la prueba de la Proposición 6 muestra que se pueden utilizar dos momentos de los datos para identificar la distribución de r: la proporción de elecciones interiores y la elasticidad de la respuesta de las asignaciones de fichas a las probabilidades de pago o a las tasas de retorno. Utilizando estos momentos para identificar (μ_r, σ_r), uno puede entonces centrarse en identificar $(\mu_\delta, \sigma_\delta, \rho)$ a partir de las soluciones de esquina en los datos. En concreto, cuando $r < 0$, el problema del sujeto es análogo a la elección discreta en menús de tiempo estudiada en los entornos DMPL. Por tanto, se pueden usar las mismas condiciones utilizadas para identificar θ_δ en la Proposición 3 para identificar estos parámetros a partir del comportamiento previsto en las asignaciones de esquina.

Al igual que en el caso de las loterías DMPL, también es posible inferir θ utilizando un pequeño número de menús convexos.

Proposición 7. *Para cualquier* $\theta \in \Theta$*, existen cinco menús convexos que permiten inferir su valor exactamente.*

Ejemplo 3. *Considérese de nuevo el caso en que* $\omega \to 0$ *y los parámetros* $(\mu_r, \sigma_r, \mu_\delta, \sigma_\delta, \rho) = (0.7, 0.7, 0.05, 0.02, -0.5)$*. Considere primero el problema convexo definido por las probabilidades* $p^0 = 1$ *y* $p^1 = 0.8$*, los pagos* $x^0 = 15$ *y* $x^1 = 20$*, y los tiempos* $t^0 = 0$ *y* $t^1 = 0 + \epsilon$*, con* ϵ *pequeño. La probabilidad de elección de* $a = 1$ *corresponde a una desviación estándar negativa y, por tanto* $\frac{\mu_r}{\sigma_r} = 1$*. El nivel de aversión al riesgo por arriba de la cual la elección está por debajo de* $a = 0.48$ *es* 0.7*. Dado que la probabilidad de elección acumulada en* $a = 0.48$ *es* 0.5*, se deduce que* $\mu_r = 0.7$*, y de la expresión anterior,* $\sigma_r = 0.7$*. Ahora se puede considerar la versión convexa de los problemas de tiempo descritos en el ejemplo 2 fijando* $p^1 = p^0$*, y reproducir el análisis allí realizado con un hipotético problema de elección discreta en el que se agregan todas las probabilidades observadas de las opciones por debajo de* $1/2$ *y las opciones por encima de* $1/2$*.*

6. Menús Convexos: Ilustración Empírica

A continuación se ilustra la aplicación empírica del modelo RDEU a los presupuestos convexos utilizando datos del diseño experimental de AS. En este estudio, los autores presentan 84 menús convexos a 80 sujetos. En cada menú, el sujeto recibe 100 fichas y decide cuántas asignar a dos fechas: t^0 y t^1. Cada ficha asignada en t^0 se transforma en dólares a un tipo de cambio q^0, por lo que $x^0 = 100q^0$. Del mismo modo, cada ficha asignada en t^1 se transforma en dólares a un tipo de cambio q^1, de tal forma que $x^1 = 100q^1$. Los pagos x^0 y x^1 se otorgan con probabilidades p^0 y p^1, respectivamente. En caso contrario, el sujeto no recibe nada. Todos los menús fijaron t^0 en 7 días y q^0 en 0.20 USD por ficha y varían en cuanto a las características restantes del menú. En consecuencia, el diseño empírico satisface las condiciones para la identificación del modelo RDEU discutidas en la Proposición 6.

Los datos contienen un registro de la proporción de fichas $a \in [0, 1]$ asignadas en t^1 por cada sujeto i cuando se le presenta cada menú en $\{A_m\}_{m=1}^M$. Como las fichas no son divisibles, para la implementación empírica se discretiza el conjunto de elección en S opciones equidistantes $\alpha^1 = [a_1, a_2]$, $\alpha^2 = [a_2, a_3]$, $\ldots$, $\alpha^S = [a_S, a_{S+1}]$, con $a_1 = 0$ y $a_{S+1} = 1$. En los datos, el 93% de las elecciones corresponden a asignaciones de fichas que son múltiplos de 5. En consecuencia, se fija $S = 21$, de modo que $a_2 = 0.025$, $a_3 = 0.075$, $\ldots$, $a_S = 0.975$. Como resultado, el conjunto de datos $\mathcal{O}$ contiene una colección de $M = 84$ menús convexos a los que se enfrentan $I = 80$ individuos, para un total de $N = 6720$ observaciones. La observación (i, m) registra la elección del individuo i en el menú m como una función indicadora $Y_{i,m}(s)$ que adopta el valor uno cuando la asignación de fichas está contenida en α^s, y cero en caso contrario. En el análisis a continuación, se fija $\omega = 5$ USD, en consonancia con el pago por participación presentado en AS.

6.1. **Estimación Estructural.** Para estimar el modelo RDEU, se utilizan las restricciones paramétricas de este estudio y se sigue un enfoque de agente representativo en el que la probabilidad de que $a_m \in \alpha^s$, denominada como $\mathcal{P}_\theta([a_s, a_{s+1}], A_m)$, es independiente de i y está dada por la Proposición 5. Por lo tanto, se puede escribir la función log-verosimilitud de los datos, condicionada al vector de parámetros

θ, como:

$$\log \mathcal{L}\left(\theta|\mathcal{O}\right) = \frac{1}{N}\sum_{i,m}\sum_{s}\Big[Y_{i,m}(s)\log \mathcal{P}_\theta([a_s, a_{s+1}], A_m)\Big].$$

La maximización de la log-verosimilitud anterior produce un estimador consistente y asintóticamente normal de θ en condiciones de regularidad estándar. Se calculan los errores estándar robustos de las estimaciones agrupados a nivel individual y se comparan las estimaciones del modelo RDEU con dos métodos alternativos.

El primer método alternativo sigue la propuesta de AS para la estimación de r y δ a partir de la condición de primer orden asociada al problema de decisión de un individuo que enfrenta un presupuesto convexo, y donde se utilizan mínimos cuadrados no lineales (NLS, por sus siglas en inglés) para minimizar la distancia entre la asignación de fichas prevista y la observada. Este método permite estimar dos parámetros de preferencia, μ_r y μ_δ, sin tener en cuenta explícitamente su heterogeneidad.[25]

El segundo método alternativo emplea un RUM i.i.d. aditivo para estimar las preferencias por riesgo y tiempo en presupuestos convexos (véase Harrison et al. (2013) y Cheung (2015)). En este modelo, la probabilidad de elegir la alternativa a en el menú A_m está dada por:

$$\mathcal{P}_\theta([a_s, a_{s+1}], A_m) = \frac{e^{DEU(\overline{\alpha}^s)}}{e^{DEU(0)} + e^{DEU\left(\overline{\alpha}^2\right)} + \ldots + e^{DEU(1)}},$$

con $\overline{\alpha}^1 = 0$, $\overline{\alpha}^S = 1$, y $\overline{\alpha}^s = (a_s + a_{s+1})/2$ para $s = 2, \ldots, S-1$. Al igual que en el método NLS, este modelo supone que las preferencias son deterministas, de modo que μ_r y μ_δ son los dos únicos parámetros que hay que estimar. Sin embargo, la elección en este modelo es estocástica y, por tanto, potencialmente coherente con la gran heterogeneidad en las asignaciones observada en este tipo de datos.

El cuadro 3 presenta los parámetros estimados según cada modelo. En cuanto a la aversión al riesgo, el modelo RDEU estima $\mu_r = 0.207$ y $\sigma_r = 0.752$. Obsérvese que estas estimaciones son inferiores a las de los diseños DMPL, probablemente debido a que aquí la curvatura r representa tanto la aversión al riesgo como la sustitución

[25]Note que el método NLS impone penalizaciones mayores a las desviaciones más grandes de las condiciones de primer orden, lo que puede interpretarse como una forma de penalizar las desviaciones del promedio de los parámetros subyacentes. El modelo RDEU formaliza este principio en términos de variación del comportamiento, permitiendo producir de forma explícita predicciones probabilísticas.

intertemporal. Sin embargo, su valor promedio es positivo y estadísticamente diferente de cero. Esto contrasta con las estimaciones negativas del RUM i.i.d. aditivo. La razón de estas diferencias es sencilla: alrededor del 48% de las elecciones en los datos corresponden a asignaciones extremas $a = 0$ o $a = 1$. Para explicar el gran número de soluciones de esquina observadas, el RUM i.i.d. aditivo requiere la estimación de funciones de utilidad convexas. En comparación, la flexibilidad adicional del modelo RDEU le permite ajustarse a la gran fracción de soluciones de esquina con una función de utilidad ligeramente cóncava.[26]

En cuanto a la distribución de δ, se estima una tasa de descuento anual promedio de aproximadamente 34%. Esta estimación se aproxima al 26% estimado mediante NLS y es casi la mitad de la tasa anual estimada utilizando el RUM i.i.d. aditivo Esta diferencia puede explicarse por la mayor concavidad de la utilidad estimada en el modelo RDEU, que sustituye a un mayor factor de descuento para explicar las elecciones entre $a = 0$ y $a = 1$ en las esquinas. Por último, se estima una gran variabilidad en el factor de descuento, con una desviación típica de 1.8 y un coeficiente de correlación negativa de -0.16, inferior al obtenido con los datos de DMPL.

La diferencia en las estimaciones de población y las log-verosimilitudes presentan una imagen incompleta de las diferencias entre los modelos. La figura 3 muestra la distribución de las elecciones a en los datos y la compara con la distribución correspondiente predicha por cada modelo.

Lo primero que hay que señalar es que el ajuste del modelo RDEU en la muestra completa es bastante bueno. El modelo se ajusta bien tanto a la proporción de soluciones de esquina como a la presencia de elecciones interiores distribuidas en torno a $a = 0.5$. En comparación, el modelo RUM i.i.d. aditivo no acierta a la hora de predecir la gran proporción de aignaciones de esquina observadas en los datos; tammpoco predice bien la masa de elecciones interiores en torno a $a = 0.5$. Por otro lado, el modelo NLS no modela de forma explicita la estocasticidad de

[26]Una ventaja de disponer de la distribución completa de r es que se pueden estimar momentos adicionales de interés y calcular sus errores estándar correspondientes utilizando el método delta. Por ejemplo, uno podría estar interesado en $E[r|r > 0]$, que proporciona información sobre la curvatura promedio de la función de utilidad inferida de las asignaciones interiores. Según nuestras estimaciones y supuestos paramétricos, estimamos el valor de este momento en 0.68, con un error estándar de 0.067.

la elección, por lo que no es posible usarlo para explicar la heterogeneidad en la elección observada para un menú determinado. En su lugar, se pueden mostrar las elecciones predichas tanto a nivel de menú como entre menús. La segunda columna de la figura 3 muestra la frecuencia de elección observada y la predicha para un menú en particular en la base de datos, usando los tres modelos considerados.[27] Puede verse que el modelo RDEU hace un buen trabajo a la hora de predecir los patrones de elección observados en los presupuestos convexos, tanto para la muestra completa como para menús concretos.

El modelo RDEU explica bien los patrones generales de frecuencia de elección en todos los menús del conjunto de datos. Esto no quiere decir que el modelo RDEU sea capaz de racionalizar cualquier dato. El modelo hereda muchas de las debilidades de los supuestos de utilidad esperada y descuento exponencial. Un ejemplo es la propiedad de la "proporción común", discutida en AS. La figura 4 en el Apéndice muestra la distribución de elección prevista para las tareas que comparten las mismas probabilidades de resultado. Se puede ver que la distribución es idéntica en todos los menús con la misma proporción p^0/p^1, lo que no concuerda con los patrones de elección observados en los datos. No obstante, las herramientas que se presentan en este estudio pueden utilizarse para ampliar el modelo y tener en cuenta estas y otras consideraciones de comportamiento. Lo anterior se ilustra en el Apéndice D mediante la estimación de una pequeña extensión del modelo asumiendo preferencias $\beta-\delta$ como en AS. El análisis teórico de ésta y otras extensiones del modelo se remite a futuras investigaciones.

7. Comentarios finales

En este trabajo se ha estudiado la heterogeneidad de preferencias en el contexto del tratamiento más estándar de las preferencias de riesgo y tiempo, y se propone y estudia el modelo de utilidad esperada descontada aleatoria. Al utilizar la estructura ordenada que vincula los parámetros del modelo con la heterogeneidad en las elecciones observadas, se muestra que el modelo es conveniente desde el punto de vista computacional y está bien fundamentado en términos de estática comparativa.

[27] La figura 3 reporta sobre el menú $(x_0, x_1, p_0, p_1, t_0, t_1) = (20, 20, 0.4, 0.5, 7, 45)$. Se obtienen resultados similares usando cualquiera de los 84 menús en la base de datos. Los gráficos de todos los menús pueden encontrarse en el material suplementario de este artículo.

Además, el modelo se aplica a dos conjuntos de datos muy diferentes, demostrándose que el modelo explica el comportamiento considerablemente bien en ambos casos.

Se considera que este es un enfoque prometedor para el tratamiento de la heterogeneidad cuando intervienen múltiples parámetros, como en el estudio de las preferencias sociales, la ambigüedad, la atención limitada y otras consideraciones conductuales relevantes y sus aplicaciones.

References

[1] Aguiar, V.H. y N. Kashaev (2021). "Stochastic Revealed Preferences with Measurement Error." *Review of Economic Studies* 88(4):2042–2093.

[2] Ahlbrecht, M. y M. Weber (1997). "An Empirical Study on Intertemporal Decision Making under Risk." *Management Science*, 43(6):813–826.

[3] Alan, S. y S. Ertac (2018). "Fostering Patience in the Classroom: Results from Randomized Educational Intervention." *Journal of Political Economy*, 126(5):1865–1911.

[4] Andersen, S., G.W. Harrison, M.I. Lau y E.E. Rutstrom (2008). "Eliciting Risk and Time Preferences." *Econometrica* 76(3):583–618.

[5] Andersen, S., G.W. Harrison, M.I. Lau y E.E. Rutstrom (2014). "Discounting Behavior: A Reconsideration." *European Economic Review* 71:15–33.

[6] Andreoni, J. y C. Sprenger (2012a). "Estimating Time Preferences from Convex Budgets." *American Economic Review*, 102(7):3333–3356.

[7] Andreoni, J. y C. Sprenger (2012b). "Risk Preferences Are Not Time Preferences." *American Economic Review*, 102(7):3357–3376.

[8] Apesteguia, J. y M.A. Ballester (2018). "Monotone Stochastic Choice Models: The Case of Risk and Time Preferences." *Journal of Political Economy*, 126(1):74–106.

[9] Augenblick, N., M. Niederle y C. Sprenger. (2015) "Working over Time: Dynamic Inconsistency in Real Effort Tasks." *Quarterly Journal of Economics*, 130(3):1067–1115.

[10] Barseghyan, L., F. Molinari y M. Thirkettle (2021). "Discrete Choice under Risk with Limited Consideration," *American Economic Review*, 111(6):1972–2006.

[11] Baucells, M. y F. H. Heukamp (2012). "Probability and Time Tradeoff." *Management Science*, 58(4):831–842.

[12] Benjamin, D.J., S.A. Brown y J.M. Shapiro (2013). "Who Is 'Behavioral'? Cognitive Ability and Anomalous Preferences." *Journal of the European Economic Association.* 11(6):1231–1255.

[13] Burks, S.V, J.P. Carpenter, L. Goette y A. Rustichini (2009). "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment." *Proceedings of the National Academy of Sciences*, 106(19):7745–7750.

[14] Carvalho, L. S., S. Meier y S.W. Wang (2016). "Poverty and Economic Decision-Making: Evidence from Changes in Financial Resources at Payday." *American Economic Review*, 106(2):260–284.

[15] Cattaneo, M., X. Ma, Y. Masatlioglu y E. Suleymanov (2020). "A Random Attention Model," *Journal of Political Economy*, 128(7):2796–2836.

[16] Cheung, S. (2015). "Risk Preferences Are Not Time Preferences: On the Elicitation of Time Preference under Conditions of Risk: Comment." *American Economic Review*, 105 (7): 2242-60.

[17] Choi, S., R. Fisman, D. Gale y S. Kariv (2007). "Consistency and Heterogeneity of Individual Behavior under Uncertainty," *American Economic Review*, 97(5):1921–1938.

[18] Coble, K.H. y J. Lusk (2010). "At the Nexus of Risk and Time Preferences: An Experimental Investigation." *Journal of Risk and Uncertainty*, 41(1):67–79.

[19] Dardanoni, V., P. Manzini, M. Mariotti y C. Tyson (2020). "Inferring Cognitive Heterogeneity from Aggregate Choices." *Econometrica*, 88(3):1269–1296.

[20] DellaVigna, S. (2018). "Structural Behavioral Economics." *Handbook of Behavioral Economics-Foundations and Applications 1*, B.D. Bernheim, S. DellaVigna y D. Laibson (eds.), Elsevier, 613–723.

[21] Dohmen T., A. Falk, D. Huffman y U. Sunde (2010). "Are Risk Aversion and Impatience Related to Cognitive Ability?" *American Economic Review*, 100:1238–1260.

[22] Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman y U. Sunde (2018). "Global Evidence on Economic Preferences." *Quarterly Journal of Economics*, 133(4):1645–1692.

[23] Fisman, R., S. Kariv y D. Markovits (2007). "Individual Preferences for Giving." *American Economic Review*, 97(5):1858–1876.

[24] Harrison, G., M. Lau y E.E. Rutstrom (2013). "Identifying Time Preferences with Experiments: Comment." *Center for the Economic Analysis of Risk Working Paper.*

[25] Harrison, G., M. Lau y H. Yoo (2020). "Risk Attitudes, Sample Selection and Attrition in a Longitudinal Field Experiment." *The Review of Economics and Statistics*, 102(3):552–568.

[26] Holt, C. y S. Laury (2002). "Risk Aversion and Incentive Effects ." *American Economic Review*, 92 (5): 1644-1655.

[27] Jagelka, T. (2021). "Are Economists' Preferences Psychologists' Personality Traits?" Mimeo.

[28] Kim, H.B., S. Choi, B. Kim y C. Pop-Eleches (2018). "The Role of Education Interventions in Improving Economic Rationality." *Science*, 362:83–86.

[29] Tanaka, T., C.F. Camerer y Q. Nguyen (2010). "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam." *American Economic Review*, 100(1):557–571.

[30] Wilcox, N. (2011). "Stochastically More Risk Averse: A Contextual Theory of Stochastic Discrete Choice under Risk." *Journal of Econometrics*, 162(1):89–104.

Appendix A. Demostraciones

Demostración de la Proposición 1: Considere un menú $A_{\mathcal{R}} = \{0_{\mathcal{R}}, 1_{\mathcal{R}}\} = \{([p, 1-p; x_1^0, x_2^0], 0), ([p, 1-p; x_1^1, x_2^1], 0)\}$, tal que $x_1^1 > x_1^0 > x_2^0 > x_2^1$ y $p \in (0, 1)$. Considere $r < r'$. Construya las transformaciones afines $v_r, v_{r'}$ de $u_r, u_{r'}$ que satisfacen $v_r(\omega + x_2^1) = v_{r'}(\omega + x_2^1) = 0$ y $v_r(\omega + x_2^0) = v_{r'}(\omega + x_2^0) = 1$. Por la monotonía estricta de las funciones de utilidad originales, debe cumplirse $v_r(\omega + x_1^0) > 1$ y $v_{r'}(\omega + x_1^0) > 1$. Ahora afirmamos que debe cumplirse $v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^0)$, lo cual probaremos por contradicción. Suponga que $v_{r'}(\omega + x_1^0) \geq v_r(\omega + x_1^0) > 1$. En este caso, podemos considerar las loterías $[p^*, 1-p^*; x_1^0, x_2^1]$ y $[1; x_2^0]$, con $\frac{1}{v_{r'}(\omega + x_1^0)} \leq p^* \leq \frac{1}{v_r(\omega + x_1^0)}$. Es inmediato ver que la utilidad esperada construida sobre v_r lleva a, al menos, preferir débilmente la lotería $[1; x_2^0]$, mientras que la utilidad esperada construida sobre $v_{r'}$ lleva a, al menos, preferir débilmente la lotería $[p^*, 1-p^*; x_1^0, x_2^1]$. Esto contradice el hecho de que $v_{r'}$, siendo una transformación estrictamente cóncava de v_r, debe tener un equivalente de certeza estrictamente menor para la segunda lotería, que es más riesgosa, y por lo tanto, hemos probado que $v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^0)$.

Ahora afirmamos que debe cumplirse $v_{r'}(\omega + x_1^1) - v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^1) - v_r(\omega + x_1^0)$, y lo probamos por contradicción. Si no fuera cierto, dado que ya probamos $v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^0)$, tendríamos $\frac{v_{r'}(\omega + x_1^1) - v_{r'}(\omega + x_1^0)}{v_{r'}(\omega + x_1^0)} > \frac{v_r(\omega + x_1^1) - v_r(\omega + x_1^0)}{v_r(\omega + x_1^0)}$. Considerando las loterías $[p', 1-p'; x_1^1, x_2^1]$ y $[1; x_1^0]$, con $\frac{v_{r'}(\omega + x_1^1) - v_{r'}(\omega + x_1^0)}{v_{r'}(\omega + x_1^0)} > \frac{1-p'}{p'} > \frac{v_r(\omega + x_1^1) - v_r(\omega + x_1^0)}{v_r(\omega + x_1^0)}$, la utilidad esperada construida sobre v_r llevaría a elegir $[1; x_1^0]$, mientras que la utilidad esperada construida sobre $v_{r'}$ llevaría a elegir $[p', 1-p'; x_1^1, x_2^1]$. Esto es nuevamente una contradicción a el supuesto de concavidad, con lo que concluye el argumento; entonces debe cumplirse $v_{r'}(\omega + x_1^1) - v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^1) - v_r(\omega + x_1^0)$.

Ahora afirmamos que para cada (r, δ) y (r', δ') tal que $r' > r$, si $DEU_{r,\delta}(0_{\mathcal{R}}) \geq DEU_{r,\delta}(1_{\mathcal{R}})$, entonces $DEU_{r',\delta'}(0_{\mathcal{R}}) > DEU_{r',\delta'}(1_{\mathcal{R}})$. Para verlo, suponga que $DEU_{r,\delta}(0_{\mathcal{R}}) \geq DEU_{r,\delta}(1_{\mathcal{R}})$. Esto equivale a afirmar que la utilidad esperada de la lotería $[p, 1-p; x_1^0, x_2^0]$ es mayor que la utilidad esperada de la lotería $[p, 1-p; x_1^1, x_2^1]$ cuando se utiliza la utilidad monetaria u_r. Esto a su vez es equivale a afirmar que la utilidad esperada de la lotería $[p, 1-p; x_1^0, x_2^0]$ es mayor que la utilidad esperada de la lotería $[p, 1-p; x_1^1, x_2^1]$ cuando se utiliza la utilidad monetaria v_r, y puede escribirse como

$\frac{1-p}{p} \geq \frac{v_r(\omega+x_1^1)-v_r(\omega+x_1^0)}{v_r(\omega+x_2^0)-v_r(\omega+x_2^1)}$. A partir de nuestras afirmaciones anteriores, sabemos que debe cumplirse $\frac{1-p}{p} > v_{r'}(\omega+x_1^1)-v_{r'}(\omega+x_1^0) = \frac{v_{r'}(\omega+x_1^1)-v_{r'}(\omega+x_1^0)}{v_{r'}(\omega+x_2^0)-v_{r'}(\omega+x_2^1)}$, lo que implica que la primera lotería es estrictamente preferida a la segunda utilizando $v_{r'}$ o, alternativamente, utilizando $u_{r'}$. Esto implica que $DEU_{r',\delta'}(0_{\mathcal{R}}) > DEU_{r',\delta'}(1_{\mathcal{R}})$ y concluye el argumento. Con la suposición de curvatura no acotada, el equivalente de certeza de ambas loterías debe converger al pago máximo y mínimo cuando r tiende a $-\infty$ y $+\infty$, respectivamente. Es decir, hay valores de r para los cuales $0_{\mathcal{R}}$ y $1_{\mathcal{R}}$ son preferidos. Como resultado, debe existir un único $K(A_{\mathcal{R}}) \in \mathbb{R}$ tal que la alternativa $0_{\mathcal{R}}$ sea preferida si y solo si $r \geq K(A_{\mathcal{R}})$, lo que lleva a la Afirmación 1 en la proposición. Para la Afirmación 2, observe que siempre que F^r domina a G^r, debe cumplirse $\mathcal{P}_f(0_{\mathcal{R}}, A_{\mathcal{R}}) = 1 - F^r(K(A_{\mathcal{R}})) > 1 - G^r(K(A_{\mathcal{R}})) = \mathcal{P}_g(0_{\mathcal{R}}, A_{\mathcal{R}})$. Para la Afirmación 3, observe que la suposición requiere considerar dos casos, $\text{med}(F^r) > K(A_{\mathcal{R}})$ o $\text{med}(F^r) < K(A_{\mathcal{R}})$. En el primer caso, dado que F^r expande G^r, debe cumplirse $F^r(K(A_{\mathcal{R}})) > G^r(K(A_{\mathcal{R}})) > 1/2$, mientras que en el segundo caso, debe cumplirse $F^r(K(A_{\mathcal{R}})) < G^r(K(A_{\mathcal{R}})) < 1/2$, concluyendo la prueba. ■

Demostración de la Proposición 2: Considere un menú $A_{\mathcal{T}} = \{0_{\mathcal{T}}, 1_{\mathcal{T}}\} = \{([1; x^0], t^0), ([1; x^1], t^1)\}$ tal que $t^0 < t^1$ y $x^0 < x^1$. A partir de la definición de DEU, se deduce inmediatamente que

$$DEU_{r,\delta}(0_{\mathcal{T}}) \geq DEU_{r,\delta}(1_{\mathcal{T}}) \Leftrightarrow \delta \geq K(A_{\mathcal{T}}|r) = \frac{1}{t^1 - t^0} \log \frac{u_r(\omega + x^1)}{u_r(\omega + x^0)}.$$

La monotonía estricta de u_r garantiza que este umbral sea siempre un valor real positivo y, por lo tanto, la primera expresión en la Afirmación 1, y la Afirmación 2a, se deducen inmediatamente.

Ahora afirmamos que el mapa de umbrales $\{K(A_{\mathcal{T}}|r)\}_{r\in\mathbb{R}}$ es estrictamente decreciente en r. Para verlo, observe que las transformaciones escalares no afectan las decisiones DEU. Por lo tanto, podemos seleccionar la transformación escalar v_r de u_r para la cual se cumple $v_r(\omega + x^1) = 1$, y luego, solo necesitamos demostrar que $\frac{-\log v_r(\omega+x^0)}{t^1-t^0}$ es estrictamente decreciente en r o, de manera equivalente, que $v_r(\omega + x^0)$ es estrictamente creciente en r. Supongamos por contradicción que este no es el caso, es decir, que $v_r(\omega + x^0) \geq v_{r'}(\omega + x^0)$ con $r < r'$. Considerando las loterías $[p^*, 1 - p^*; \omega, \omega + x^1]$ y $[1; \omega + x^0]$, donde $v_r(\omega + x^0) \geq p^* \geq v_{r'}(\omega + x^0)$, es

inmediato ver que la utilidad esperada, usando v_r, es mayor para la primera lotería que para la segunda, mientras que la utilidad esperada, usando $v_{r'}$, es mayor para la segunda lotería que para la primera, lo que contradice la suposición de estricta concavidad. El mapa de umbrales es, por lo tanto, estrictamente decreciente en r.

La suposición de curvatura no acotada también demuestra que el mapa cubre completamente $\mathbb{R}_{++}$ y, por lo tanto, es una aplicación biyectiva entre $\mathbb{R}$ y $\mathbb{R}_{++}$. Así, puede invertirse para obtener los umbrales estrictamente decrecientes $\{K(A_\mathcal{T}|\delta)\}_{\delta\in\mathbb{R}_{++}}$, tales que, para un $\delta > 0$, la alternativa $0_\mathcal{T}$ es elegida si y solo si r está por encima de este umbral. Para $\delta \leq 0$, la alternativa $1_\mathcal{T}$ siempre es elegida. Por lo tanto, $\mathcal{P}_f(0_\mathcal{T}, A_\mathcal{T}) = \int_r(1 - F_{\delta|r}(K(A_\mathcal{T}|r)))f^r(r)\mathrm{d}r = \int_{\delta>0}(1 - F_{r|\delta}(K(A_\mathcal{T}|\delta)))f^\delta(\delta)\mathrm{d}\delta$, donde f^δ es la densidad marginal de δ. La segunda expresión en la Afirmación 1, y la Afirmación 2b, se deducen.

Para la Afirmación 3a, solo necesitamos reproducir la lógica de la Proposición 1, expandiendo por separado cada una de las distribuciones condicionales $F_{\delta|r}$. Esto siempre crea una mayor estocasticidad condicional. A partir de ahí, debemos demostrar que el argumento se extiende a la agregación ponderada de todas estas distribuciones condicionales. Para verlo, observe que la continuidad del mapa $\{K(A_\mathcal{T}|r)\}_{r\in\mathbb{R}}$ garantiza que todas las medianas condicionales de δ se encuentran en el mismo lado del mapa de umbrales. Como resultado, la misma alternativa, ya sea $0_\mathcal{T}$ o $1_\mathcal{T}$, se elige más a menudo en cada una de las condicionales, y el argumento de expansión se extiende a la agregación. Para la Afirmación 3b, un argumento similar se aplica expandiendo las condicionales $F_{r|\delta}$ y utilizando la continuidad del mapa inverso. ■

Demostración de la Proposición 3: Comenzamos mostrando que $\theta_r \equiv (\mu_r, \sigma_r)$ está identificado. Supongamos, de forma contraria, que este no es el caso: existen θ'_r y θ^*_r en Θ_r tales que $\theta'_r \neq \theta^*_r$ y $\mathcal{P}_{\theta'}(1_\mathcal{R}, A_\mathcal{R}) = \mathcal{P}_{\theta^*}(1_\mathcal{R}, A_\mathcal{R})$. Usando la Proposición 1, podemos escribir esta igualdad como:

$$\mathcal{P}_{\theta'_r}(r \leq K(A_\mathcal{R})) = \Phi\left(\frac{K(A_\mathcal{R}) - \mu'_r}{\sigma'_r}\right) = \Phi\left(\frac{K(A_\mathcal{R}) - \mu^*_r}{\sigma^*_r}\right) = \mathcal{P}_{\theta^*_r}(r \leq K(A_\mathcal{R})).$$

Dado que $\Phi(\cdot)$ es una función estrictamente monótona, la última igualdad implica:

$$-\frac{\mu_r^*}{\sigma_r^*}+\frac{1}{\sigma_r^*}K\left(A_{\mathcal{R}}\right)=-\frac{\mu_r'}{\sigma_r'}+\frac{1}{\sigma_r'}K\left(A_{\mathcal{R}}\right), \tag{A.1}$$

para cada menú $A_{\mathcal{R}}$. Dado que $\theta_r' \neq \theta_r^*$, existen tres casos posibles: $\mu_r' \neq \mu_r^*$ y $\sigma_r' = \sigma_r^*$; $\mu_r' = \mu_t^*$ y $\sigma_r' \neq \sigma_r^*$; y $\mu_r' \neq \mu_r^*$ y $\sigma_r' \neq \sigma_r^*$. En el primer caso, la ecuación (A.1) implica $\mu_r' = \mu_r^*$, llevando a una contradicción. Consideremos ahora los casos segundo y tercero donde $\sigma_r' \neq \sigma_r^*$. Evaluando (A.1) para $A_{\mathcal{R},a}$ y $A_{\mathcal{R},b}$ y combinando las expresiones resultantes obtenemos:

$$\left(\frac{1}{\sigma_r^*}-\frac{1}{\sigma_r'}\right)\left(K\left(A_b^{\mathcal{R}}\right)-K\left(A_a^{\mathcal{R}}\right)\right)=0. \tag{A.2}$$

Dado que $K\left(A_{\mathcal{R},a}\right) \neq K\left(A_{\mathcal{R},b}\right)$ por la suposición (a), debe ser el caso que $\sigma_r' = \sigma_r^*$, llegando a una contradicción. Concluimos así que θ_r está identificado. El siguiente paso es demostrar que $\theta_\delta \equiv (\mu_\delta, \sigma_\delta, \rho)$ está identificado. Fijamos (μ_r, σ_r) y suponemos, por el contrario, que θ_δ no está identificado: existen θ_δ' y θ_δ^* en Θ tales que $\theta_\delta' \neq \theta_\delta^*$ y $\mathcal{P}_{\theta_\delta'}(0_{\mathcal{T}}, A_{\mathcal{T}}) = \mathcal{P}_{\theta_\delta^*}(0_{\mathcal{T}}, A_{\mathcal{T}})$. Usando la Proposición 2, la igualdad de probabilidades implica:

$$\int_{-\infty}^{\infty}\left\{\Phi\left(\frac{K\left(A_{\mathcal{T}}|r\right)-\mu_{\delta|r}'}{\sigma_{\delta|r}'}\right)-\Phi\left(\frac{K\left(A_{\mathcal{T}}|r\right)-\mu_{\delta|r}^*}{\sigma_{\delta|r}^*}\right)\right\}\phi\left(\frac{r-\mu_r}{\sigma_r}\right)dr=0,$$

con $\mu_{\delta|r} = \mu_\delta + \rho\sigma_\delta v(r)$, $\sigma_{\delta|r} = \sigma_\delta\sqrt{1-\rho^2}$, y $\nu(r) \equiv \frac{r-\mu_r}{\sigma_r}$. El término entre corchetes en la expresión anterior está acotado en $[-1, 1]$. Por la continuidad y monotonicidad de $\phi(\cdot)$ y $\Phi(\cdot)$, existe $\overline{r}_m \in \mathbb{R}$ para cada uno de los menús $\{A_{\mathcal{T},c}, A_{\mathcal{T},d}, A_{\mathcal{T},e}\}$ tal que:

$$-\frac{\mu_\delta'}{\sigma_{\delta|r}'}-\alpha' v\left(\overline{r}_m\right)+\frac{1}{\sigma_{\delta|r}'}K\left(A_{\mathcal{T},m}|\overline{r}_m\right)=-\frac{\mu_\delta^*}{\sigma_{\delta|r}^*}-\alpha^* v\left(\overline{r}_m\right)+\frac{1}{\sigma_{\delta|r}^*}K\left(A_{\mathcal{T},m}|\overline{r}_m\right), \tag{A.3}$$

donde $\alpha \equiv \rho/\sqrt{1-\rho^2}$ y usamos el hecho de que $\sigma_{\delta|r}$ es independiente de $\overline{r}_m$. Ahora, cualquiera de las tres condiciones en la Suposición (b) implica $K\left(A_{\mathcal{T},c}|r\right) \neq K\left(A_{\mathcal{T},d}|r\right) \neq K\left(A_{\mathcal{T},e}|r\right) \neq K\left(A_{\mathcal{T},c}|r\right)$ para cualquier $r \in \mathbb{R}$. Usando este resultado y el Teorema de la Función Implícita, concluimos que $\overline{r}_c \neq \overline{r}_d \neq \overline{r}_e \neq \overline{r}_c$. Estableciendo $m = c$ en (A.3) y restando la expresión correspondiente para $m = d$, obtenemos:

$$\left[\alpha'-\alpha^*\right]\left(v\left(\overline{r}_c\right)-v\left(\overline{r}_d\right)\right)+\left[\frac{1}{\sigma_{\delta|r}'}-\frac{1}{\sigma_{\delta|r}^*}\right]\left(K\left(A_{\mathcal{T},c}|\overline{r}_c\right)-K\left(A_{\mathcal{T},d}|\overline{r}_d\right)\right)=0. \tag{A.4}$$

Repitiendo este procedimiento para los menús $A_{\mathcal{T},c}$ y $A_{\mathcal{T},e}$, obtenemos:

$$[\alpha' - \alpha^*]\left(v\left(\overline{r}_c\right) - v\left(\overline{r}_e\right)\right) + \left[\frac{1}{\sigma'_{\delta|r}} - \frac{1}{\sigma^*_{\delta|r}}\right]\left(K\left(A_{\mathcal{T},c}|\overline{r}_c\right) - K\left(A_{\mathcal{T},e}|\overline{r}_e\right)\right) = 0. \tag{A.5}$$

Usando (A.5) para reemplazar $\left[1/\sigma'_{\delta|r} - 1/\sigma^*_{\delta|r}\right]$ en (A.4), obtenemos:

$$[\alpha' - \alpha^*]\left[\overline{r}_c - \overline{r}_d\right]\left[1 - \left(\frac{\overline{r}_c - \overline{r}_e}{\overline{r}_c - \overline{r}_d}\right)\left(\frac{K\left(A_{\mathcal{T},c}|\overline{r}_c\right) - K\left(A_{\mathcal{T},d}|\overline{r}_d\right)}{K\left(A_{\mathcal{T},c}|\overline{r}_c\right) - K\left(A_{\mathcal{T},e}|\overline{r}_e\right)}\right)\right] = 0. \tag{A.6}$$

Dado que $K\left(A_{\mathcal{T},c}|\overline{r}_c\right) \neq K\left(A_{\mathcal{T},d}|\overline{r}_d\right) \neq K\left(A_{\mathcal{T},e}|\overline{r}_e\right) \neq K\left(A_{\mathcal{T},c}|\overline{r}_c\right)$, la ecuación (A.6) implica que $\alpha' = \alpha^*$, lo que a su vez implica $\rho' = \rho^*$. Este resultado y la ecuación (A.5) implican que $\sigma'_{\delta|r} = \sigma^*_{\delta|r}$, por lo que $\sigma'_\delta = \sigma^*_\delta$. Usando la ecuación (A.3), concluimos que $\mu'_\delta = \mu^*_\delta$, llegando a una contradicción. Esto concluye la prueba. ■

Demostración de la Proposición 4: Para simplificar la presentación del resultado y proporcionar intuiciones claras, utilizamos dos supuestos relativamente suaves: (i) la probabilidad del evento $\{r > 1, \delta < 0\}$ es pequeña y (ii) existen dos menús de tiempo $A_{\mathcal{T}_1}$ y $A_{\mathcal{T}_2}$ con los mismos pagos donde las probabilidades de seleccionar las opciones $0_{\mathcal{T}_1}$ y $1_{\mathcal{T}_2}$ son mayores que $\frac{1}{2}$. Para motivar (i), nótese que $\delta < 0$ ya corresponde al evento raro en el que el individuo tiene una preferencia estricta por el futuro, y estamos combinando esto con el efecto adicional de una curvatura mayor que logarítmica. Para motivar (ii), nótese que podemos hacer que las diferencias en los tiempos sean tan pequeñas o grandes como deseemos.

Ahora, a partir de la Proposición 1, sabemos que $\mathcal{P}_f(0_\mathcal{R}, A_\mathcal{R}) = 1 - F^r(K(A_\mathcal{R}))$, que en la versión paramétrica se lee como $\mathcal{P}_f(0_\mathcal{R}, A_\mathcal{R}) = 1 - \Phi(\frac{K(A_\mathcal{R}) - \mu_r}{\sigma_r})$. Fijemos cualquier valor de $p \in (0,1)$, tres pagos $x_1^0 > x_2^0 > x_2^1 > 0$, y consideremos menús de riesgo que varían solo en el pago x_1^1. Los umbrales de estos menús son estrictamente crecientes en x_1^1 y forman una biyección con los números reales. Por lo tanto, existen dos menús $A_{\mathcal{R}_1}$ y $A_{\mathcal{R}_2}$ tales que las probabilidades de elección para las opciones $0_{\mathcal{R}_1}$ y $0_{\mathcal{R}_2}$ son iguales a $1 - \Phi(0)$ y $1 - \Phi(1)$, respectivamente. Debe ser entonces $\mu_r = K(A_{\mathcal{R}_1})$ y $\sigma_r = K(A_{\mathcal{R}_2}) - \mu_r = K(A_{\mathcal{R}_2}) - K(A_{\mathcal{R}_1})$.

Ahora discutimos los menús de tiempo y, usando pagos suficientemente grandes en todos nuestros argumentos, podemos asumir sin pérdida de generalidad que el comportamiento corresponde al caso $\omega \to 0$. En este caso límite, el mapa umbral

condicional se vuelve lineal por partes: (a) cuando $r < 1$, $K(A_{\mathcal{T}}|r) = \frac{\log \frac{x^1}{x^0}}{t^1 - t^0}(1-r) \equiv K(A_{\mathcal{T}})(1-r) > 0$, donde $K(A_{\mathcal{T}})$ es una constante dependiente del menú, y (b) cuando $r \geq 1$, $K(A_{\mathcal{T}}|r)$ se vuelve nulo. Por lo tanto, la probabilidad de elegir $0_{\mathcal{T}}$ corresponde a la probabilidad de que δ esté por encima de $\min\{0, K(A_{\mathcal{T}})(1-r)\}$ y, dada nuestra suposición (i), esto puede aproximarse por la probabilidad de que δ esté por encima de $K(A_{\mathcal{T}})(1-r)$. Esto equivale a la probabilidad de que la variable aleatoria $K(A_{\mathcal{T}})(1-r) - \delta$ sea menor que cero. Dada la normalidad, esta variable aleatoria también es normal con media $-\mu_\delta - K(A_{\mathcal{T}})\mu_r + K(A_{\mathcal{T}})$ y desviación estándar $\sqrt{K^2(A_{\mathcal{T}})\sigma_r^2 + \sigma_\delta^2 + 2\rho K(A_{\mathcal{T}})\sigma_r\sigma_\delta}$. Así, la probabilidad de elección de $0_{\mathcal{T}}$ es aproximadamente $\Phi\big(\frac{\mu_\delta + K(A_{\mathcal{T}})\mu_r - K(A_{\mathcal{T}})}{\sqrt{K^2(A_{\mathcal{T}})\sigma_r^2 + \sigma_\delta^2 + 2\rho K(A_{\mathcal{T}})\sigma_r\sigma_\delta}}\big)$.

Por la suposición (ii), existen dos menús de tiempo $A_{\mathcal{T}_1}$ y $A_{\mathcal{T}_2}$ con los mismos pagos x^0, x^1 tales que las probabilidades de elección de $0_{\mathcal{T}}$ están por encima y por debajo de $\Phi(0) = \frac{1}{2}$, respectivamente. Debido a la estacionariedad de DEU, podemos asumir sin pérdida de generalidad que el pago más temprano ocurre en el presente en ambos casos. Por continuidad, debe existir un único t, y por lo tanto, un menú $A_{\mathcal{T}_3} = \{([1; x^0], 0), ([1; x^1], t)\}$ tal que la probabilidad de elección de la opción $0_{\mathcal{T}_3}$ sea exactamente $\Phi(0)$. Por lo tanto, $\mu_\delta = K(A_{\mathcal{T}_3})(1-\mu_r) = K(A_{\mathcal{T}_3})(1 - K(A_{\mathcal{R}_1}))$. Segundo, consideremos cualquier secuencia de problemas de tiempo $\{A_{\mathcal{T}^n}\}$ tal que $\lim_n K(A_{\mathcal{T}^n}) = 0$. Denotemos por q el límite de las probabilidades de elección de la opción $0_{\mathcal{T}^n}$. Sabemos que $q = \Phi(\frac{\mu_\delta}{\sigma_\delta})$ y, por lo tanto, debe ser $\sigma_\delta = \frac{\mu_\delta}{\Phi^{-1}(q)} = \frac{K(A_{\mathcal{T}_3})(1-K(A_{\mathcal{R}_1}))}{\Phi^{-1}(q)}$. Finalmente, al fijar nuevamente cualquier tres parámetros en un menú de tiempo y variar el cuarto, sabemos que existe un único menú de tiempo $A_{\mathcal{T}_4}$ en dicha familia para el cual $K(A_{\mathcal{T}_4}) = \frac{\sigma_\delta}{\sigma_r}$.[28] Denotemos por q' la probabilidad de elección de la opción $0_{\mathcal{T}_4}$. Entonces, debe ser $q' = \Phi(\frac{\frac{\mu_r - 1}{\sigma_r} + \frac{\mu_\delta}{\sigma_\delta}}{\sqrt{2(1+\rho)}})$. Nótese que el mapeo del lado derecho es estrictamente creciente o estrictamente decreciente en ρ, lo que permite obtener ρ: $\rho = \frac{1}{2}\left[\frac{\frac{\mu_r - 1}{\sigma_r} + \frac{\mu_\delta}{\sigma_\delta}}{\Phi^{-1}(q')}\right]^2 - 1 = \frac{1}{2}\left[\frac{\frac{K(A_{\mathcal{R}_1}) - 1}{K(A_{\mathcal{R}_2}) - K(A_{\mathcal{R}_1})} + \Phi^{-1}(q)}{\Phi^{-1}(q')}\right]^2 - 1$. Esto concluye la demostración. ■

Demostración de la Proposición 5: Considere un menú $A_{\mathcal{C}}$ definido por $(p^0, x^0, t^0; p^1, x^1, t^1)$. Primero afirmamos que, para todo $r \in \mathbb{R}$, el argumento que

[28] Esto es, en general, diferente de $A_{\mathcal{T}_3}$. De lo contrario, observe que el mapeo $\Phi\big(\frac{\mu_\delta + K(A_{\mathcal{T}})\mu_r - K(A_{\mathcal{T}})}{\sqrt{K^2(A_{\mathcal{T}})\sigma_r^2 + \sigma_\delta^2 + 2\rho K(A_{\mathcal{T}})\sigma_r\sigma_\delta}}\big)$ es estrictamente monótono en ρ, y por lo tanto el parámetro puede recuperarse usando algún otro menú de tiempo.

maximiza $DEU_{r,\delta}$ es decreciente en δ. Para ver esto, considere cualquier par de parámetros (r, δ), y sea $a^* \in [0, 1]$ el argumento que maximiza $DEU_{r,\delta}$. Si $a^* = 1$, hemos terminado. Considere entonces el caso de $a^* < 1$ y cualquier alternativa $a^* < a \leq 1$. Dada la optimalidad de a^*, sabemos que $DEU_{\delta,r}(a) \leq DEU_{\delta,r}(a^*)$, es decir, se cumple que $e^{-\delta t^0} p^0 u_r(\omega + (1-a)x^0) + e^{-\delta t^1} p^1 u_r(\omega + ax^1) \leq e^{-\delta t^0} p^0 u_r(\omega + (1-a^*)x^0) + e^{-\delta t^1} p^1 u_r(\omega + a^* x^1)$. La última desigualdad es equivalente a $p^0 u_r(\omega + (1-a)x^0) + e^{-\delta(t^1-t^0)} p^1 u_r(\omega + ax^1) \leq p^0 u_r(\omega + (1-a^*)x^0) + e^{-\delta(t^1-t^0)} p^1 u_r(\omega + a^* x^1)$. Ahora, es evidente que un aumento de δ no afecta al primer término en ambos lados de la desigualdad, pero disminuye más significativamente el segundo término del lado izquierdo, porque la función u_r es estrictamente creciente. Por lo tanto, la alternativa a^* debe ser preferida a la alternativa a para el δ más grande, y el argumento que maximiza DEU debe ser a^* o menor. Hemos probado nuestra afirmación. Por lo tanto, dado $r \in \mathbb{R}$, podemos definir $K(a, A_{\mathcal{C}}|r)$, $a \in [0, 1)$, como el ínfimo de los valores de δ para los cuales cualquier alternativa en $[0, a]$ es el maximizador de DEU. Por lo tanto, el maximizador de $DEU_{r,\delta}$ está por debajo de a si y solo si δ está por encima de $K(a, A_{\mathcal{C}}|r)$, es decir, $\Gamma(a, A) = \{(r, \delta) : \delta \leq K(a, A_{\mathcal{C}}|r)\}$, y las Afirmaciones 1 y 2a se derivan.

Estudiamos ahora la estructura de los umbrales. Comenzamos con el caso de utilidades monetarias convexas, es decir, $r \leq 0$. La convexidad y el hecho de que $u_r(\omega) = 0$ garantizan que $e^{-\delta t^0} p^0 u_r(\omega + (1-a)x^0) + e^{-\delta t^1} p^1 u_r(\omega + ax^1) \leq e^{-\delta t^0} p^0 [a u_r(\omega) + (1-a) u_r(\omega + x^0)] + e^{-\delta t^1} p^1 [(1-a) u_r(\omega) + a u_r(\omega + x^1)] = e^{-\delta t^0} p^0 (1-a) u_r(\omega + x^0) + e^{-\delta t^1} p^1 a u_r(\omega + x^1) \leq \max\{e^{-\delta t^0} p^0 u_r(\omega + x^0), e^{-\delta t^1} p^1 u_r(\omega + x^1)\}$. Por lo tanto, solo las alternativas 0 o 1 pueden ser los maximizadores de $DEU_{r,\delta}$. Así, para cada menú $A_{\mathcal{C}}$ existe un único umbral $K(a, A_{\mathcal{C}}|r) \in \mathbb{R}$, independiente de a, que corresponde al δ que, dado r, iguala el valor de DEU de 0 y 1. Este valor es $\frac{1}{t^1 - t^0} \log \left[\frac{p^1 u_r(\omega + x^1)}{p^0 u_r(\omega + x^0)} \right]$, que también se puede escribir como $\frac{1}{t^1 - t^0} \log \frac{p^1}{p^0} + K(A_{\mathcal{T}}|r)$, donde $K(A_{\mathcal{T}}|r)$ se refiere al menú temporal hipotético en el que los premios x^0 y x^1 se ofrecen en los períodos t^0 y t^1, sin considerar la probabilidad de estos premios. La Proposición 2 argumentó que $K(A_{\mathcal{T}}|r)$ es estrictamente decreciente, y por lo tanto $K(a, A_{\mathcal{C}}|r)$ también es estrictamente decreciente siempre que $r \leq 0$.

Analizamos ahora las utilidades estrictamente cóncavas, $r > 0$. Comenzamos afirmando que el umbral $K(a, A_{\mathcal{C}}|r)$ es decreciente para cada $a > \bar{e}$, y creciente

para cada $a < \bar{e}$. Comenzamos con el primero, asumiendo por contradicción que $0 < r < r'$ pero $K(a, A_{\mathcal{C}}|r) < K(a, A_{\mathcal{C}}|r')$ para algún $a > \bar{e}$. Usando la continuidad y la definición de los umbrales, debe existir un δ^* con $K(a, A_{\mathcal{C}}|r) < \delta^* < K(a, A_{\mathcal{C}}|r')$ tal que el maximizador para DEU_{r,δ^*} es a^*, con $\bar{e} < a^* < a$. Considere cualquier $a' > a^*$. Debe cumplirse que $DEU_{r,\delta^*}(a^*) \geq DEU_{r,\delta^*}(a')$, es decir, $e^{-\delta^* t^0} p^0 u_r(\omega + (1-a^*)x^0) + e^{-\delta^* t^1} p^1 u_r(\omega + a^* x^1) \geq e^{-\delta^* t^0} p^0 u_r(\omega + (1-a')x^0) + e^{-\delta^* t^1} p^1 u_r(\omega + a' x^1)$. Dividiendo ambos términos por la constante positiva $p^0 e^{-\delta^* t^0} + p^1 e^{-\delta^* t^1}$ y denotando $p = \frac{p^0 e^{-\delta^* t^0}}{p^0 e^{-\delta^* t^0} + p^1 e^{-\delta^* t^1}}$, la expresión anterior se puede escribir como $p u_r(\omega + (1 - a^*)x^0) + (1-p)u_r(\omega + a^* x^1) \geq p u_r(\omega + (1-a')x^0) + (1-p)u_r(\omega + a' x^1)$. Por lo tanto, la comparación de estas dos alternativas es equivalente a la de un menú de riesgo, con a^* correspondiente a la alternativa $0_{\mathcal{R}}$ y a' a la alternativa $1_{\mathcal{R}}$. Así, dado que a^* es preferido en (r, δ^*), sabemos por la Proposición 1 que a^* también será preferido en (r', δ^*) porque $r' > r$.[29] Por lo tanto, el maximizador de DEU_{r',δ^*} no puede estar por encima de a^*. Esto contradice la definición de $K(a, A_{\mathcal{C}}|r')$. Es decir, el umbral debe ser decreciente siempre que $a > \bar{e}$. Dado que la familia $\{u_r\}$ está estrictamente ordenada por concavidad, el umbral debe ser estrictamente decreciente. La demostración de que el umbral es estrictamente creciente siempre que $a < \bar{e}$ es análoga y, por lo tanto, se omite.

Considere ahora $0 < \underline{a} < \bar{e} < \overline{a} < 1$ y denote por $\delta^{\bar{e}}$ el valor de δ que hace indiferentes a todas las alternativas cuando $r = 0$. A partir del razonamiento anterior, siempre que $r > 0$, $K(\underline{a}, A_{\mathcal{C}}|r)$ (resp., $K(\overline{a}, A_{\mathcal{C}}|r)$) está por encima (resp., por debajo) de $\delta^{\bar{e}}$. Por lo tanto, para cualquier $\delta > \delta^{\bar{e}}$ (resp., $\delta < \delta^{\bar{e}}$), un cambio dominante en la distribución condicional de r genera un aumento en la masa condicional del conjunto de valores de r que están por encima del inverso del umbral $K(\underline{a}, A_{\mathcal{C}}|r)$ (resp., $K(\overline{a}, A_{\mathcal{C}}|r)$). La Afirmación 2b se deduce entonces.

Ahora demostramos la Afirmación 3a. Sabemos que la elección pertenece a $[0, \bar{e}]$ si y solo si $\delta > K(\bar{e}, A_{\mathcal{C}}|r)$. Podemos reproducir el análisis de la Proposición 2 para el caso de expansiones en la distribución condicional de δ, y así, se deriva la Afirmación 3a. Para mostrar la Afirmación 3b, considere cualquier secuencia de valores $\{a^n\}$, con $a^n > \bar{e}$, tal que $\lim_n a^n = \bar{e}$. Sabemos que la elección pertenece a $[0, a^n]$ si y solo si $\delta > K(a^n, A_{\mathcal{C}}|r)$. Dado que para cada a^n el umbral es estrictamente decreciente,

[29] Observe que mantenemos δ^* constante porque este valor es parte de la definición de las loterías.

podemos invertir estos mapeos y luego reproducir el análisis de la Proposición 2 para el caso de expansiones en la distribución condicional de r, y así, se deriva la Afirmación 3b. ■

Demostración de la Proposición 6: Comenzamos demostrando que $\theta_r \equiv (\mu_r, \sigma_r)$ está identificado. Supongamos, por el contrario, que este no es el caso: existen θ' y θ^* en Θ con $\theta'_r \neq \theta^*_r$ tal que la distribución de los datos es la misma bajo ambos parámetros. Sea $\log \widetilde{R} \equiv \log \frac{x^1}{x^0} + \log \frac{p^1}{p^0}$ y $k \equiv t^1 - t^0$. Nótese que, bajo la suposición de que $x^0/\omega \to 0$ y $x^1/\omega \to 0$, la probabilidad de asignaciones extremas, condicionado a $r > 0$, converge a cero. En consecuencia, la probabilidad de observar asignaciones extremas $a \in \{0, 1\}$ está dada por $\mathcal{P}_{\theta'}(a \in \{0, 1\}, A_{\mathcal{C}}) = \mathcal{P}_{\theta'}(r \leq 0) = \Phi(-\mu'_r/\sigma'_r)$ la cual, por la suposición de no-identificación, es igual a $\Phi(-\mu^*_r/\sigma^*_r)$. Se deduce que $\mu'_r/\sigma'_r = \mu^*_r/\sigma^*_r$. En consecuencia, debe cumplirse tanto $\mu'_r \neq \mu^*_r$ como $\sigma'_r \neq \sigma^*_r$. De lo contrario, $\theta'_r = \theta^*_r$. Ahora, usando la condición de primer orden del problema, tenemos:

$$\mathbb{E}_{\theta'}\left[\log\left(\frac{ax^1_m + \omega}{(1-a)\,x^0_m + \omega}\right)|r > 0\right] = -\mathbb{E}_{\theta'}\left[\frac{\delta}{r}|r > 0\right] k_m + \mathbb{E}_{\theta'}\left[\frac{1}{r}|r > 0\right] \log \widetilde{R}_m,$$

para cualquier menú $A_{\mathcal{C},m}$. Combinando las expresiones correspondientes para $A_{\mathcal{C},a}$ y $A_{\mathcal{C},b}$, y usando la suposición de que $k_a = k_b$, obtenemos:

$$\mathbb{E}_{\theta'}[\Delta c|r > 0] = \mathbb{E}_{\theta'}\left[\frac{1}{r}|r > 0\right]\left(\log \widetilde{R}_a - \log \widetilde{R}_b\right), \tag{A.7}$$

con

$$\Delta c \equiv \log\left(\frac{ax^1_a + \omega}{(1-a)\,x^0_a + \omega}\right) - \log\left(\frac{ax^1_b + \omega}{(1-a)\,x^0_b + \omega}\right).$$

Dado que el modelo no está identificado, también debe ser el caso que $\mathbb{E}_{\theta'}[\Delta c|r > 0] = \mathbb{E}_{\theta^*}[\Delta c|r > 0]$. De la ecuación (A.7), se sigue que:

$$\left(\mathbb{E}_{\theta'}\left[\frac{1}{r}|r > 0\right] - \mathbb{E}_{\theta^*}\left[\frac{1}{r}|r > 0\right]\right)\left(\log \widetilde{R}_a - \log \widetilde{R}_b\right) = 0.$$

Dado que $\widetilde{R}_a \neq \widetilde{R}_b$ por suposición, la expresión anterior implica que $\mathbb{E}_{\theta'}\left[\frac{1}{r}|r > 0\right] = \mathbb{E}_{\theta^*}\left[\frac{1}{r}|r > 0\right]$. Usando el hecho de que r sigue una distribución normal, podemos

escribir $\mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right]$ como:

$$\mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right] = \frac{1}{1-\Phi\left(-\mu'_r/\sigma'_r\right)}\int_0^\infty \frac{1}{x}\left(\frac{1}{\sigma'_r}\right)\phi\left(\frac{x-\mu'_r}{\sigma'_r}\right)$$
$$= \frac{1}{\Phi\left(\mu'_r/\sigma'_r\right)\sigma'_r}\int_0^\infty \frac{1}{z}\phi\left(z-\mu'_r/\sigma'_r\right)dz,$$

donde la segunda línea usa el cambio de variable $z \equiv x/\sigma$ y la simetría de la función de distribución acumulativa normal. Finalmente, para θ^* también debe cumplirse que:

$$\mathbb{E}_{\theta^*}\left[\frac{1}{r}|r>0\right] = \frac{1}{\Phi\left(\mu^*_r/\sigma^*_r\right)\sigma^*_r}\int_0^\infty \frac{1}{z}\phi\left(z-\mu^*_r/\sigma^*_r\right)dz.$$

Dado que $\mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right] = \mathbb{E}_{\theta^*}\left[\frac{1}{r}|r>0\right]$ y $\mu'_r/\sigma'_r = \mu^*_r/\sigma^*_r$, las dos expresiones anteriores implican que $\sigma'_r = \sigma^*_r$, llegando a una contradicción. Por lo tanto, θ_r está identificado.

El siguiente paso es demostrar que $\theta_\delta \equiv (\mu_\delta, \sigma_\delta, \rho)$ está identificado. Fijemos $\theta_r \equiv (\mu_r, \sigma_r)$ y supongamos, por el contrario, que θ_δ no está identificado: existen θ'_δ y θ^*_δ en Θ tales que $\theta'_\delta \neq \theta^*_\delta$ y $\mathcal{P}_{\theta'_\delta}(a=0, A_\mathcal{C}) = \mathcal{P}_{\theta^*_\delta}(a=0, A_\mathcal{C})$. Según la Proposición 5, esta igualdad implica: $\mathcal{P}_{\theta'_\delta}(a=0, A_\mathcal{C}|r\leq 0) = \mathcal{P}_{\theta'_\delta}(K(0, A_\mathcal{C}|r)\leq\delta|r\leq 0)\,\mathcal{P}_{\theta_r}(r\leq 0)$, que es igual a $\mathcal{P}_{\theta^*_\delta}(K(0, A_\mathcal{C}|r)\leq\delta|r\leq 0)\,\mathcal{P}_{\theta_r}(r\leq 0)$. Esta igualdad a su vez implica:

$$\int_{-\infty}^0 \left\{\Phi\left(\frac{K\left(0, A_\mathcal{C}|r\right)-\mu'_{\delta|r}}{\sigma'_{\delta|r}}\right) - \Phi\left(\frac{K\left(0, A_\mathcal{C}|r\right)-\mu^*_{\delta|r}}{\sigma^*_{\delta|r}}\right)\right\} dr = 0,$$

En este punto, la identificación de θ_δ a partir de menús convexos es análoga a su identificación usando menús temporales. Por lo tanto, podemos usar la suposición (b) y los mismos pasos utilizados en la segunda parte de la demostración de la Proposición 3 para llegar a una contradicción. ■

Demostración de la Proposición 7: Comenzamos discutiendo menús en los que $t^1 \to t^0$, que pueden interpretarse como problemas de riesgo únicamente. En estos menús, la elección está determinada por la optimización de la función objetivo $p^0\frac{(\omega+(1-a)x^0)^{1-r}-\omega^{1-r}}{1-r} + p^1\frac{(\omega+ax^1)^{1-r}-\omega^{1-r}}{1-r}$. Siempre que $r\leq 0$, hay dos casos. Primero, si $p^1x^1 \geq p^0x^0$, la elección es $a=1$. Segundo, si $p^1x^1 < p^0x^0$, existe $r^* < 0$ tal que siempre que $r \leq r^*$ la elección es $a=1$ y siempre que $r > r^*$ la elección es $a=0$. Ahora, cuando $r > 0$, la solución es interior y las elecciones forman un mapeo

continuo desde la esquina con mayor esperanza hacia el punto $\bar{e} = \frac{x^0}{x^0+x^1}$, que es el límite de las elecciones cuando $r \to \infty$. Por lo tanto, podemos considerar uno de estos problemas, digamos, uno tal que $p^1x^1 > p^0x^0$. La masa observada de la alternativa $a = 1$, que denotamos por q_1, debe ser igual a $\Phi(\frac{0-\mu_r}{\sigma_r})$. Ahora, sea a cualquier valor en $(\bar{e}, 1)$ y denote por r_a el valor por encima del cual la elección óptima cae por debajo de a. Denote por q' la probabilidad de elección observada por debajo de a, que debe ser igual a $1 - \Phi(\frac{r_a-\mu_r}{\sigma_r})$. Esto permite obtener los parámetros μ_r y σ_r: $\sigma_r = \frac{r_a}{1-\Phi^{-1}(q_1)-\Phi^{-1}(q')}$ y $\mu_r = -\sigma_r\Phi^{-1}(q_1) = -\frac{r_a\Phi^{-1}(q_1)}{1-\Phi^{-1}(q_1)-\Phi^{-1}(q')}$.

El resto de los parámetros pueden identificarse de la siguiente manera. Como se comentó en la demostración de la Proposición 4, para cualquier ω dado, el uso de pagos grandes es equivalente al caso $\omega \to 0$ y, en lo que sigue, asumimos pagos grandes. Siempre que $p_1 \to p_0$, la probabilidad de seleccionar opciones por debajo de $\frac{1}{2}$ es la probabilidad de que δ esté por encima de $K(A_\mathcal{T})(1 - r)$.[30] Se puede reproducir entonces la demostración de la Proposición 4 reemplazando la masa de $0_\mathcal{T}$ por la masa acumulada por debajo de $\frac{1}{2}$. ■

Appendix B. Menús Híbridos

En un menú híbrido, cada una de las dos alternativas corresponde a una lotería contingente a dos estados, con la lotería más segura otorgada en un momento anterior en el tiempo. Formalmente, $A_\mathcal{H} = \{0_\mathcal{H}, 1_\mathcal{H}\}$ con $0_\mathcal{H} = ([p, 1-p; x_1^0, x_2^0], t^0)$ y $1_\mathcal{H} = ([p, 1-p; x_1^1, x_2^1], t^1)$ tales que $x_1^1 > x_1^0 > x_2^0 > x_2^1$, $p \in (0, 1)$ y $t^0 < t^1$. El análisis de los menús híbridos es análogo al de los menús temporales al condicionar nuevamente sobre el parámetro r. Para cualquier menú híbrido dado $A_\mathcal{H}$ y cualquier valor de r, existe una constante dependiente del menú $K(A_\mathcal{H}|r) \in \mathbb{R}$ tal que se selecciona la alternativa $0_\mathcal{H}$ si y solo si $\delta \geq K(A_\mathcal{H}|r)$.[31] Como resultado, la probabilidad de elección de la alternativa $0_\mathcal{H}$ es:

$$\mathcal{P}_f(0_\mathcal{H}, A_\mathcal{H}) = 1 - \int_r F_{\delta|r}K(A_\mathcal{H}|r)f^r(r)\mathrm{d}r.$$

El efecto de desplazamientos y expansiones de δ se entiende trivialmente a partir de esta estructura, aplicando la lógica de la Proposición 2. Entender el efecto de r

[30] Nótese que, a diferencia del caso de la Proposición 4, la probabilidad obtenida aquí es exacta.

[31] Es fácil ver que $K(A_\mathcal{H}|r) = \frac{1}{t^1-t^0}\log\left[\frac{pu_r(x_1^0)+(1-p)u_r(x_2^0)}{pu_r(x_1^1)+(1-p)u_r(x_2^1)}\right]$.

requiere cierta precaución, ya que la razón de las utilidades esperadas es un objeto que puede ser difícil de manejar. Afortunadamente, puede verse que para familias estándar de utilidades monetarias (por ejemplo, CRRA y CARA), el umbral $K(A_{\mathcal{H}}|r)$ es decreciente en r siempre que $\delta > 0$.

Appendix C. Evaluación Numérica de las Probabilidades de Elección

DMPL. Dado un valor para el vector de parámetros $\theta \in \Theta$, el cálculo de la log-verosimilitud definida en la Sección 4 requiere computar $\mathcal{P}_\theta(0_m, A_m)$ para cada menú A_m en el conjunto de datos $\mathcal{O}$. En general, esto requiere evaluar un integral doble de forma numérica. Dada la estructura del modelo RDEU, esto puede hacerse de manera eficiente utilizando métodos de Cuasi-Monte Carlo (QMC): comience por discretizar el soporte de f_θ en N_{QMC} nodos $\{r_k, \delta_k\}_{k=1}^{N_{QMC}}$ utilizando secuencias de baja discrepancia. Sea $\mathcal{I}(0_m, A_m|r_k, \delta_k)$ una función indicadora que toma el valor uno cuando $DEU_{r_k,\delta_k}(0_m) > DEU_{r_k,\delta_k}(1_m)$, y cero en caso contrario. Para N_{QMC} suficientemente grande, tenemos:

$$\mathcal{P}_\theta(0_m, A_m) \approx \frac{V}{N_{QMC}} \sum_{k=1}^{N_{QMC}} \mathcal{I}(0_m, A_m|r_k, \delta_k) f_\theta(r_k, \delta_k),$$

donde $V \equiv \int_r \int_\delta dr d\delta = (\overline{r} - \underline{r})\left(\overline{\delta} - \underline{\delta}\right)$ es una constante de normalización. Podemos controlar la precisión de la aproximación aumentando el número de nodos N_{QMC}. Es importante destacar que la función indicadora $\mathcal{I}(0_m, A_m|r_k, \delta_k)$ es independiente de θ. Por lo tanto, para calcular el estimador de máxima verosimilitud de θ, esta función indicadora solo necesita calcularse una vez antes de iniciar la búsqueda del máximo, reduciendo drásticamente el tiempo de estimación.

También podemos usar los resultados del artículo, junto con el supuesto de que f sigue una distribución normal bivariada, para mejorar aún más el algoritmo de estimación. Para los menús de riesgo, la Proposición 1 implica que $\mathcal{P}_\theta(0_m, A_{\mathcal{R},m}) = 1 - \Phi((K(A_{\mathcal{R},m}) - \mu_r)/\sigma_r)$. El cálculo de $K(A_{\mathcal{R},m})$ requiere resolver una ecuación no lineal de forma numérica para cada menú de riesgo. Sin embargo, estos umbrales son independientes de θ, por lo que solo necesitan calcularse una vez antes de la estimación. Para los menús temporales, la Proposición 2 simplifica el integral doble

que caracteriza $\mathcal{P}_\theta\left(0_m, A_{\mathcal{T},m}\right)$ en el siguiente integral de una sola variable:

$$\mathcal{P}_\theta\left(0_m, A_{\mathcal{T},m}\right) = 1 - \int_r \Phi\left(\frac{K\left(A_{\mathcal{T},m}|r\right) - \mu_{\delta|r}}{\sigma_{\delta|r}}\right) \phi\left(\frac{r-\mu_r}{\sigma_r}\right) dr,$$

con $\mu_{\delta|r} \equiv \mu_\delta + \rho\frac{\sigma_\delta}{\sigma_r}\left(r-\mu_r\right)$ y $\sigma_{\delta|r} \equiv \sigma_\delta\sqrt{1-\rho^2}$. Este integral simplificada también puede evaluarse numéricamente utilizando métodos QMC, como se discutió anteriormente.

Presupuestos Convexos. El cálculo de la función de log-verosimilitud con presupuestos convexos, como se define en la Sección 6, es computacionalmente más exigente ya que requiere evaluar $\mathcal{P}_\theta\left(a \in \alpha^s, A_m\right)$ para cada uno de los M menús en el conjunto de datos y cada una de las S opciones en las que se discretiza el conjunto de elección. Al igual que en el modelo RUM aditivo iid, podemos proceder redondeando cada asignación observada a al punto medio de la opción α^s para la cual $a \in \alpha^s$. Esto da como resultado S posibles asignaciones observadas en los datos: $\overline{\alpha}_1 = 0$, $\overline{\alpha}_2 = \left(a_2 + a_3\right)/2$, ..., $\overline{\alpha}_{S-1} = \left(a_{S-1} + a_S\right)/2$ y $\overline{\alpha}_S = 1$. Sea $\mathcal{I}\left(\overline{\alpha}, A_m|r_k, \delta_k\right)$ una función indicadora que toma el valor de 1 cuando $DEU_{r_k,\delta_k}\left(\overline{\alpha}\right) \geq DEU_{r_k,\delta_k}\left(\overline{\alpha}^*\right)$ en el menú A_m para todo $\overline{\alpha}^* \in \{\overline{\alpha}_1, \ldots, \overline{\alpha}_S\}$, y cero en caso contrario. Podemos entonces usar la aproximación numérica $\mathcal{P}_\theta\left(a \in \alpha^s, A_m\right) \approx \left(V/N_{QMC}\right)\sum_{k=1}^{N_{QMC}} \mathcal{I}\left(\overline{\alpha}, A_m|r_k, \delta_k\right) f_\theta\left(r_k, \delta_k\right)$. Al igual que en el caso de DMPL, la función indicadora $\mathcal{I}\left(\overline{\alpha}, A_m|r_k, \delta_k\right)$ es independiente de θ y puede precomputarse antes de maximizar la función de log-verosimilitud. Este método permite especificaciones flexibles de f y puede extenderse fácilmente a modelos más generales, como se ilustra en el Apéndice D.

Alternativamente, podemos aprovechar los resultados de la Proposición 5 y el supuesto de que f es normal para calcular $\mathcal{P}_\theta\left(\left[a_s, a_{s+1}\right], A_m\right)$ como el siguiente integral de variable única:

$$\int_r \left[\Phi\left(\frac{K\left(A_{s,\mathcal{C}}|r\right) - \mu_{\delta|r}}{\sigma_{\delta|r}}\right) - \Phi\left(\frac{K\left(A_{s+1,\mathcal{C}}|r\right) - \mu_{\delta|r}}{\sigma_{\delta|r}}\right)\right] \phi\left(\frac{r-\mu_r}{\sigma_r}\right) dr,$$

donde los umbrales $K\left(A_{s,\mathcal{C}}|r\right)$ se definen en la Sección 5 y son independientes de θ.

Appendix D. Extensiones

Los métodos introducidos en este trabajo pueden extenderse para incluir características de comportamiento adicionales y distribuciones alternativas para los

parámetros aleatorios. En esta sección, ilustramos estas extensiones utilizando datos de presupuestos convexos provenientes del diseño experimental de Andreoni y Sprenger (2012a). Este conjunto de datos es similar al utilizado en la Sección 5 y presenta 97 sujetos enfrentando 45 menús convexos con pagos ciertos. La principal diferencia con los datos utilizados en el análisis hecho en el texto principal es que algunos menús incluyen pagos en el presente, lo que permite estimar sesgo hacia el presente en el descuento.

La Tabla 6 resume los resultados de este ejercicio. Para facilitar la comparación entre modelos con diferentes supuestos de distribución, se reportan la mediana y el rango intercuartil (IQR) de las distribuciones estimadas. La segunda columna de la Tabla 6 resume las estimaciones base obtenidas utilizando el modelo RDEU siguiendo el procedimiento descrito en la Sección 6. La curvatura estimada de la función de utilidad es estadísticamente cero, similar a los resultados obtenidos por Andreoni y Sprenger (2012b), lo que podría deberse a la falta de variación en las probabilidades de pago en este conjunto de datos, indicando que la curvatura captura únicamente la sustitución intertemporal. La mediana estimada de la tasa de descuento anual es pequeña y estadísticamente cercana a cero. No obstante, el modelo estima un alto grado de heterogeneidad en ambos parámetros.

La tercera columna de la Tabla 6 muestra los resultados utilizando métodos de Quasi-Monte Carlo (QMC), como se discute en la Sección C de este Apéndice. Los resultados son prácticamente idénticos a los obtenidos utilizando el algoritmo base desarrollado en el trabajo. Sin embargo, estimar el modelo utilizando métodos QMC toma cuatro veces más tiempo, lo que confirma los beneficios de aprovechar la estructura económica del problema. A pesar de ello, el método QMC es útil para estimar el modelo con distribuciones alternativas y características de comportamiento adicionales, como se discute a continuación.

La cuarta columna de la Tabla 6 presenta los resultados de una versión "restringida" del modelo RDE donde δ sigue una distribución normal truncada en cero, descartando la posibilidad de preferencia por el futuro. En este caso, utilizamos una cópula Gaussiana para permitir correlación entre r y δ. La mediana estimada de la tasa de descuento anual es mayor y estadísticamente diferente de cero. La mediana de la curvatura de la función de utilidad permanece cercana a cero, y el IQR de

ambos parámetros es ahora menor. Como era de esperarse, restringir el dominio de δ reduce el ajuste a los datos, como se refleja en el log-verosimilitud resultante.

Las dos últimas columnas de la Tabla 6 muestran los resultados para el modelo extendido que permite sesgo hacia el presente en el descuento, de modo que el factor de descuento en el modelo es $\beta e^{-\delta}$ cuando $t^0 = 0$, y $e^{-\delta}$ en caso contrario. La primera de las dos columnas muestra los resultados para un modelo "sin restricciones" que asume que los parámetros r, δ y β siguen una distribución normal multivariante con una matriz de correlación arbitraria. Los resultados en este caso son muy similares a los obtenidos en el modelo RDEU, lo que indica un bajo grado de sesgo hacia el presente. La segunda columna muestra los resultados para un modelo "restringido" donde r sigue una distribución normal, δ sigue una distribución normal truncada en cero, y β sigue una distribución beta con soporte en el intervalo unitario. Las distribuciones estimadas de r y δ son similares a las obtenidas para el modelo RDEU restringido, pero ahora el sesgo hacia el presente mediano es estadísticamente diferente de 1. No obstante, la mejora en el ajuste es relativamente baja en comparación con el modelo RDEU restringido.

Appendix E. Riqueza inicial

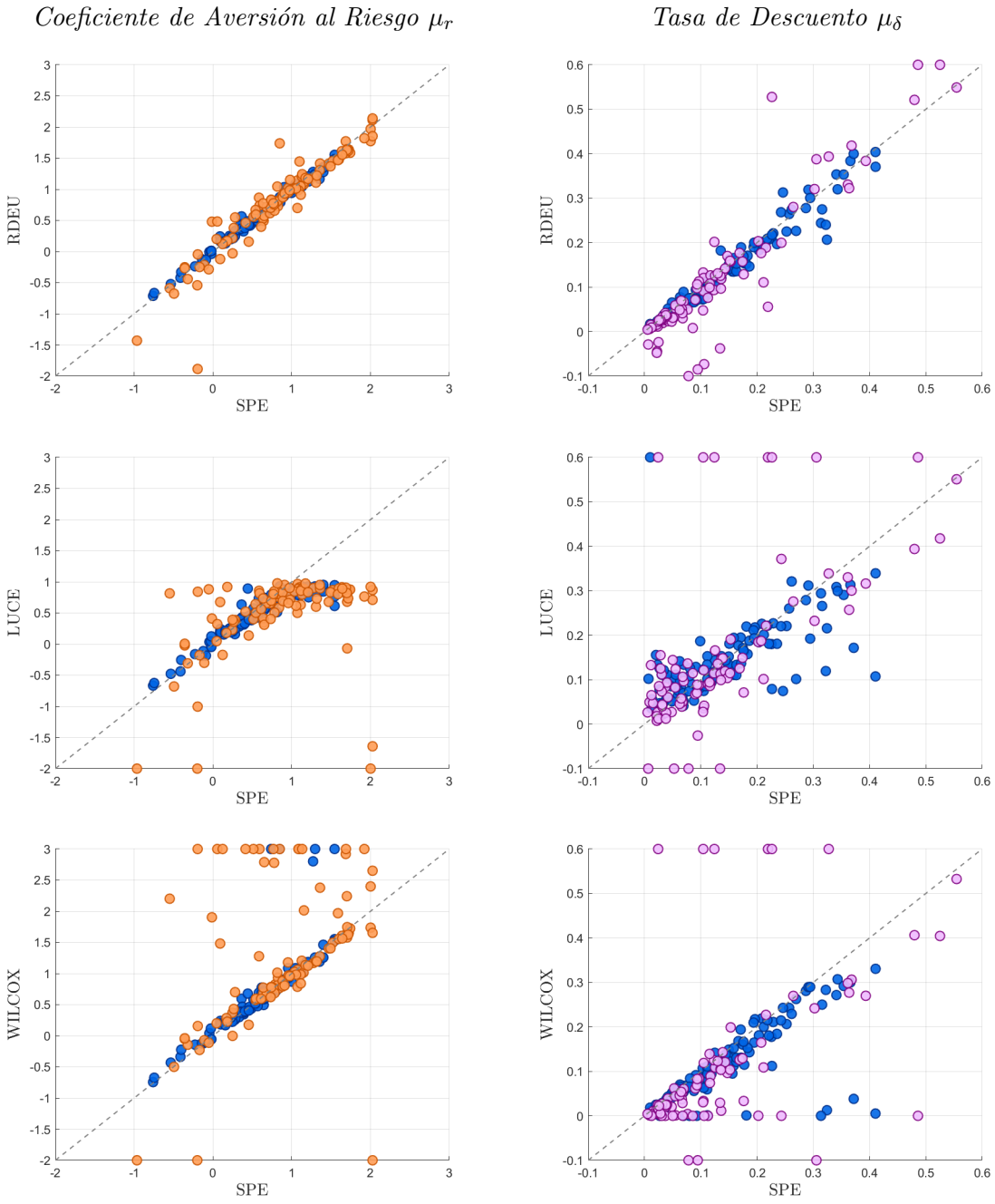
Comentamos ahora brevemente el papel de ω con utilidades monetarias CRRA. Es inmediato ver que, en menús de riesgo $A_{\mathcal{R}}$ tales que $K(A_{\mathcal{R}}) \neq 0$, $K(A_{\mathcal{R}})$ es estrictamente creciente (respectivamente, decreciente) en ω siempre que $K(A_{\mathcal{R}}) > 0$ (resp., $K(A_{\mathcal{R}}) < 0$).[32] En consecuencia, ceteris paribus, la alternativa con mayor valor esperado será elegida con mayor frecuencia. En los menús de tiempo $A_{\mathcal{T}}$, cada umbral $K(A_{\mathcal{T}}|r)$ converge monótonamente a la constante $K(A_{\mathcal{T}}|0)$ a medida que ω aumenta. El comportamiento condicionado de cada r se alinea más con las elecciones condicionadas de $r = 0$. Es decir, ceteris paribus, los individuos más adversos al riesgo (resp., amantes del riesgo) elegirán con mayor frecuencia la opción presente (resp., futura). De manera similar, en menús convexos $A_{\mathcal{C}}$, cada mapa de umbral $K(a, A_{\mathcal{C}}|r)$ converge monótonamente al mapa constante $K(a, A_{\mathcal{C}}|0)$ a medida que ω aumenta. El comportamiento condicionado de cada r se alinea más con las elecciones condicionadas de $r = 0$ (desapareciendo las soluciones interiores).

[32]En el caso degenerado donde los valores esperados de ambas loterías coinciden, obviamente se tiene $K(A_{\mathcal{R}}) = 0$ para todos los niveles de ω.

Dado que en la práctica a menudo se asumen niveles de riqueza inicial nulos, es interesante discutir teóricamente el modelo límite cuando la riqueza inicial tiende a cero. De la discusión anterior, sabemos que este caso límite crearía las mejores condiciones para que el parámetro r influya en todos los menús. Curiosamente, para el caso de menús de tiempo $A_{\mathcal{T}}$, como se discutió en la demostración de la Proposición 3, el mapa de umbral condicionado se vuelve lineal por tramos: (a) cuando $r < 1$, $K(A_{\mathcal{T}}|r) = \frac{\log \frac{x^1}{x^0}}{t^1 - t^0}(1-r) \equiv K(A_{\mathcal{T}})(1-r) > 0$, donde $K(A_{\mathcal{T}})$ es una constante dependiente del menú, y (b) cuando $r \geq 1$, $K(A_{\mathcal{T}}|r)$ se vuelve nulo. Por lo tanto, para parámetros (r, δ), con $r < 1$, la opción más temprana $0_{\mathcal{T}}$ es preferida a la opción más tardía si y solo si $\frac{\delta}{1-r} \geq K(A_{\mathcal{T}})$. Es decir, la expresión $\frac{\delta}{1-r}$ representa una corrección simple de δ basada en el parámetro de riesgo r que captura completamente las consideraciones temporales. En otras palabras, el comportamiento de $DEU_{r,\delta}$ es equivalente al comportamiento de $DEU_{0,\frac{\delta}{1-r}}$, y si el analista está dispuesto a considerar que la aversión al riesgo por encima de 1 no es crucial o que la aversión al riesgo y la aversión al retraso son fenómenos algo independientes para valores estándar, se pueden considerar distribuciones independientes de r y $\frac{\delta}{1-r}$. Es importante destacar que el comportamiento para $r \geq 1$ se vuelve extremo cuando la riqueza es despreciable, ya que la alternativa $0_{\mathcal{T}}$ siempre es preferida.

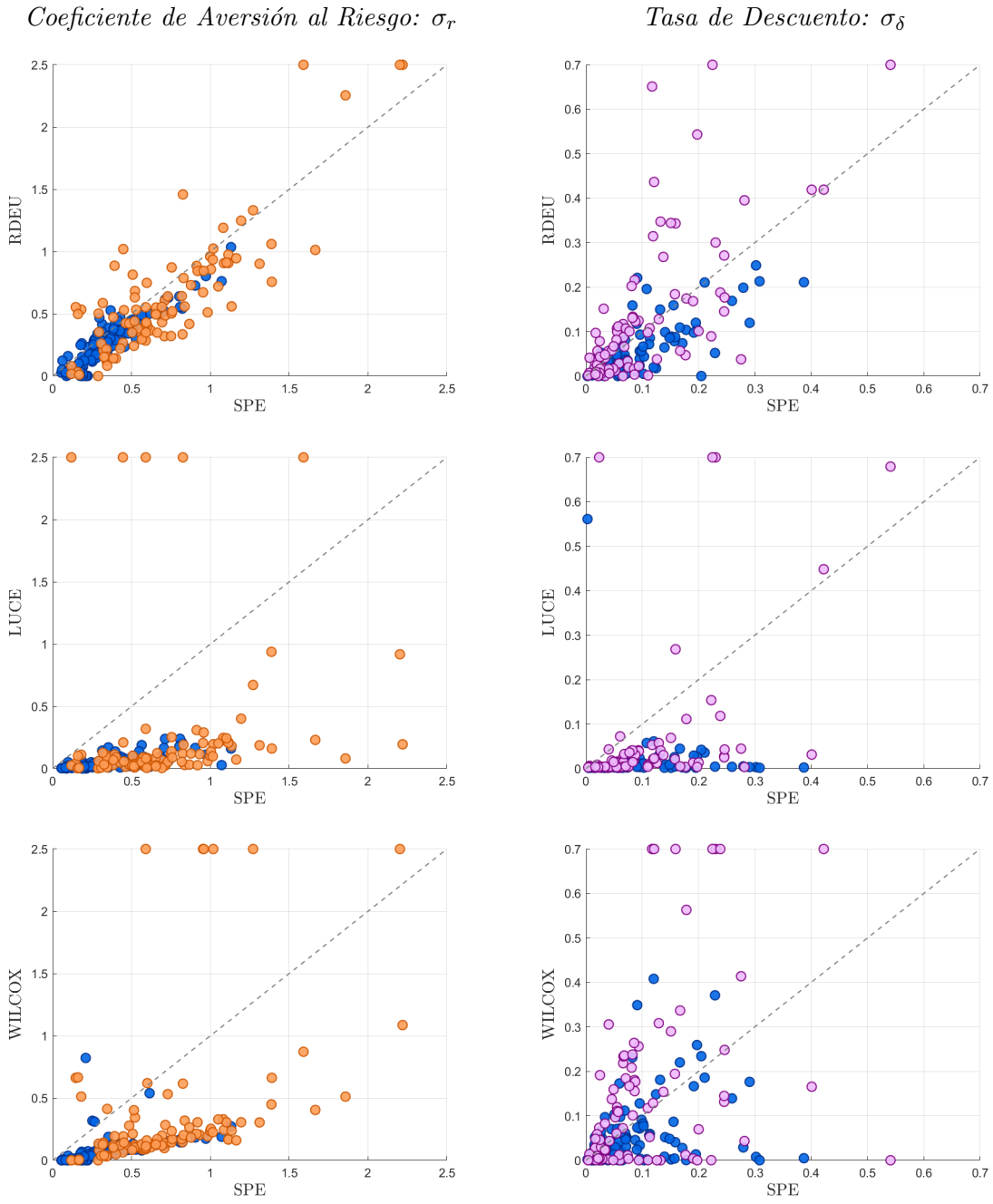
Para observar el papel de la riqueza inicial en las aplicaciones empíricas estudiadas en este trabajo, la Tabla 5 compara las estimaciones bajo la elección base de ω con aquellas obtenidas al fijar ω a un valor positivo cercano a cero. Este ejercicio confirma la discusión teórica previa: el promedio estimado y la desviación estándar de la aversión al riesgo disminuyen con el valor de ω, y la correlación entre r y δ aumenta. No obstante, estos cambios son cuantitativamente pequeños. Nótese también que la distribución marginal estimada de δ permanece prácticamente inalterada entre los valores de ω.

FIGURE 1. Preferencias por Riesgo y Tiempo Individuales: AHLR



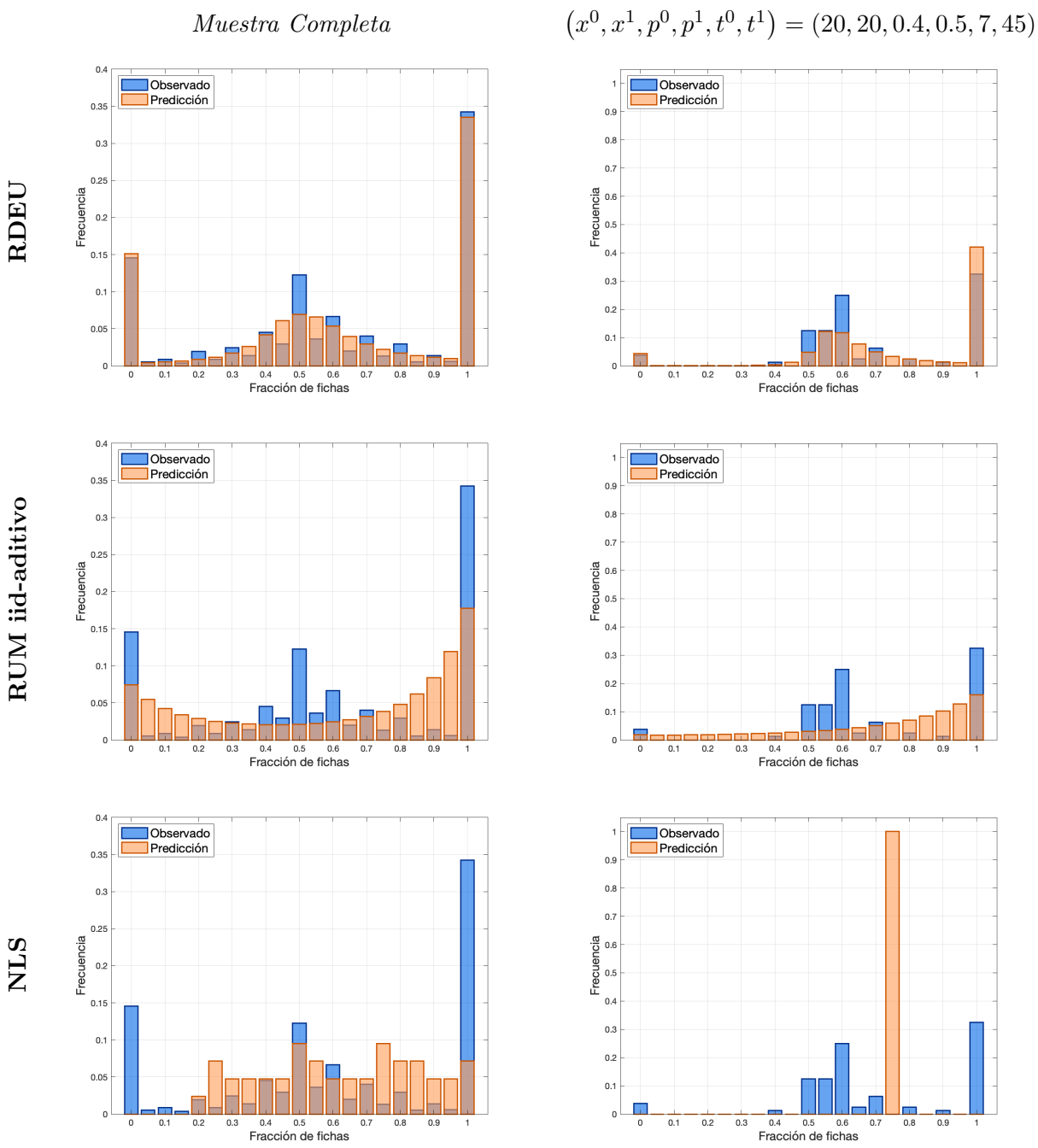
NOTAS. La figura muestra el coeficiente promedio estimado de aversión al riesgo μ_r (primera columna) y la tasa de descuento anual μ_δ (segunda columna) para cada individuo en los datos de las listas de precios dobles-múltiples de Andersen et al. (2008). Cada punto representa la estimación de un sujeto utilizando el modelo estructural correspondiente y la compara con la estimación semiparamétrica basada en los menús adyacentes en cada tarea de riesgo/tiempo donde la elección del sujeto cambió de la lotería segura/temprana a la lotería arriesgada/tardía. Los sujetos que no cambiaron de elección en al menos una de las cuatro tareas de riesgo se muestran en naranja. Los sujetos que no cambiaron de elección en al menos una de las seis tareas de tiempo se muestran en morado. Las estimaciones están truncadas para ajustarse a los rangos en los gráficos.

FIGURE 2. Volatilidad de las Preferencias por Riesgo y Tiempo Individuales: AHLR



NOTAS. La figura muestra la desviación estándar estimada del coeficiente de aversión al riesgo σ_r (primera columna) y la tasa de descuento anual σ_δ (segunda columna) para cada individuo en los datos de listas de precios dobles múltiples de Andersen et al. (2008). Cada punto muestra la estimación de un sujeto utilizando el modelo estructural correspondiente y la compara con la estimación semi-paramétrica basada en los menús adyacentes en cada tarea de riesgo/tiempo donde la elección del sujeto cambió de la lotería segura/temprana a la lotería riesgosa/demorada. Los sujetos que no cambiaron sus elecciones en al menos una de las cuatro tareas de riesgo se muestran en naranja. Los sujetos que no cambiaron sus elecciones en al menos una de las seis tareas de tiempo se muestran en morado. Las estimaciones están truncadas para ajustarse a los rangos en las gráficas.

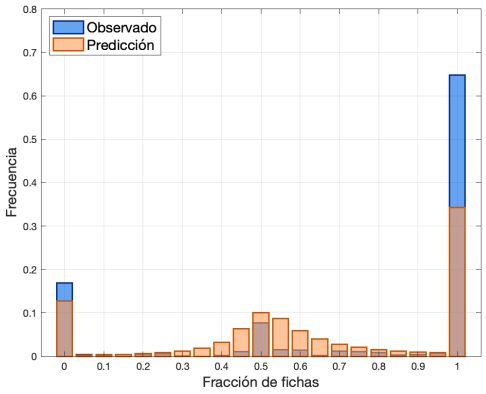
FIGURE 3. Elección Obervada vs. Predicción: AS



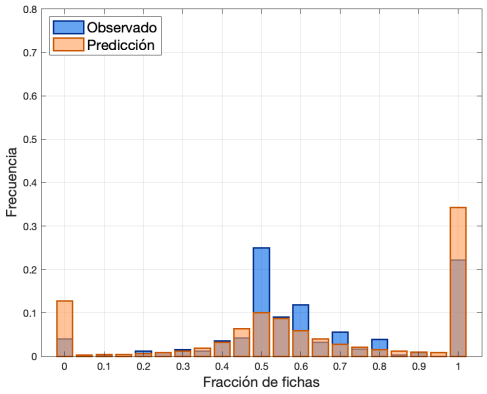
NOTAS. La figura muestra la frecuencia observada y predicha de elegir cada proporción de fichas a en el conjunto de datos de menús convexos de Andreoni y Sprenger (2012b). Las barras azules muestran la frecuencia observada en los datos. Las barras naranjas muestran la frecuencia predicha por el modelo correspondiente utilizando las estimaciones poblacionales reportadas en el Cuadro 3. La columna izquierda muestra los resultados para toda la muestra, mientras que la columna derecha muestra los resultados para un menú convexo con pagos de 20 USD entregados en 7 y 35 días con probabilidades de 0.4 y 0.5, respectivamente.

FIGURE 4. Elección Observada vs. Predicción por Tratamiento: AS

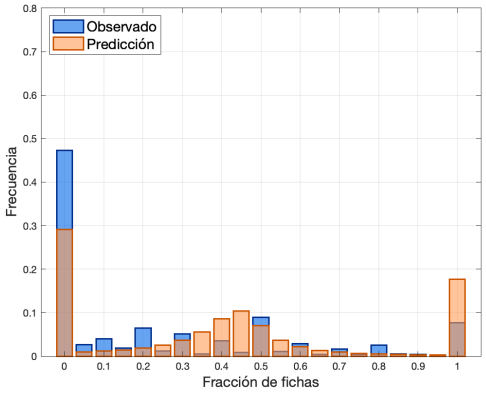
Tratamiento 1: $(p_t, p_{t+k}) = (1, 1)$



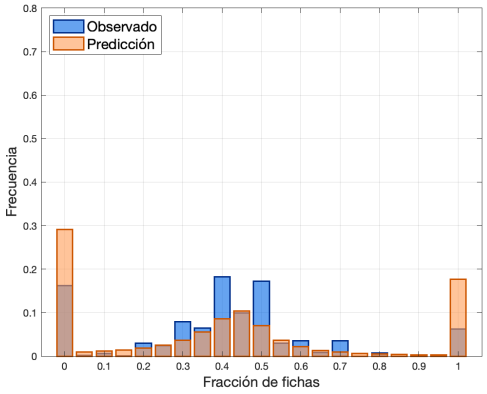
Tratamiento 4: $(p_t, p_{t+k}) = (0.5, 0.5)$



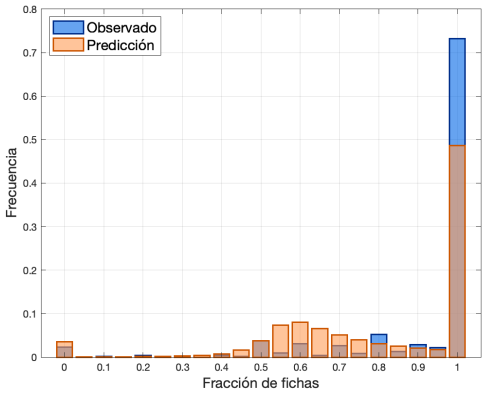
Tratamiento 2: $(p_t, p_{t+k}) = (1, 0.8)$



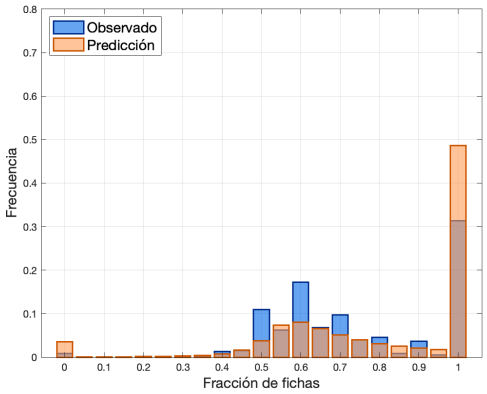
Tratamiento 5: $(p_t, p_{t+k}) = (0.5, 0.4)$



Tratamiento 3: $(p_t, p_{t+k}) = (0.8, 1)$



Tratamiento 6: $(p_t, p_{t+k}) = (0.4, 0.5)$



NOTAS. La figura muestra la frecuencia observada y predicha de elegir cada proporción de fichas a en el conjunto de datos de menús convexos de Andreoni y Sprenger (2012b). Las barras azules muestran la frecuencia observada en los datos. Las barras naranjas muestran la frecuencia predicha por el modelo RDEU utilizando las estimaciones poblacionales reportadas en la Tabla 3.

TABLE 1. Preferencias por Riesgo y Tiempo Agregadas: AHLR

	RDEU	**LUCE**	**WILCOX**	**SPE**
μ_r	0.781	0.726	0.777	0.715
	[0.053]	[0.058]	[0.045]	–
σ_r	0.895	0.086	0.220	0.833
	[0.049]	[0.016]	[0.011]	–
μ_δ	0.125	0.101	0.095	0.138
	[0.008]	[0.009]	[0.007]	–
σ_δ	0.125	0.020	0.222	0.165
	[0.010]	[0.006]	[0.017]	–
ρ	−0.958	–	–	−0.761
	[0.016]	–	–	–
Log-Verosim. $\mathcal{L}$				
Menús de Riesgo	−0.450	−0.448	−0.445	–
Menús de Tiempo	−0.495	−0.554	−0.557	–
Todos los Menús	−0.481	−0.521	−0.522	–

NOTAS. La tabla reporta las tasas estimadas de aversión al riesgo y de descuento a nivel agregado utilizando diferentes modelos y datos del diseño de listas de precios doble múltiple en Andersen et al. (2008). Las primeras tres columnas muestran las estimaciones de máxima verosimilitud de los tres modelos estructurales descritos en el texto principal. La última columna muestra la media y la desviación estándar agregadas de las estimaciones semiparamétricas de r y δ obtenidas a partir de los menús adyacentes en cada tarea de riesgo/tiempo donde la elección de cada individuo en la muestra cambió de la lotería segura/temprana a la lotería arriesgada/demorada. Los errores estándar para cada estimación de máxima verosimilitud se muestran entre corchetes y están agrupados a nivel individual.

TABLE 2. Preferencias por Riesgo y Tiempo Individuales: AHLR

	Coeficiente de Aversión al Riesgo μ_r				**Tasa de Descuento μ_δ**			
Momento	**RDEU**	**LUCE**	**WILCOX**	**SPE**	**RDEU**	**LUCE**	**WILCOX**	**SPE**
Media	0.698	−4.875	−0.305	0.715	0.133	1.525	2.634	0.138
Desv. Est.	0.629	45.04	19.59	0.608	0.167	34.412	20.31	0.111
Min	−1.886	−458.4	−276.7	−0.964	−0.443	−315.6	−34.03	0.006
Pctl. 10	−0.122	−0.052	−0.054	−0.064	0.018	0.040	0.001	0.027
Pctl. 25	0.366	0.333	0.389	0.285	0.042	0.074	0.024	0.055
Mediana	0.708	0.656	0.782	0.714	0.098	0.115	0.067	0.110
Pctl. 75	1.125	0.842	1.271	1.128	0.182	0.184	0.152	0.194
Pctl. 90	1.507	0.894	2.782	1.557	0.315	0.303	0.273	0.308
Max	2.139	0.977	8.123	2.026	1.658	344.4	153.9	0.555
Correlación con SPE								
Pearson r	0.961	0.079	−0.111	1	0.814	0.029	0.023	1
Kendall τ	0.899	0.597	0.690	1	0.836	0.544	0.591	1
Spearman ρ	0.980	0.728	0.756	1	0.943	0.702	0.686	1

NOTAS. La tabla reporta estadísticas descriptivas de las tasas promedio estimadas de aversión al riesgo (μ_r) y de descuento (μ_δ) entre individuos utilizando datos del diseño de listas de precios doble múltiple de Andersen et al. (2008). Cada columna corresponde a un modelo descrito en el texto principal. Las últimas tres filas reportan, respectivamente, el coeficiente de correlación de Pearson, el coeficiente de correlación de rango de Kendall y el coeficiente de correlación de rango de Spearman entre las estimaciones de los sujetos utilizando un modelo estructural y las estimaciones semiparamétricas obtenidas de los menús adyacentes en cada tarea de riesgo/tiempo donde la elección del individuo cambió de la lotería segura/temprana a la lotería arriesgada/demorada.

TABLE 3. Preferencias por Riesgo y Tiempo Agregadas: AS

	RDEU	**RUM iid-aditivo**	**NLS**
μ_r	0.207	−0.133	0.317
	[0.062]	[0.020]	[0.017]
σ_r	0.752	–	–
	[0.079]	–	–
μ_δ	0.339	0.571	0.262
	[0.108]	[0.081]	[0.079]
σ_δ	1.805	–	–
	[0.124]	–	–
ρ	−0.164	–	–
	[0.053]	–	–
$\mathcal{L}$	−2.108	−2.519	–

NOTAS. La tabla reporta las estimaciones de máxima verosimilitud de la aversión al riesgo y el descuento a nivel agregado utilizando datos de menús convexos del diseño experimental de Andreoni y Sprenger (2012). Cada columna reporta las estimaciones para el modelo estructural correspondiente discutido en el texto principal. Los errores estándar de todas las estimaciones, mostrados entre paréntesis, están agrupados a nivel individual.

TABLE 4. Volatilidad de las Preferencias por Riesgo y Tiempo Individuales: AHLR

	Coeficiente de Aversión al Riesgo σ_r				**Tasa de Descuento σ_δ**			
Momento	**RDEU**	**LUCE**	**WILCOX**	**SPE**	**RDEU**	**LUCE**	**WILCOX**	**SPE**
Media	0.473	19.56	5021	0.539	0.099	0.510	19802	0.091
Desv. Est.	0.568	208.1	70187	0.374	0.247	6.373	139319	0.086
Min	0.001	0.001	0.001	0.056	0.001	0.002	0.001	0.003
Pctl. 10	0.050	0.006	0.002	0.172	0.005	0.002	0.001	0.017
Pctl. 25	0.190	0.021	0.055	0.273	0.015	0.003	0.005	0.029
Mediana	0.362	0.051	0.104	0.446	0.042	0.004	0.041	0.061
Pctl. 75	0.554	0.112	0.189	0.710	0.107	0.017	0.120	0.120
Pctl. 90	0.903	0.211	0.407	1.028	0.211	0.042	0.241	0.207
Max	5.963	2895	1e6	2.219	3.127	90.59	1e6	0.541
Correlación con SPE								
Pearson r	0.805	−0.055	0.014	1	0.591	0.126	0.136	1
Kendall τ	0.681	0.471	0.635	1	0.520	0.366	0.310	1
Spearman ρ	0.850	0.644	0.778	1	0.702	0.504	0.421	1

NOTAS. La tabla reporta las estimaciones de máxima verosimilitud de la aversión al riesgo y el descuento a nivel agregado utilizando datos de menús convexos del diseño experimental de Andreoni y Sprenger (2012). Cada columna reporta las estimaciones para el modelo estructural correspondiente discutido en el texto principal. Los errores estándar de todas las estimaciones, mostrados entre paréntesis, están agrupados a nivel individual.

TABLE 5. Preferencias por Riesgo y Tiempo Estimadas cuando $\omega \approx 0$

	DMPL-AHLR		**CB-AS**	
	Base	$\omega \approx 0$	*Base*	$\omega \approx 0$
μ_r	0.781	0.681	0.207	0.095
	[0.053]	[0.032]	[0.062]	[0.045]
σ_r	0.895	0.768	0.752	0.562
	[0.049]	[0.039]	[0.079]	[0.054]
μ_δ	0.125	0.102	0.339	0.383
	[0.008]	[0.007]	[0.108]	[0.110]
σ_δ	0.125	0.116	1.805	1.821
	[0.010]	[0.007]	[0.124]	[0.125]
ρ	-0.958	-0.999	-0.164	-0.202
	[0.016]	[0.001]	[0.053]	[0.052]

NOTAS.- La tabla reporta el coeficiente de aversión al riesgo y la tasa de descuento anual a nivel poblacional estimados por el modelo RDEU bajo dos supuestos diferentes sobre el valor de la riqueza integrada ω.

TABLE 6. Preferencias por Riesgo y Tiempo Estimadas en Andreoni y Sprenger (2012a)

	RDEU			**Descuento Hiperbólico**	
	Base	*QMC*	*Restringido*	*Sin Restricciones*	*Restringido*
Median (r)	−0.051 [0.084]	−0.050 [0.080]	−0.037 [0.039]	−0.039 [0.071]	−0.038 [0.038]
IQR (r)	0.718 [0.166]	0.715 [0.152]	0.347 [0.072]	0.665 [0.122]	0.348 [0.070]
Median (δ)	0.043 [0.157]	0.037 [0.151]	0.504 [0.058]	0.072 [0.148]	0.490 [0.052]
IQR (δ)	2.317 [0.335]	2.310 [0.305]	0.784 [0.087]	2.217 [0.271]	0.764 [0.079]
Median (β)	–	–	–	0.991 [0.004]	0.898 [0.008]
IQR (β)	–	–	–	0.004 [0.005]	0.247 [0.056]
Cor (r, δ)	−0.306 [0.071]	−0.302 [0.068]	0.020 [0.056]	−0.284 [0.067]	0.017 [0.060]
Cor (r, β)	–	–	–	−0.002 [2.605]	−0.078 [0.459]
Cor (δ, β)	–	–	–	−0.001 [0.479]	0.134 [0.356]
$\mathcal{L}$	−1.709	−1.709	−1.835	−1.709	−1.833

NOTAS. La tabla reporta los momentos estimados de las distribuciones de aversión al riesgo (r), descuento (δ) y sesgo presente (β) estimados a nivel poblacional. La segunda columna reporta los resultados obtenidos utilizando el modelo RDEU con la misma metodología y supuestos que en la Sección 6. La tercera columna muestra los resultados utilizando métodos de Quasi-Monte Carlo (QMC), como se discute en el Apéndice C. La cuarta columna muestra los resultados utilizando QMC y asumiendo que δ sigue una distribución normal truncada en cero. La cuarta columna muestra los resultados para un modelo extendido que permite sesgo presente bajo el supuesto de que (r, δ, β) siguen una distribución normal multivariada. La última columna muestra los resultados asumiendo que δ sigue una distribución normal truncada en cero y β una distribución beta. Los errores estándar para cada MLE se muestran entre corchetes y están agrupados a nivel individual.

TABLE 7. Ejemplo de una Tarea de Riesgo en AHLR

	Lotería A				***Lotería B***			
Menú	p	Pago	p	Prize	p	Pago	p	Pago
1	0.1	2000	0.9	1600	0.1	3850	0.9	100
2	0.1	2000	0.9	1600	0.1	3850	0.9	100
3	0.1	2000	0.9	1600	0.1	3850	0.9	100
4	0.1	2000	0.9	1600	0.1	3850	0.9	100
5	0.1	2000	0.9	1600	0.1	3850	0.9	100
6	0.1	2000	0.9	1600	0.1	3850	0.9	100
7	0.1	2000	0.9	1600	0.1	3850	0.9	100
8	0.1	2000	0.9	1600	0.1	3850	0.9	100
9	0.1	2000	0.9	1600	0.1	3850	0.9	100
10	0.1	2000	0.9	1600	0.1	3850	0.9	100

NOTAS. Ejemplo de una tarea de riesgo en Andersen et al. (2008). Todos los pagos están expresados en DKK.

TABLE 8. Ejemplo de una Tarea de Tiempo en AHLR

Alternativa de Pago	***Opción de Pago A (Paga el siguiente monto en 1 mes)***	***Opción de Pago B (Paga el siguiente monto en 7 meses)***
1	3000	3075
2	3000	3152
3	3000	3229
4	3000	3308
5	3000	3387
6	3000	3467
7	3000	3548
8	3000	3630
9	3000	3713
10	3000	3797

NOTAS. Ejemplo de una tarea de riesgo en Andersen et al. (2008). Todos los premios están expresados en DKK.