

DOCUMENTO DE INVESTIGACIÓN

Empleo, productividad
e innovación

**Actitudes ante el riesgo y
preferencias intertemporales
de los trabajadores peruanos**

Miguel Jaramillo
Kristian López Vargas
Óscar Rodríguez
Diego Tocre

123

Documentos de Investigación 123

Actitudes ante el riesgo y preferencias intertemporales de los trabajadores peruanos

Miguel Jaramillo¹
Kristian López Vargas
Óscar Rodríguez
Diego Tocre

¹ Miguel Jaramillo es investigador principal del Grupo de Análisis para el Desarrollo (GRADE). Kristian López es profesor del Departamento de Economía en la Universidad de California, Santa Cruz. Óscar Rodríguez y Diego Tocre han sido asistentes de investigación en GRADE.

Grupo de Análisis para el Desarrollo (GRADE)
Av. Grau 915, Barranco, Lima 4, Perú
Teléfono: 247-9988
www.grade.org.pe



Esta publicación cuenta con una Licencia Creative Commons Atribución-No-Comercial 4.0 Internacional.

Publicación electrónica. Primera edición. Lima, marzo del 2022

En concordancia con los objetivos de GRADE, el propósito de la serie Documentos de Investigación es difundir oportunamente los estudios que realizan sus investigadores y suscitar el intercambio con otros miembros de la comunidad científica que permita enriquecer el producto final de la investigación, de modo que esta apruebe sólidos criterios técnicos para el proceso político de toma de decisiones. Su publicación requiere de un proceso interno de revisión de pares ciego.

Las opiniones y recomendaciones vertidas en este documento son responsabilidad de sus autores y no representan necesariamente los puntos de vista de GRADE u otros auspiciadores. Los autores declaran que no tienen conflicto de interés vinculado a la realización del presente estudio, sus resultados o la interpretación de estos. Este estudio se basa en datos que se analizaron como parte de una consultoría realizada a Helvetas Perú.

Directora de Investigación: María Balarin
Asistente de edición: Diana Balcázar Tafur
Corrección de estilo: Rocío Moscoso
Diseño de carátula: Elena González
Diagramación: Amaurí Valls / Impresiones y Ediciones Arteta E.I.R.L.

Hecho el Depósito Legal en la Biblioteca Nacional del Perú N° 2022-02386
ISBN: 978-612-4374-49-4

CENDOC / GRADE

JARAMILLO, Miguel; Kristian LÓPEZ VARGAS, Óscar RODRÍGUEZ
y Diego TOCRE

Actitudes ante el riesgo y preferencias intertemporales de los trabajadores peruanos. Lima: GRADE, 2022. (Documentos de Investigación, 123).

TRABAJADORES, RIESGO, MODELOS ECONÓMICOS, TEORÍA
ECONÓMICA, PERÚ

ÍNDICE

PRINCIPALES ABREVIACIONES	7
INTRODUCCIÓN	9
I. TEORÍA	13
1.1. Riesgo: la elección bajo incertidumbre	13
1.2. Elección intertemporal	23
2. EXPERIMENTO	31
2.1. Medición de preferencias	31
2.1.1. <i>Actitudes ante el riesgo</i>	31
2.1.2. <i>Preferencias intertemporales</i>	32
2.1.3. <i>Actitudes ante el riesgo y elecciones intertemporales</i>	34
2.2. Diseño experimental	38
2.2.1. <i>Decisiones intertemporales</i>	38
2.2.2. <i>Decisiones frente al riesgo</i>	41
2.2.3. <i>Otras consideraciones</i>	44
2.3. Procedimiento	45
2.4. Características de la muestra	47
3. MODELOS ESTOCÁSTICOS Y ESTRATEGIAS EMPÍRICAS	51
3.1. Modelos estocásticos	54
3.1.1. <i>Modelos de parámetros aleatorios</i>	56
3.1.2. <i>Modelos de utilidad aleatoria</i>	58
3.1.3. <i>Parámetro de tiemble</i>	63

3.2. Preferencias heterogéneas	64
3.2.1. <i>Según características no observables</i>	64
3.2.2. <i>Según características observables</i>	65
4. RESULTADOS	67
4.1. Estimación de actitudes ante el riesgo	67
4.1.1. <i>Modelos de agente representativo</i>	68
4.1.2. <i>Incorporando heterogeneidad no observable</i>	75
4.1.3. <i>Incorporando heterogeneidad en características observables</i>	79
4.2. Estimados de preferencias intertemporales	81
4.2.1. <i>Modelos de agente representativo</i>	82
4.2.2. <i>Incorporando heterogeneidad no observable</i>	89
4.2.3. <i>Incorporando heterogeneidad en características observables</i>	93
5. Aplicación de nuestros resultados: un ejemplo	97
5.1. Modelo	97
5.1.1. <i>Seguro de desempleo</i>	99
5.1.2. <i>Compensación por tiempo de servicios</i>	100
5.2. Datos	101
5.3. Resultados	102
6. DISCUSIÓN Y CONCLUSIONES	107
REFERENCIAS BIBLIOGRÁFICAS	115
ANEXOS	125
A. Detalles adicionales del procedimiento	125
B. Rondas de práctica	128
C. Implementación de efectos aleatorios	131
D. Resultados adicionales	135
D.1. <i>Actitudes ante el riesgo</i>	135
D.2. <i>Preferencias intertemporales</i>	137

PRINCIPALES ABREVIACIONES

AIC	Criterio de información de Akaike
BDM	Método Becker-DeGroot-Marschak
CRRA	Función de aversión relativa al riesgo constante
CTB	Método Convex Time Budget
CTS	Compensación por tiempo de servicios
DARA	Aversión absoluta al riesgo decreciente
DU	Modelo de utilidad-descontada
DMPL	Double Multiple Price List
EUT	Teoría de la utilidad esperada
FED	Mecanismo <i>front-end-delay</i>
MPL	Multiple Price List
NFED	No recibe el tratamiento <i>front-end-delay</i>
QHD	Descuento cuasihiperbólico
TEA	Tasa de interés efectiva anual

INTRODUCCIÓN

Todo individuo, al tomar decisiones de carácter económico, debe evaluar distintas opciones con resultados inciertos que afectarán su bienestar de manera distinta en más de un periodo. Así, en el ámbito laboral, un trabajador, ante la posibilidad de perder su empleo en el futuro, puede decidir ahorrar parte de sus ingresos regulares con la finalidad de contar con un seguro para esta eventualidad. O bien podría decidir no ahorrar tanto y alcanzar un mayor nivel de consumo durante sus periodos de empleo, con el costo de eventualmente encontrarse sin una reserva de dinero suficiente en caso de perder su empleo. De manera análoga, alguien que busca empleo debe considerar cuánto tiempo esperar por una mejor oferta de trabajo, sacrificando ingresos por un tiempo adicional. Así, si la recompensa de esperar es mayor que el costo de un menor consumo durante el periodo actual, entonces decidirá esperar; o, de estimar lo contrario, podría tomar la oferta actual. Como estos dos ejemplos de decisiones frente a un futuro incierto, hay muchos otros que se pueden aplicar a la vida económica y, en especial, al ámbito laboral. Estos subrayan la importancia de las actitudes ante el riesgo y de las preferencias por consumir ahora o en el futuro, así como la necesidad de entenderlas.

El presente estudio busca aportar en esta dirección de entender mejor las actitudes ante el riesgo y las preferencias por consumir ahora o en el futuro. Específicamente, tiene dos grandes objetivos. El primero es caracterizar a los trabajadores peruanos en cuanto a sus preferencias

ante el riesgo y el consumo intertemporal. Amparados en las teorías estructurales o *determinísticas* de utilidad esperada (EUT) y de descuento cuasihiperbólico (QHD), intentaremos identificar cuán aversos al riesgo e impacientes son los trabajadores peruanos. Para ello, empleamos un diseño experimental que, mediante el uso de incentivos, nos permite estimar los parámetros necesarios para explicar las elecciones de los individuos según las teorías que asumimos. Usando información de varias ciudades a nivel nacional, encontramos que el trabajador peruano tiene un grado de aversión al riesgo de moderado para alto y que descuenta mucho su utilidad futura. Al permitir que las preferencias sean heterogéneas, hallamos que existe un sesgo por el presente. Además, también realizamos un análisis sobre cómo estas preferencias por el riesgo y el consumo intertemporal se relacionan con ciertas características de los individuos, como su disponibilidad de liquidez, edad, género y estatus ocupacional.

El segundo gran objetivo es identificar qué modelos estocásticos son los que mejor se ajustan a la toma de decisiones del trabajador peruano. De hecho, la capacidad predictiva de las teorías determinísticas tiene un límite. Siempre va a existir una variabilidad en las elecciones de los agentes que no se va a poder explicar y, ante la cual, lo mejor que podemos hacer es tratar de comprender cómo se comporta este componente estocástico. Hay estudios que hasta sostienen que elegir el modelo estocástico correcto puede pesar incluso más que elegir el modelo determinístico correcto. Sobre esto, hallamos que el modelo que suele predecir mejor el componente estocástico es el de parámetros aleatorios, que asume que los parámetros son las variables aleatorias que provienen de ciertas distribuciones, y no las utilidades.

Este documento creemos que es valioso, además, porque es un resultado ilustrativo del encuentro entre la Economía teórica y la experimental. Desde el inicio de la Economía como ciencia social, los

pensadores han buscado explicar las decisiones de los agentes de manera determinística; es decir, sobre la base de teorías que caracterizan a un agente representativo que, dados ciertos parámetros o condiciones, siempre debería actuar de la misma manera. La Economía experimental, por otro lado, reconoce que el individuo, por más de verse enfrentado a los mismos incentivos económicos, no siempre toma las mismas decisiones —y más aún si son decisiones complicadas—, lo cual puede deberse a varios sesgos, como la falta de atención, el grado de capacidad cognitiva u otros factores que, a veces, no se observan o no se pueden controlar. De esta constatación surgió la necesidad de que, al explicar la toma de decisiones de las personas, no solo se considere el rol de los incentivos observables o “racionales”, sino también el de aquellos factores que no se observan y que influyen en la decisión del agente. En este documento, a estas teorías las llamamos *determinísticas* y *estocásticas*, respectivamente.

Los resultados de nuestro estudio buscan proveer una mejor base sobre la cual evaluar la viabilidad de las políticas públicas. Usando los parámetros que encontramos, se pueden calcular niveles de bienestar dados ciertos escenarios que involucren varios periodos de tiempo y realizaciones asociadas a probabilidades, como, por ejemplo, la conveniencia de la imposición de un seguro de desempleo sobre el esquema actual de la CTS.

Este estudio se presenta de la siguiente manera. En el capítulo 1, detallamos las teorías determinísticas necesarias para comprender nuestros resultados y métodos de estimación. En particular, primero se relata la teoría sobre elección ante el riesgo y luego la teoría sobre preferencias de consumo intertemporal. En el capítulo 2, se explica cómo se efectúa la medición de preferencias, el diseño experimental que empleamos, el procedimiento y las características de la muestra. El capítulo 3 trata sobre los modelos estocásticos que se han considerado

para este estudio y cómo cada uno de estos se vincula a una estrategia empírica particular. El capítulo 4 muestra los resultados que estimamos según los modelos que consideramos. El capítulo 5 ilustra cómo aplicar nuestros resultados a la discusión de políticas públicas en el mercado laboral. Finalmente, el capítulo 6 discute y concluye.

1. TEORÍA

1.1. Riesgo: la elección bajo incertidumbre

Los matemáticos evalúan el dinero en proporción a su cantidad, mientras que, en la práctica, los hombres con sentido común lo hacen en proporción al uso que pueden obtener de él.

Gabriel Cramer, 1728

El mundo en el que vivimos es intrínsecamente incierto. Casi todas las decisiones que tomamos implican atenerse a cierta variabilidad en sus posibles consecuencias. Las decisiones bajo riesgo –el foco de este capítulo– fueron distinguidas de aquellas bajo incertidumbre por Knight (1921). Para Knight, atenerse a riesgo es tomar decisiones bajo incertidumbre, situación en la que no podemos determinar el resultado de la resolución, pero podemos atribuir, de manera veraz, probabilidad a cada una de las posibles consecuencias. En cambio, la verdadera incertidumbre “no es susceptible a la medición”. En lo que sigue, nos centraremos en estudiar la toma de decisiones entre alternativas riesgosas, en las cuales los resultados son contingentes o dependientes de una distribución de probabilidad objetiva.

Dada su ubicuidad, el estudio de las decisiones bajo incertidumbre –o riesgo– es indispensable para entender –y describir– la economía a través de las preferencias de los actores que la conforman. Por ejemplo, un individuo se puede preguntar: ¿cuánto debo ahorrar

para la jubilación si hay incertidumbre en las rentas futuras, la salud personal, y la solidez de las instituciones políticas y financieras? De la misma manera, debemos cuestionar las decisiones que nos afectan a todos: ¿qué políticas públicas deben regir cuando el futuro de la productividad, la inflación y el desempleo es incierto?

Alrededor de 300 años atrás, Gabriel Cramer y Daniel Bernoulli (1738) dieron el primer paso en desarrollar la base conceptual para responder estas preguntas. La teoría de la utilidad esperada (EUT por sus siglas en inglés) surge por la necesidad de entender la aversión al riesgo. Previa a esta, la norma de evaluación de alternativas riesgosas se basaba en maximizar el valor esperado. Sin embargo, varias anomalías, como la paradoja de San Petersburgo, demostraron que adoptar esta norma conlleva a la contradicción entre la teoría y la praxis. Así, la creación de EUT hizo avanzar nuestra comprensión del comportamiento mediante la mejora del poder descriptivo de las decisiones bajo incertidumbre. Von Neumann y Morgenstern (1944) postularon el axioma de la independencia, un supuesto crucial para la racionalidad en la toma de decisiones bajo incertidumbre, que formalizó la teoría de la utilidad esperada. Pocos años después, con el apoyo de Leonard Savage, Jacob Marschak, Milton Friedman y Paul Samuelson, la utilidad esperada se consagró como la teoría ortodoxa para las decisiones económicas bajo riesgo.²

Desde entonces, varias teorías alternativas de decisión han sido desarrolladas. Notablemente, Savage (1954) axiomatiza una versión de utilidad esperada basada en la probabilidad subjetiva. Similarmente, Kahneman y Tversky (1979) desarrollaron la teoría prospectiva; así, marcaron un hito en el estudio de las preferencias, racionalizando

2 En el histórico Colloque Internationale d'Econometrie de París, en 1952. Para más información sobre los sucesos que llevaron a la aceptación de EUT, véase Moscati (2017).

evidencia contra EUT con hipótesis de la Psicología. Siguiendo otra línea conceptual, una rama de la literatura ha desarrollado modelos estocásticos en los cuales el determinismo inherente a la versión de von Neumann y Morgenstern es sustituido por decisiones afectadas por una variable de utilidad con un componente aleatorio, como fue popularizada en los modelos de utilidad aleatoria de Luce (1959).³

Estas perspectivas han hecho avanzar el entendimiento de las decisiones económicas, aportando nuevas maneras de abordar el riesgo y nuestra relación con él. De aquí en adelante, el capítulo se enfoca en la teoría de la utilidad esperada bajo riesgo como es empleada comúnmente.

La teoría de la utilidad esperada: von-Neumann y Morgenstern

Partiendo de la perspectiva de la probabilidad como concepto fundamental que representa objetivamente la incertidumbre, el propósito de la utilidad esperada es caracterizar las preferencias hacia la toma de decisiones bajo riesgo. Para ello, postula que las preferencias se revelan sobre la base de decisiones entre alternativas riesgosas, las cuales llamamos *loterías*. Estas son el contenido empírico de la teoría, y son definidas por el conjunto de sus posibles resultados $X = [x_1, \dots, x_n]$. Cada lotería X se denota como $(x_1, p_1; \dots; x_n, p_n)$, donde cada p_i se refiere a la probabilidad de que el resultado x_i se realice. En síntesis, la teoría de la utilidad esperada propone que las loterías se evalúan —o deberían ser evaluadas— por la sumatoria de las utilidades esperadas,

$$\sum_{i=1}^n u(x_i)p_i \quad (1.1)$$

3 Esta corriente de pensamiento está endeudada con conceptos y metodologías de medición de estímulos de la literatura en psicofísica, iniciada por Fechner (1860).

donde u es una función que representa las preferencias ante el riesgo del agente –o grupo de agentes– cuando toma una decisión.

Esta teoría es fruto de una lógica *consecuencialista*, justificada por las siguientes tres premisas. Primero, la toma de decisiones bajo riesgo es –o debe ser– un proceso que involucra la evaluación de resultados *mutuamente excluyentes* y, por lo tanto, la evaluación de los posibles resultados debe hacerse de manera separada. Segundo, la evaluación de los resultados es –o debe ser– expresable en términos de una utilidad cardinal, u . Tercero, la utilidad final de cada lotería es la suma de las utilidades de todos sus posibles resultados, ponderados por la probabilidad de ocurrencia de cada uno, p_i .

Von Neumann y Morgenstern (1944) fueron los primeros en derivar EUT formalmente. Estos autores postulan que las decisiones bajo riesgo se basan en cuatro axiomas que definen la racionalidad. Los primeros tres son comunes a la teoría del consumidor bajo certidumbre y a la mayoría de los modelos de elección previos. El cuarto axioma, y el más controversial, explota el hecho de que los sucesos transcurren de manera mutuamente excluyente y le dan a EUT su estructura distintiva.

1. *Axioma de la completitud*

Las preferencias de un agente son completas si para cada par de loterías X y X' , $X' \succeq X$, $X \succeq X'$, o ambos.⁴

2. *Axioma de la transitividad*

Las preferencias de un agente son transitivas si para cualquier grupo de loterías X , X' y X'' , si es que $X \succeq X'$ y $X' \succeq X''$, entonces $X \succeq X''$.

⁴ La relación \succeq denota preferencia, con lo cual $X' \succeq X$ significa que X' es preferido a X . También usaremos la relación \sim , que significa indiferencia entre ambas opciones.

Estos dos axiomas postulan propiedades convencionales sobre cualquier orden jerárquico. La completitud requiere que todas las loterías sean comparadas entre sí y que se juzgue la preferencia —estricta o no— de una lotería sobre otra, o la indiferencia entre sí. Similarmen- te, la transitividad implica que todas las loterías sean clasificadas de acuerdo con esta evaluación.

3. *Axioma de continuidad (o axioma arquimediano):*

Las preferencias de un agente son continuas si para cualquier conjunto de tres loterías X , X' y $X'' \in X$, donde $X'' \succeq X' \succeq X$, existe un $\alpha \in [0,1]$, tal que $\alpha X + (1-\alpha) X'' \sim X'$.

Cualitativamente, el axioma describe una situación en la que se presentan dos alternativas: una compuesta por las loterías X y X'' con probabilidad α y $1-\alpha$, respectivamente, y la otra es la lotería X' , que sucede con certeza. El axioma de la continuidad dice que existe un valor de α que causa indiferencia entre las dos alternativas. Por ende, un cambio en α suficientemente pequeño no alterará el orden de la preferencia entre las dos alternativas. Sin embargo, también implica que existen valores de α que evalúan las dos alternativas en forma dife- rente. Esto descarta la posibilidad de que haya un orden lexicográfico, en el cual un resultado es infinitamente superior que los demás. Este axioma permite que preferencias completas y transitivas sean repre- sentadas por una función de utilidad.

4. *Axioma de la independencia*

Las preferencias de un agente satisfacen la independencia si para las loterías X , X' y X'' , y $\alpha \in [0,1]$, tenemos

$$X' \succeq X \Leftrightarrow \alpha X' + (1-\alpha) X'' \succeq \alpha X + (1-\alpha) X''.$$

Este último axioma impone que la clasificación entre loterías no sea afectada por resultados contingentes a otra lotería. Cuando comparamos las alternativas $\alpha X + (1-\alpha) X''$ y $\alpha X' + (1-\alpha) X''$, solo debemos fijarnos en las virtudes relativas entre X y X' , sin importar las posibles consecuencias de X'' o el valor de α . Esto implica que las preferencias cambian de manera lineal con la probabilidad asociada a los posibles resultados de cada lotería. Es decir, evaluar la utilidad que generan dos loterías en conjunto, X y X' , $U(\alpha X + (1-\alpha) X')$ es lo mismo que hacerlo por separado, $\alpha U(X) + (1-\alpha) U(X')$.⁵ Para que esto ocurra, debemos concluir que la utilidad es una función lineal de la probabilidad.

Esta línea argumentativa conlleva al teorema de la representación. Este teorema dice que una clasificación completa y transitiva en el espacio de consecuencias de las loterías relevantes, X , es representada por una función de utilidad con la estructura de la utilidad esperada solo si se satisfacen los axiomas de la continuidad y de la independencia. La estructura de esta función proviene de asignar un valor cardinal de utilidad (u_1, \dots, u_i) a cada uno de los N resultados posibles (x_1, \dots, x_n), de manera que la utilidad de una lotería esté dada por (1.1). La idea importante es que esta representación de la función de utilidad esperada evalúa los posibles resultados de cada lotería de manera independiente, lo cual equivale a la evaluación de loterías de manera lineal en la probabilidad. Matemáticamente, $U(\alpha X + (1-\alpha) X') = \alpha U(X) + (1-\alpha) U(X')$, como implica el axioma de la independencia. Entonces, una clasificación de loterías que cumpla los cuatro axiomas anteriores solo puede ser representada por una función con la estructura de la utilidad esperada.

5 Esta conclusión es aparente si cambiamos la relación entre X y X' por una de indiferencia. Entonces, si $X \sim X'$, aplicando la independencia tenemos $\alpha X + (1-\alpha)X'' \sim \alpha X' + (1-\alpha) X''$. Ahora, si esta relación no es lineal en α , $U(\alpha X + (1-\alpha)X'')$ no es igual que $\alpha U(X) + (1-\alpha)U(X'')$ para todos los valores de α .

Aversión al riesgo: la función de Bernoulli

El debate en esta sección se centra en loterías exclusivamente definidas en términos monetarios, en las cuales el espacio de consecuencias es la línea real. En el caso en que las consecuencias son continuas, cada lotería es representada por una función de distribución acumulada F , donde $F(x)$ es la probabilidad de que la consecuencia monetaria de la lotería sea igual o menor que x soles. Si U es una función que representa las preferencias con la estructura de la utilidad esperada, entonces para cualquier distribución acumulada F ,

$$U(F) = \int u(x) dF(x). \quad (1.2)$$

Nos referimos a U como la función de utilidad de von Neumann-Morgenstern y a u como la función de utilidad de Bernoulli. Notablemente, la representación de la utilidad esperada sobre un espacio continuo de consecuencias no solo depende de la media y varianza de la distribución, sino también de sus demás momentos.

Los axiomas de la utilidad esperada no imponen restricción alguna sobre la forma de la función de utilidad de Bernoulli. Desde el punto de vista económico, gran parte del poder analítico yace en la especificación de esta función, ya que caracteriza aspectos interesantes de la naturaleza del comportamiento. En la literatura se suele restringir u a funciones crecientes y continuas en x .

Llamamos a una persona *aversa al riesgo* si la función de utilidad de Bernoulli u asociada a sus preferencias es cóncava. Formalmente, denotemos como γ_x a una lotería degenerada –es decir, que da x con certeza–. Una persona actúa de manera estrictamente aversa al riesgo si, entre una lotería no degenerada $F(\cdot)$ con valor esperado $E_F = \int dF(x)$ y la lotería γ_{E_F} , prefiere γ_{E_F} a $F(\cdot)$. Esto es equivalente a decir, de forma más general, que para todas las loterías F ,

$$\int u(x) dF(x) \leq u(\int x dF(x)). \quad (1.3)$$

Esta expresión matemática es llamada la desigualdad de Jensen, y es la propiedad que define una función cóncava. En nuestro contexto, la desigualdad quiere decir que para una persona aversa al riesgo la utilidad del valor esperado de cualquier lotería es mayor que la utilidad esperada de esta.

De similar manera, es útil definir el concepto de *equivalente cierto*, o monto x por el cual un agente sería indiferente entre x y tomar una lotería $F(\cdot)$. Dada una función de utilidad de Bernoulli, u , podemos definir el equivalente cierto, $c(F, u)$, de cualquier lotería $F(\cdot)$ como

$$u(c(F, u)) = \int u(x) dF(x). \quad (1.4)$$

Para un agente averso al riesgo –cuya u es cóncava–, $c(F, u)$ es necesariamente menor que el valor esperado de la lotería.

Grados de aversión al riesgo

Intuitivamente, la curvatura de la función de Bernoulli está relacionada con la aversión al riesgo –entre más cóncava, más aversión–. Sin embargo, la segunda derivada de la función, $u''(x)$, no es buen indicador, ya que varía con transformaciones lineales positivas. Este es un punto importante, ya que la concepción moderna de las preferencias es ordinal, no cardinal. No obstante, normalizar la segunda derivada de u al dividirla por la primera conserva esta propiedad deseable. Así, definimos el índice de Arrow-Pratt de aversión absoluta al riesgo,

$$r_a(x) = - \frac{u''(x)}{u'(x)}. \quad (1.5)$$

Este índice es apto para hacer comparaciones en la toma de decisiones de un individuo con diferentes niveles de riqueza. Es decir, compara actitudes hacia la toma de decisiones entre alternativas riesgosas cuyos resultados sean ganancias o pérdidas absolutas de la riqueza actual. Sin embargo, esta medida depende del nivel de riqueza del individuo, y dos individuos dispuestos a arriesgar el mismo valor monetario en la misma lotería tendrían el mismo grado de aversión al riesgo independientemente de que uno sea rico y el otro pobre.

Por esto, para establecer comparaciones entre individuos también es importante juzgar el grado de aversión al riesgo cuando las loterías están definidas sobre ganancias o pérdidas relativas. Cuando los resultados de la lotería están expresados por cambios del porcentaje de la riqueza actual, el índice adecuado es el coeficiente de aversión relativa al riesgo de Arrow-Pratt, dado por:

$$r_r(x) = -x \frac{u''(x)}{u'(x)}. \quad (1.6)$$

Utilizando los coeficientes de Arrow-Pratt, nos es posible discutir formas funcionales de u . Una posibilidad atractiva es asumir que las preferencias de los sujetos se rigen bajo una función de aversión relativa al riesgo constante (CRRA, por sus siglas en inglés):

$$u(x) = \frac{x^{1-r} - 1}{1 - r}, \quad (1.7)$$

donde x es el resultado evaluado, y r es el parámetro de aversión al riesgo. Esta función es cóncava cuando r es mayor que 0 –el caso del averso al riesgo–, lineal cuando es igual que 0 –neutral al riesgo– y convexa cuando es menor que 0 –amante del riesgo–. En esta función, r es la elasticidad de la utilidad y, en la teoría de la demanda del consumidor, refleja la elasticidad de la sustitución (Wakker, 2008). Al

derivar la elasticidad de la función CRRA aplicando la ecuación (1.6) obtenemos

$$r_r(x) = r.$$

Cuando derivamos el índice de aversión relativa al riesgo con respecto al resultado, vemos que

$$\frac{\partial r_r(x)}{\partial x} = 0.$$

De esta manera, concluimos que la aversión al riesgo relativo no cambia con x . Es decir, la aversión al riesgo con forma funcional CRRA es constante en la riqueza y , por lo tanto, independiente del valor del resultado.

Otra consideración que tomar antes de asumir la forma funcional de u es la idea de que un orden de preferencias que exhibe aversión absoluta al riesgo decreciente (DARA) es intuitivo. Para preferencias de este tipo, derivamos la expresión (1.5) con respecto al consumo o la riqueza y podemos apreciar que el coeficiente de aversión absoluta al riesgo decrece con estos. En el caso de utilidad CRRA –ecuación (1.7)–, se halla que

$$\frac{\partial r_a(x)}{\partial x} = -\frac{r}{x^2}.$$

Esto implica que una persona con preferencias CRRA está dispuesta a pagar una cantidad menor por eliminar el riesgo de una lotería X cuando su riqueza, o el conjunto de resultados, x_n , se incrementa.

1.2. Elección intertemporal

Las preferencias intertemporales condicionan la forma en la que una persona distribuye sus beneficios y costos a lo largo del tiempo. Su estudio no solo ayuda a predecir mejor las decisiones que un individuo tomará sobre los recursos que asignará a su salud, vivienda o educación; también es crucial para la elaboración y evaluación de políticas públicas. El estudio de las preferencias intertemporales es tan antiguo como el de la economía, y hasta el día de hoy es sujeto a revisión y avance, dada la naturaleza única y compleja del comportamiento individual.

La forma de modelar las decisiones intertemporales ha sido, principalmente, mediante el modelo de utilidad-descontada (DU, en adelante). En este esquema, un individuo asigna menos peso a su bienestar futuro de manera constante y consistente por medio de un parámetro llamado *tasa de descuento*. Debido a su simplicidad y elegancia, este modelo se usa ampliamente hasta el día de hoy. Sin embargo, no se ha salvado de diversas críticas con respecto a su validez empírica.

Desde el surgimiento de la Economía, hay una preocupación por la dimensión temporal de las decisiones individuales. Los economistas clásicos se enfocaron en un primer momento en identificar los patrones psicológicos detrás de estas decisiones. En específico, Rae (1834) proponía que existe un “deseo efectivo de acumulación” en cada individuo y que este determina la forma en la que una persona distribuye sus acciones en el tiempo. Este deseo se vería influenciado por factores relacionados con el anhelo de dejar un legado, con la capacidad de autorrestricción, con la incertidumbre en la vida diaria y con el nivel de gratificación del consumo inmediato.

La naturaleza psicológica presente en estos cuatro factores fue desarrollada ampliamente hasta inicios del siglo XX. Un aporte valioso es el de Jevons (1871), en el que sostiene que a una persona solo

le importa su utilidad inmediata y que solo retrasaría una acción si la utilidad anticipada que le brindara fuera mayor. Una perspectiva distinta de analizar este fenómeno fue propuesta por Senior (1836), quien sugiere que tanto el presente como el futuro son valorados de igual manera, pero que es la falta de abstinencia la que lo lleva a uno a consumir más en el presente. A pesar de contar con diferentes mecanismos de postergación, ambas visiones coinciden en la presencia de sentimientos inmediatos –sean positivos o negativos– en el proceso de decisión intertemporal. Sumado a estos dos, Böhm-Bawerk (1889) analiza que otro factor que se debe tomar en cuenta es el hecho de que las personas subestiman sus deseos futuros de forma sistemática. Cree que esto puede deberse a falta de imaginación, incapacidad analítica o, simplemente, a que no hacen el esfuerzo de pensar en el futuro. A pesar de no llegar a un consenso sobre el motivo detrás de la decisión intertemporal, cada vez resultaba más claro que el comportamiento humano en este ámbito seguía ciertos patrones que merecían ser analizados.

La contribución de Böhm-Bawerk (1889) no solo fue importante por el motivo psicológico adicional que propuso, sino también por su manera de tratar el asunto como cualquier tipo de decisión “técnica” en la que uno distribuye recursos siguiendo un criterio – en este caso temporal–, tal y como se compararía el consumo entre dos bienes físicos, como la papa y el arroz. Gracias a esta forma de enfrentar el problema, Fisher (1930) logra contemplar el consumo intertemporal como un diagrama de curvas de indiferencia entre dos niveles de consumo en periodos distintos. Un aspecto relevante que brinda este diagrama es que permite identificar que, además de la preferencia temporal, otra noción que se debe tener en cuenta es la utilidad marginal decreciente que se asume del individuo representativo, lo cual lo lleva a suavizar su consumo entre ambos periodos. Aun así, por más de ser pionero en analizar el tema en forma matemática y

en traer a discusión el concepto de sustitución intertemporal, todavía existían varias limitaciones que necesitaban ser ajustadas, entre las que resaltaba la de no poder resolver el problema con más de dos periodos (Frederick et al., 2002).

El punto de inflexión en cuanto al estudio de decisiones intertemporales aparece con Samuelson (1937). Inspirado en el concepto de valor presente utilizado en las finanzas, y en un intento por demostrar cómo insertar consideraciones temporales en la medición de utilidad –de una manera que sea matemáticamente manejable–, esboza un modelo que agrega todos los motivos profundizados por sus antecesores y los sintetiza en un único parámetro al que denomina *tasa de descuento*. Esta tasa se define como un índice del deseo de postergar el consumo para recibir utilidad futura. Asume, también, que esta tasa es constante a través del tiempo –esto es, que forma parte de las preferencias inmutables del individuo–. Por su aporte claro, simplificado y amigable a la expresión matemática, el modelo DU pasó, en poco tiempo, a convertirse en la forma más prevalente de modelar las decisiones intertemporales. Con el tiempo, el modelo DU se convirtió en el enfoque por defecto y dejó de cuestionarse. Esto a pesar de que Samuelson expresara su preocupación por la gran cantidad de permisos que su modelo requiere y la falta de conexión empírica.

Modelo de utilidad descontada (DU)

El modelo DU se construye sobre las siguientes premisas. Un individuo vive desde el periodo 0 –que, en ocasiones, representa el momento de la decisión– hasta el T . En este lapso, se asume que las preferencias no cambian y que están representadas por una función de utilidad $u(\cdot)$. El objetivo del individuo es maximizar su bienestar

agregado considerando todos los periodos en los que sabe que vivirá. Para hacerlo, se descuenta de un periodo al siguiente según una tasa δ , de manera que, mientras mayor es la diferencia entre periodos de tiempo, el descuento aumenta en forma exponencial. En tal sentido, el individuo busca maximizar:

$$U(x_0, \dots, x_T) = \sum_{t=0}^T \left(\frac{1}{1 + \delta} \right)^t u(x_t) \quad (1.8)$$

donde x_t es el nivel de consumo en el periodo t y $U(\cdot)$ es la función de utilidad agregada. Como Samuelson (1937) indicó en su momento, esta representación solo sirve como una primera aproximación al verdadero mecanismo de decisión. Por esto, detrás del modelo DU existe un conjunto de propiedades que necesita ser revisado para entender sus virtudes y limitaciones. Koopmans (1960) elaboró una serie de axiomas que, de cumplirse, harían posible la preferencia temporal positiva en los individuos.

Para empezar, se asume que las preferencias de una persona son invariantes en el tiempo, lo cual se representa mediante una utilidad instantánea estacionaria. Luego, el nivel de bienestar en un periodo determinado es independiente del consumo de algún otro periodo —por lo que la tasa de sustitución entre un periodo y otro tampoco depende del consumo en algún otro periodo—. Asimismo, la utilidad también se trata como independiente en el sentido de que no hay una preferencia *per se* por algún patrón particular de consumo —ya sea creciente o decreciente—, sino solo el valor numérico de la utilidad agregada. Se considera que la utilidad marginal es decreciente y que consumir en el presente genera mayor utilidad que hacerlo en el futuro. Estos últimos puntos son interesantes porque generan el *trade-off* entre suavizar el consumo a lo largo del tiempo y mandar el consumo a la inmediatez. Por otro lado, el descuento temporal no depende del

bien consumido; es decir, a pesar de que el bien sea comida o videojuegos, se asume que la tasa de descuento asociada es independiente. Más aun, en la especificación original, la tasa de descuento entre dos periodos consecutivos se supone constante, sin importar la posición temporal en la que se encuentren ambos periodos. El hecho de que la tasa sea constante brinda consistencia dinámica al modelo: una persona asignará el mismo nivel de consumo al periodo t sin importar el periodo en el que lo decida. Por último, se asume que, cuando se le presenta al individuo una nueva alternativa de consumo, este la evalúa integrándola a sus planes existentes. En otras palabras, si la alternativa de consumo X implica cambiar el sendero de consumo de (x_0, \dots, x_T) a (x'_0, \dots, x'_T) , la persona solo elegirá el nuevo plan si $U(x_0, \dots, x_T) < U(x'_0, \dots, x'_T)$.

El modelo DU en sus diferentes versiones –aun amparado por un grupo igual o muy parecido a estos axiomas– fue dominante en la comunidad académica durante todo lo que restaba de la segunda mitad del siglo XX. A pesar de las reservas de sus creadores, los supuestos fueron en gran medida aceptados y pocos se atrevieron a criticar su validez. Recién a finales de los años 1990 empezaron a realizarse nuevas investigaciones y la anomalía más relevante que encuentran gira en torno a que la tasa de descuento no es estática en el tiempo, sino que va decreciendo –a lo que se denomina *tasa de descuento hiperbólico*–. Existen otras anomalías, pero esta ha sido la más documentada por su drástico impacto sobre el análisis de preferencias a corto plazo.

Modelo de descuento hiperbólico

El término *descuento hiperbólico* surge para referirse a una tasa de descuento que disminuye en el tiempo. Según esta modificación, el

descuento que se aplique a la utilidad del periodo inmediatamente posterior va a ser mayor mientras más cercano sea el periodo inicial. Thaler (1981) ilustra muy bien este concepto con el siguiente ejemplo. Se le presentan al sujeto dos loterías con dos opciones cada una: en la primera, se le ofrece una manzana hoy o dos manzanas mañana; y en la segunda, se le ofrece una manzana en un año o dos manzanas en un año más un día. El autor sugiere que, mientras algunos estarían tentados a elegir la manzana hoy en la primera lotería, nadie elegiría la manzana en un año en la segunda lotería, a pesar de que la diferencia temporal sea de un día para ambos casos. En efecto, si la tasa de descuento fuese constante, si es que el sujeto eligiese la manzana hoy, también debería elegir la manzana en un año, lo cual no ocurre.

Desde Böhm-Bawerk (1889) persistía un escepticismo entre los economistas con respecto a la validez de una tasa de descuento constante en el tiempo, pero tendría que pasar más de medio siglo para que Strotz (1955) planteara formalmente la falta de sentido de que un individuo tenga una función de descuento tan especial como la exponencial. De hecho, Strotz reconoce que la función de descuento exponencial es la única que implica preferencias temporalmente consistentes e imposibilita al sujeto de reconsiderar sus decisiones anteriores. Por el contrario, él sostiene que, generalmente, las preferencias de los individuos son inconsistentes, y que sus decisiones se ven mejor modeladas por una función de descuento que dependa de la diferencia entre el periodo actual y el periodo en cuestión.

Sobre la base de estas reflexiones, hacia finales del siglo XX, se empezaron a proponer modelos específicos que contemplaran el caso de una tasa de descuento no constante. En particular, surgió un interés por representar las implicancias de una tasa decreciente de preferencias temporales. Tal es así que Laibson (1997), con la finalidad de identificar la relación entre preferencias por la inmediatez y patrones

de ahorro y liquidez, utiliza la siguiente representación de la utilidad agregada del individuo:⁶

$$U(x_0, \dots, x_T) = u(x_0) + \beta \sum_{t=1}^T \left(\frac{1}{1 + \delta} \right)^t u(x_t) \quad (1.9)$$

donde, a diferencia del modelo anterior, se añade un parámetro β , que representa el sesgo por el presente. Este parámetro mide si las personas castigan más los flujos de consumo futuros cuando se tiene que decidir contra la inmediatez. En este sentido, si $\beta < 1$, la tasa de descuento entre el periodo actual y el subsiguiente es mayor que la tasa de descuento entre dos periodos consecutivos en el futuro. Si $\beta = 1$, la función colapsa a la versión estándar de descuento exponencial. Como en este modelo la tasa de descuento solo es estrictamente decreciente entre el primer y segundo periodo, se lo conoce mejor como modelo de descuento *cuasihiperbólico* (QHD, por sus siglas en inglés), pues la función de descuento no decrece continuamente como una hipérbole, pero igual sirve para distinguir entre los conceptos de corto y largo plazo. Existe vasta evidencia que soporta un mejor ajuste de los datos utilizando un modelo de descuento hiperbólico o cuasihiperbólico sobre uno exponencial (Kirby, 1997; Myerson y Green, 1995; Rachlin et al., 1991). Sobre estos avances, nuevas investigaciones trataron de relacionar el sesgo por el presente con la procrastinación o la falta de autocontrol. Una de las más relevantes es la de O'Donoghue y Rabin (1999), en la que el descuento hiperbólico lleva a una persona a postergar ciertas actividades, a pesar de no haberlo deseado en periodos anteriores. Ellos incluso identifican dos tipos de consumidores, a los que denominan *ingenuos* y *sofisticados*, que reaccionan de manera

6 Los primeros en proponer esta forma funcional fueron Phelps y Pollak (1968) en un estudio que buscaba analizar la relación entre altruismo y herencia.

distinta frente a la inconsistencia temporal inherente de su sesgo por el presente.

Aparte del descuento hiperbólico, existen otras anomalías del modelo DU para las que se han propuesto modelos alternativos de decisión intertemporal. Algunas de las alternativas más relevantes las constituyen el modelo de formación de hábitos, ideado por Duesenberry (1952) y desarrollado por Pollak (1970); modelos con puntos de referencia temporales, basados en la teoría prospectiva de Kahneman y Tversky (1979); modelos que incorporan utilidad por anticipación, iniciados por Loewenstein (1987), entre otros.

2. EXPERIMENTO

2.1. Medición de preferencias

2.1.1. *Actitudes ante el riesgo*

En general, los economistas optan por medir la actitud ante el riesgo de un sujeto usando métodos compatibles con incentivos reales. Dos métodos con esta característica son prominentes en la literatura. El primero, y más utilizado, es el de las decisiones binarias. Los economistas experimentales suelen preferir este método porque requiere una carga cognitiva relativamente baja y se asemeja a las decisiones de la vida real, donde se presentan decisiones binarias entre las cuales el sujeto debe optar por una de dos loterías. Por ejemplo, dos estudios característicos de este método son Hey y Orme (1994) y Andersen et al. (2008). Los primeros expusieron a sus sujetos experimentales a 100 decisiones binarias entre 4 loterías diferentes, en las que cambiaban las probabilidades de los resultados. Andersen et al. (2008) hacen que los sujetos respondan 40 preguntas de riesgo sobre 4 loterías diferentes, pero cuyo orden juega un papel importante. Para hacer las decisiones más realistas, en ambos experimentos se escogió una de las decisiones que los sujetos tomaron de manera aleatoria y se pagó por el resultado de la lotería. Usando los resultados de las decisiones binarias, los investigadores infirieron coeficientes de aversión al riesgo.

El otro método comúnmente utilizado en la literatura es el de Becker et al. (1964) y es llamado Becker-DeGroot-Marschak (o BDM), por sus creadores. Este método se centra en obtener el equivalente cierto de una lotería cuando los sujetos revelan su disposición a pagar por ella o su disposición de venderla. Para lograr esto, el investigador presenta una lotería en la que, por ejemplo, el resultado mayor es x_1 y el menor es x_2 . Entonces, el investigador le pregunta al sujeto por qué monto, $c(F, u)$, estaría dispuesto a vender la lotería. Como puede notar, si este es un ejercicio pagado, el sujeto está incentivado a responder con el monto mayor. Para evitar que el sujeto haga esto y sea en su mejor interés responder honestamente, solo podrá cambiar la lotería si otro sujeto está dispuesto a pagar $z \geq c(F, u)$. Este proceso puede ser emulado con un generador de números aleatorio, donde z sea sacado de una distribución uniforme entre x_1 y x_2 , y se le pague si $c(F, u) \leq z$; de lo contrario, el sujeto jugará la lotería. De esta manera, asumiendo cierta forma funcional u , el investigador puede inferir el coeficiente de aversión al riesgo usando la ecuación (1.4).

2.1.2. Preferencias intertemporales

Para medir las preferencias intertemporales, se deben poder observar las decisiones del individuo en distintos escenarios, en los que varíe la cantidad de dinero o el tiempo de espera. Existen diversas formas de hacerlo, que varían según el marco teórico que está detrás y los objetivos de la investigación. La medición puede llevarse a cabo mediante el análisis del comportamiento en los quehaceres cotidianos del individuo o, como es más común, mediante experimentos. Según Frederick et al. (2002), las mediciones de preferencias nunca serán perfectas, pues, por mucho que se pueda controlar, siempre existirá un

componente no observable, cuya minimización debe ser el enfoque de metodologías de estimación que se empleen.⁷

Precisamente, los métodos experimentales son los más empleados debido a su mayor capacidad de control. El mecanismo usual en estos experimentos consiste básicamente en preguntar al sujeto si prefiere un monto de dinero menor en un periodo más cercano o un monto de dinero mayor en un periodo más lejano –por ejemplo, $S/100$ hoy o $S/150$ en un año–. Así, si en la decisión anterior el sujeto prefiere la alternativa en el periodo más cercano, entonces el valor de δ es tal que $u(100) > 1/(1 + \delta) \cdot u(150)$. Si en la siguiente decisión que se le plantea al sujeto –ahora con un monto cercano de $S/110$ – este responde que prefiere la alternativa cercana, entonces se sabe que $u(110) < 1/(1 + \delta) \cdot u(150)$. De ambas respuestas se tiene que δ se encuentra en el intervalo $\left[\frac{u(150)}{u(110)} - 1; \frac{u(150)}{u(100)} - 1 \right]$.⁸ Se presentan varias de estas decisiones a distintos individuos, de tal forma que sea posible estimar las preferencias intertemporales de un grupo de sujetos.

Desde los primeros intentos de medir las preferencias intertemporales, han sido dos los cambios de mayor relevancia en este proceso: (i) la incorporación de mecanismos para estimar el descuento hiperbólico y (ii) la corrección por la concavidad de la función de utilidad. Respecto al primer punto, como los primeros modelos sobre los que se construyeron las metodologías de estimación eran de utilidad-descontada (DU) exponencial, el interés giraba en torno a la tasa de descuento exponencial y no al descuento hiperbólico. Coller y Williams (1999) fueron los primeros en llevar a cabo estimaciones

7 Frederick et al. (2002) sugieren que es casi imposible un adecuado control por las expectativas de inflación del sujeto, sus capacidades de arbitraje, la forma en la que incorpora nuevas opciones de consumo en sus planes, su aversión a la incertidumbre o la verdadera forma de su función de utilidad.

8 Esto aplica al utilizar el modelo DU como marco teórico detrás.

experimentales teniendo en cuenta la forma no constante de la tasa de descuento, para lo cual utilizan el mecanismo *front-end-delay* (FED), que consiste en volver a presentar los mismos montos de dinero, pero ambos retrasados en la misma cantidad de tiempo.⁹ Sobre el segundo punto, debido a la mayor simplicidad en el procedimiento y a que la función de utilidad puede considerarse muy cercana a ser lineal para cantidades de dinero bajas, los primeros estudios sobre el tema no controlan por una posible concavidad de la utilidad. Por ende, dichas mediciones podrían encontrarse sesgadas hacia arriba. En cuanto nos concierne, Chapman (1996) es la primera en utilizar funciones de utilidad cóncavas para la estimación de tasas de descuento y halla diferencias bastante significativas con respecto a una función lineal. A continuación, se detallan los procedimientos necesarios para controlar por la curvatura de la utilidad en la medición de preferencias intertemporales, haciendo uso de las actitudes frente al riesgo.

2.1.3. Actitudes ante el riesgo y elecciones intertemporales

Para poder llevar a cabo inferencias sobre las preferencias intertemporales de un sujeto, es necesario conocer la forma de su función de utilidad y el grado de concavidad de esta. A pesar de que en los primeros estudios que modelan las preferencias intertemporales se asumió que la utilidad individual se podía representar por una función lineal, hoy existe una vasta literatura que cuestiona esta metodología e incorpora la aversión individual al riesgo en su medición. Por ende, argumentamos que es necesario estimar las preferencias ante el riesgo y el tiempo

9 Por ejemplo, se presenta la decisión entre S/100 hoy o S/150 en un año, y también se presenta la decisión entre S/100 en un mes o S/150 en un año más un mes.

en un mismo diseño experimental. Solo así es posible identificar las preferencias intertemporales de manera coherente con la teoría de la utilidad esperada. Para la estimación utilizamos un método bastante similar al DMPL (Double Multiple Price List) propuesto por Andersen et al. (2008), por su simplicidad y efectividad.

En un principio, las mediciones de ambas preferencias se han llevado a cabo de manera separada. Los estudios previos se solían enfocar en una de las dos actitudes, pero no en ambas en simultáneo. Siendo las actitudes hacia el riesgo y hacia el tiempo de suma importancia para las decisiones, no tomó mucho tiempo hasta que se notara su interdependencia teórica, ante lo cual surge la necesidad de medir ambas actitudes en un mismo diseño experimental. Chapman (1996) fue la primera en medir tasas de descuento teniendo en cuenta la concavidad de la función de utilidad; no obstante, no mide las actitudes al riesgo de los sujetos que utiliza –sus resultados dependen de supuestos arbitrarios sobre la concavidad de la utilidad–. Sin embargo, demuestra que las tasas de descuento –asumiendo aversión al riesgo– son distintas de las estimadas bajo el supuesto de neutralidad al riesgo. Otros pioneros de diseños que midan ambas preferencias en conjunto fueron Anderhub et al. (2001), que realizan un experimento usando el método BDM (ver sección 2.1.1), y estiman tasas de descuento y coeficientes de aversión al riesgo sobre la base de una sola pregunta; Eckel et al. (2002), que, en lugar de hacer una misma pregunta para medir ambas actitudes, realizan dos experimentos separados para cada persona, pero no estiman parámetros de una función de utilidad; y Tanaka et al. (2010), que también miden las mismas actitudes de manera completamente separada, aunque no condicionan sus mediciones de tiempo usando las de riesgo.

En un intento por armonizar los aportes anteriores, Andersen et al. (2008) se basan en las propuestas de medición de riesgo y tiempo

de Holt y Laury (2002), y Coller y Williams (1999), respectivamente, para construir su metodología de medición conjunta. A diferencia de los diseños anteriores, los mismos autores sostienen que la propuesta que ellos escogen supone un reto cognitivo menos restrictivo para los participantes, pues las preguntas que se les hacen son más directas y simples. En efecto, las preguntas de riesgo y tiempo se hacen por separado y utilizando la metodología MPL (Multiple Price List). MPL es un procedimiento simple en el que se confronta al sujeto con varias preguntas en las que se le pide elegir una de las dos alternativas. Estas preguntas se hacen de manera ordenada, de forma que el valor de una de las alternativas sea cada vez menos deseado por el sujeto. El punto en el que el sujeto cambia de decisión –si es que lo hace– de una alternativa a la otra es indicador de sus preferencias. A esta versión de Andersen et al. (2008) se le denomina DMPL (Double Multiple Price List) porque utiliza dos versiones del MPL: una para las actitudes ante el riesgo y otra para las preferencias de elección intertemporal. Este es el diseño en el que más nos basamos para estimar las preferencias en este estudio, aunque nuestro procedimiento tiene algunas ligeras diferencias. La primera es que no le mostramos todas las alternativas desde un principio al sujeto –para evitar efectos de *framing*–; y la segunda, que no realizamos una sola estimación para obtener todos los parámetros de preferencias, sino que primero realizamos la estimación de preferencias ante el riesgo y luego estimamos las preferencias de elección intertemporal, condicionándolas por los parámetros de preferencia ante el riesgo.

Por supuesto que la metodología DMPL no se encuentra exenta de críticas y alternativas. La más relevante es quizás la de Andreoni y Sprenger (2012a) y su método CTB (Convex Time Budget). Ellos proponen que, en lugar de presentar al sujeto opciones binarias como en el DMPL, se le pida optimizar su consumo intertemporal sujeto

a un presupuesto convexo. Según los autores, esto permite obtener parámetros más precisos y una mejor predicción fuera de la muestra (Andreoni et al., 2015). Para hacerlo, el método CTB consiste en pedirles a los participantes que distribuyan un monto de “tokens” entre un periodo más cercano y otro periodo más lejano. El valor de los “tokens” destinados al periodo más lejano es mayor que el de los “tokens” del periodo más cercano.¹⁰ Para poder identificar las preferencias, permiten variación en el espacio temporal entre ambos periodos y en el valor relativo de los “tokens” en el periodo más lejano. Sobre la base de estos escenarios, se pueden estimar conjuntamente las preferencias por riesgo y tiempo con una misma tarea. En respuesta a esta alternativa y a sus críticas al DMPL, Andersen et al. (2016) sostienen que el método CTB tiene supuestos muy fuertes con respecto a la relación entre aversión al riesgo y elasticidad de sustitución intertemporal.¹¹ Según estos autores, la aversión al riesgo es un fenómeno atemporal y que describe el comportamiento del individuo en un solo periodo de tiempo.

10 En lugar de maximizar una función de utilidad $U(x_t, x_{t+k})$ sujeta a un presupuesto binario: $((1+r) \cdot x_t, x_{t+k}) \in \{(m,0), (0,m)\}$, el método CTB permite optimizar sujeto a un presupuesto convexo: $(1+r)x_t + x_{t+k} = m$.

11 Específicamente, indican que lo que el CTB identifica como aversión al riesgo describe mejor otro fenómeno denominado *aversión a la correlación intertemporal*—desarrollado por Epstein (1983), Epstein y Zin (1989) y (Farmer 1990)—, que explicaría la magnitud de la preferencia por la variabilidad de la utilidad a lo largo del tiempo. Los autores proponen la siguiente función de utilidad agregada para describir la aversión a la correlación intertemporal: $U(x_t, x_{t+k}) = E \left[\frac{1}{1-\eta} \left[\left(\frac{1}{1+\delta} \right)^t \cdot u(x_t) + \left(\frac{1}{1+\delta} \right)^{t+k} \cdot u(x_{t+k}) \right]^{1-\eta} \right]$, donde η es el parámetro de aversión relativa al riesgo intertemporal. Cuando $\eta \neq 0$, la utilidad sigue siendo separable, pero no aditiva, lo que significa que la correlación de la utilidad en distintos periodos de tiempo es relevante para el sujeto.

2.2. Diseño experimental

Para obtener una medición de las actitudes de interés según el modelo de preferencias amparado en EUT y QHD, utilizamos una metodología basada en el DMPL. En este método, el individuo toma varias decisiones en las que revela tanto sus actitudes frente al riesgo como sus preferencias intertemporales en tareas específicas y separadas. En particular, hay un conjunto de decisiones que corresponden a la medición de actitudes frente al riesgo y otro conjunto que corresponde a la medición de actitudes frente al tiempo. Ambos conjuntos comprenden 6 tareas con 10 *decisiones* cada una; en total, 30 decisiones para riesgo y otras 30 para tiempo.¹²

2.2.1. Decisiones intertemporales

La primera parte del diseño corresponde a las decisiones frente al tiempo y se basa en Coller y Williams (1999). En cada una de las 30 decisiones de tiempo, el sujeto elige entre dos alternativas A y B . Las alternativas difieren en dos aspectos: la cantidad de dinero que se ofrece y el plazo de tiempo que debe pasar antes de que se reciba el dinero. Así, se tiene como alternativa A un monto X asociado a un plazo de tiempo t –en meses– y, en la alternativa B , otro monto $Y > X$ que será pagado en un plazo de tiempo $t + k$, donde $k > 0$.

En las decisiones dentro de cada tarea, X y k se mantienen constantes, e Y cambia de modo que la tasa de interés efectiva anual (TEA) empiece en 10% y llegue al 100% en incrementos del 10%.¹³ Una vez

12 Cada decisión se presenta al encuestado de manera individual; es decir, al tomar una decisión, el participante no sabe qué otras elecciones deberá decidir luego.

13 Estos incrementos homogéneos son aproximados para garantizar que Y tenga solo un decimal.

que se complete una tarea, el plazo de tiempo k cambia. En las tareas 1, 2 y 3, los valores de k son 3, 6 y 12, respectivamente. A diferencia de Coller y Williams (1999), nosotros no mostramos a los participantes la TEA que representa la opción futura en cada decisión.

Adicionalmente, con probabilidad media, los sujetos reciben el tratamiento de *front end delay* (FED). Este consiste en que el monto de dinero de la alternativa A sea entregado a un mes del día de la decisión ($t = 1$). Para los sujetos que no reciben el tratamiento FED (NFED), la alternativa A es entregada inmediatamente ($t = 0$).

Para evitar que las decisiones se vean sesgadas por desconfianza o asimetrías en los costos de transacción, todos los pagos de tiempo se hicieron por medio de cheques, los cuales eran habilitados según el valor de $t + k$ de la tarea que se seleccionó para pago.¹⁴ Las tres tareas de tiempo están detalladas en el cuadro 2.1. En el anexo A, la figura A.1 contiene una pantalla de muestra de las que el individuo usa para tomar sus decisiones.

14 El participante fue informado, antes de tomar las decisiones, de que los pagos de las tareas de tiempo se harían mediante cheques, y que estos podían ser canjeados en el Banco de Crédito —que tiene presencia en todo el país—. Este método de pago permite mitigar la posible desconfianza acerca del pago futuro por parte de los sujetos. Similarmente, es importante que ambos pagos sean con cheques, para no incentivar una opción por la conveniencia de recibir efectivo.

Cuadro 2.1
Decisiones de tiempo

Decisión	NFED				FED				TEA (%)	Tasa de descuento s. a. $r=0,8$ (%)
	Opción A		Opción B		Opción A		Opción B			
	t	Monto	$t+k$	Monto	t	Monto	$t+k$	Monto		
<i>Tarea tiempo 1</i>										
1	0	90	3	92,2	1	90	4	92,2	10,14	1,95
2	0	90	3	94,2	1	90	4	94,2	20,01	3,72
3	0	90	3	96,1	1	90	4	96,1	29,99	5,39
4	0	90	3	97,9	1	90	4	97,9	40,01	6,96
5	0	90	3	99,6	1	90	4	99,6	49,99	8,45
6	0	90	3	101,2	1	90	4	101,2	59,86	9,84
7	0	90	3	102,8	1	90	4	102,8	70,22	11,22
8	0	90	3	104,2	1	90	4	104,2	79,68	12,43
9	0	90	3	105,7	1	90	4	105,7	90,25	13,73
10	0	90	3	107	1	90	4	107	99,79	14,85
<i>Tarea tiempo 2</i>										
1	0	90	6	94,4	1	90	7	94,4	10,02	1,93
2	0	90	6	98,6	1	90	7	98,6	20,02	3,72
3	0	90	6	102,6	1	90	7	102,6	29,96	5,38
4	0	90	6	106,5	1	90	7	106,5	40,03	6,97
5	0	90	6	110,2	1	90	7	110,2	49,93	8,44
6	0	90	6	113,8	1	90	7	113,8	59,88	9,84
7	0	90	6	117,3	1	90	7	117,3	69,87	11,18
8	0	90	6	120,7	1	90	7	120,7	79,86	12,46
9	0	90	6	124,1	1	90	7	124,1	90,13	13,71
10	0	90	6	127,3	1	90	7	127,3	100,07	14,88
<i>Tarea tiempo 3</i>										
1	0	90	12	99	1	90	13	99	10	1,92
2	0	90	12	108	1	90	13	108	20	3,71
3	0	90	12	117	1	90	13	117	30	5,39
4	0	90	12	126	1	90	13	126	40	6,96
5	0	90	12	135	1	90	13	135	50	8,45

Decisión	NFED				FED				TEA (%)	Tasa de descuento s. a. $r=0,8$ (%)
	Opción A		Opción B		Opción A		Opción B			
	t	Monto	$t+k$	Monto	t	Monto	$t+k$	Monto		
6	0	90	12	144	1	90	13	144	60	9,86
7	0	90	12	153	1	90	13	153	70	11,20
8	0	90	12	162	1	90	13	162	80	12,47
9	0	90	12	171	1	90	13	171	90	13,70
10	0	90	12	180	1	90	13	180	100	14,87

Bajo el supuesto de que el individuo sigue los modelos de EUT y QHD, estas tareas son suficientes para identificar su tasa de descuento intertemporal. En particular, condicional a la curvatura de su función de utilidad, la tasa de descuento se identifica usando el número de decisión en el que el sujeto cambia de la alternativa A a la alternativa B . Asimismo, la diferencia de elecciones entre los tratamientos NFED y FED nos permite identificar la magnitud del sesgo por el presente. Esto es porque en el tratamiento FED los pagos más cercanos son en el futuro, mientras que en el tratamiento NFED los pagos más cercanos son en el presente. La comparación de las decisiones entre ambos tratamientos nos permite estimar el parámetro del sesgo por el presente (β) de la ecuación (1.9).

2.2.2. Decisiones frente al riesgo

La segunda parte del diseño corresponde a las decisiones frente al riesgo y se basa en Holt y Laury (2002). El sujeto decide entre dos alternativas A y B , en donde la alternativa A ofrece la posibilidad de recibir el monto de dinero x con probabilidad p o recibir el monto de dinero x' con probabilidad $1-p$ ($x' > x$), y la alternativa B ofrece la posibilidad

de recibir el monto de dinero y con probabilidad p o recibir el monto de dinero y' con probabilidad $1-p$ ($y' > y$). Además, se cumple que $y' > x' > x > y$, lo que implica que la varianza de B es mayor que la varianza de A . Lo que va cambiando entre decisiones es el valor de la probabilidad p , que va ascendiendo de 0,1 a 1 a un ritmo de 0,1 por decisión. En la decisión 1, el valor esperado de A es sustancialmente mayor que el de B , y, como además B tiene mayor varianza, A es típicamente preferida. En las subsiguientes tareas, a medida que p crece, el valor esperado de la alternativa B se incrementa más rápidamente que el de A . En la novena decisión, el valor esperado de la opción B es sustancialmente mayor que el de A y, por tanto, es típicamente preferida.

Como se puede observar, hay un *trade-off* entre el pago esperado y la dispersión de los pagos. Si el individuo cambia de la opción A a la B cuando la diferencia de pagos esperados ($B - A$) pasa de ser positiva a negativa, entonces se deduce que se comporta aproximadamente de manera neutral al riesgo. Pero si el individuo “espera” que la diferencia de pagos $B - A$ sea más pronunciada para cambiar, significa que está sacrificando valor esperado por seguridad. En este caso, el individuo se comporta de manera aversa al riesgo. De esta forma, el punto en el que el sujeto cambie la alternativa A por la alternativa B (*switching point*) revela su coeficiente de aversión al riesgo.

Por construcción, la última decisión de cada tarea se reduce a escoger entre pagos x' e y' con certeza, por lo que esta decisión sirve como una prueba de racionalidad. Los pagos de esta tarea serán realizados en efectivo, de lo cual el participante es informado al inicio de la actividad. Las tres tareas correspondientes al riesgo están detalladas en el cuadro 2.2. En el anexo A, la figura A.2 contiene una pantalla de muestra de las que el individuo usa para tomar sus decisiones.

Cuadro 2.2
Decisiones de riesgo

Decisión	Lotería A				Lotería B				Dif. VE (B - A)	Intervalo abierto CRRRA	
	<i>n</i>	<i>p</i>	<i>x</i>	1- <i>p</i>	<i>x'</i>	<i>p</i>	<i>y</i>	1- <i>p</i>		<i>y'</i>	
<i>Tarea riesgo 1</i>											
1	0,9	60	0,1	75	0,9	4	0,1	150	-42,9	-∞	-1,59
2	0,8	60	0,2	75	0,8	4	0,2	150	-29,8	-1,59	-0,86
3	0,7	60	0,3	75	0,7	4	0,3	150	-16,7	-0,86	-0,41
4	0,6	60	0,4	75	0,6	4	0,4	150	-3,6	-0,41	-0,08
5	0,5	60	0,5	75	0,5	4	0,5	150	9,5	-0,08	0,2
6	0,4	60	0,6	75	0,4	4	0,6	150	22,6	0,2	0,46
7	0,3	60	0,7	75	0,3	4	0,7	150	35,7	0,46	0,73
8	0,2	60	0,8	75	0,2	4	0,8	150	48,8	0,73	1,01
9	0,1	60	0,9	75	0,1	4	0,9	150	61,9	1,01	1,41
10	0	60	1	75	0	4	1	150	75	1,41	+∞
<i>Tarea riesgo 2</i>											
1	0,9	45	0,1	100	0,9	19	0,1	200	-13,4	-∞	-0,55
2	0,8	45	0,2	100	0,8	19	0,2	200	-0,8	-0,55	-0,03
3	0,7	45	0,3	100	0,7	19	0,3	200	11,8	-0,03	0,32
4	0,6	45	0,4	100	0,6	19	0,4	200	24,4	0,32	0,60
5	0,5	45	0,5	100	0,5	19	0,5	200	37	0,60	0,86
6	0,4	45	0,6	100	0,4	19	0,6	200	49,6	0,86	1,12
7	0,3	45	0,7	100	0,3	19	0,7	200	62,2	1,12	1,40
8	0,2	45	0,8	100	0,2	19	0,8	200	74,8	1,40	1,73
9	0,1	45	0,9	100	0,1	19	0,9	200	87,4	1,73	2,24
10	0	45	1	100	0	19	1	200	100	2,24	+∞
<i>Tarea riesgo 3</i>											
1	0,9	110	0,1	125	0,9	8	0,1	250	-79,3	-∞	-1,86
2	0,8	110	0,2	125	0,8	8	0,2	250	-56,6	-1,86	-1,03
3	0,7	110	0,3	125	0,7	8	0,3	250	-33,9	-1,03	-0,53
4	0,6	110	0,4	125	0,6	8	0,4	250	-11,2	-0,53	-0,16
5	0,5	110	0,5	125	0,5	8	0,5	250	11,5	-0,16	0,15
6	0,4	110	0,6	125	0,4	8	0,6	250	34,2	0,15	0,44

► Decisión	Lotería A					Lotería B				Dif. VE (B - A)	Intervalo abierto CRRA	
	<i>n</i>	<i>p</i>	<i>x</i>	1- <i>p</i>	<i>x'</i>	<i>p</i>	<i>y</i>	1- <i>p</i>	<i>y'</i>			
7		0,3	110	0,7	125	0,3	8	0,7	250	56,9	0,44	0,72
8		0,2	110	0,8	125	0,2	8	0,8	250	79,6	0,72	1,03
9		0,1	110	0,9	125	0,1	8	0,9	250	102,3	1,03	1,46
10		0	110	1	125	0	8	1	250	125	1,46	+∞

Consideramos que el número de listas que utilizamos es el adecuado para obtener estimados precisos, ya que amplía el dominio de cantidades de dinero a las que se confronta al sujeto, por lo que nos brinda mayor variabilidad para poder obtener estimados de aversión al riesgo generales, que no sean dependientes de un conjunto específico de montos de dinero. Además, nos sentimos seguros de que los montos de dinero utilizados son relevantes para los sujetos, ya que los hemos elegido teniendo en cuenta también sus posibles ingresos.¹⁵

2.2.3. Otras consideraciones

Modificación de las tareas de tiempo

Tras recoger información de 968 individuos, se decidió modificar las cantidades de dinero dentro de las tareas de tiempo. Debido a que la mayoría de los individuos en las tareas originales siempre elegían la opción cercana y muy pocos realizaban un salto a la opción más

15 El salario mensual promedio de los encuestados es S/1626, mientras que la cantidad promedio que podría recibir el individuo por participar es un poco más de S/100 y la cantidad máxima es S/250. Estos últimos montos representan, respectivamente, un decimosexto y un sexto del salario promedio, por lo que pueden considerarse cantidades relevantes para el individuo representativo dentro del universo objetivo.

lejana, se tomó la decisión de ofrecer alternativas en las que la opción lejana brindara una tasa de interés mucho mayor, con la finalidad de que el participante tuviera mayores incentivos para realizar un salto de la alternativa *A* a la *B* dentro de las 10 decisiones que se ofrecen por cada tarea. En efecto, mientras la tasa de interés efectiva anual máxima en las tareas originales alcanzaba el 100%, la tasa de interés máxima en las tareas modificadas alcanzaba aproximadamente el 2000%. La lógica detrás de estas nuevas tareas es la misma que la de las originales, solo cambian las cantidades de dinero en juego. Los participantes que realizaron las tareas de tiempo con estas modificaciones fueron 330. Debido a que esta modificación fue imprevista en un inicio, no podemos asegurar que la muestra a la que se le aplicaron las tareas de tiempo modificadas sea representativa; no obstante, esto no impide que cumplan la función de mejorar la precisión de las preferencias agregadas, como veremos más adelante.¹⁶ El detalle de estas tareas puede verse en el cuadro A.1 en el anexo A.

2.3. Procedimiento

Con un equipo de 8 encuestadores, se recogió información acerca de 1298 personas en las ciudades de Lima Metropolitana, Arequipa, Trujillo y Tarapoto desde agosto hasta octubre del año 2017. La muestra que utilizamos se construyó sobre la base de un muestreo aleatorio hecho específicamente para este estudio. Se buscó una muestra representativa del grupo de trabajadores formales y estudiantes universitarios y técnicos en la población urbana del país, por lo que se escogieron cuatro ciudades debido al alto porcentaje que representan de

¹⁶ Las tareas modificadas de tiempo solo las realizaron 314 encuestados en Lima y 16 en Arequipa.

la población total urbana y a su distribución geográfica. Además, se aplicó un filtro antes del recojo de información, que detallamos en la siguiente subsección.¹⁷ Según las respuestas al momento del filtro, los encuestados que lo pasaron fueron divididos en cuatro categorías: (i) empleados formales, (ii) desempleados que eran formales hace menos de tres meses, (iii) estudiantes de pregrado y (iv) estudiantes de posgrado. En las cuatro categorías, se exigió que el entrevistado tuviera de 20 a 50 años.¹⁸

Se usaron *tablets* para recoger la información. La interfaz fue desarrollada para iOS, usando iOS SDK, y se distribuyó vía TestFlight. La aplicación incluía una encuesta sobre características laborales, financieras y sociodemográficas.

Las instrucciones de las tareas experimentales fueron brindadas por los encuestadores y estaban en la pantalla de la aplicación. Para optimizar la comprensión de las reglas de los juegos y para ilustrar la compatibilidad de incentivos del procedimiento de pago, se estableció una serie de tareas de práctica en las que los participantes jugaban usando dinero hipotético y caramelos.¹⁹ Los encuestadores fueron entrenados para responder preguntas siguiendo un protocolo bien definido. Para evitar que quienes tomaban decisiones tuvieran el sentimiento de ser observados, los encuestadores no tenían permitido mirar las decisiones de los jugadores tras las rondas de práctica.

17 La muestra es representativa en el nivel de la población urbana nacional, pero no necesariamente en el nivel de ciudades. Solo la muestra de Lima Metropolitana es representativa con un nivel de confianza del 95% y un error de muestreo de aproximadamente el 3%. Con el mismo nivel de confianza, los errores de muestreo para Arequipa, Trujillo y Tarpoto son 9%, 8% y 10%, respectivamente.

18 Hubo una excepción en la que se permitió a un participante de 19 años.

19 En la sección B en los anexos, describimos detalladamente en qué consistieron las rondas de práctica y hacemos un breve análisis del positivo nivel de entendimiento que demuestran los sujetos.

Como agradecimiento por su participación, se entregó a todos los individuos un polo, una gaseosa y un lapicero. Además, para incentivarlos a tomar las decisiones que más les convinieran, todos tuvieron la oportunidad de recibir un pago monetario adicional, dependiendo de las decisiones que tomaran. En específico, cada sujeto tuvo un 10% de probabilidad de recibir un pago por una de las 60 decisiones que tomó. Condicional a formar parte de este 10%, la decisión que sería pagada fue determinada de manera aleatoria al final de todas las tareas y luego de realizar la encuesta sociodemográfica. Este método de compensación –con incentivos materiales– hace que el participante se encuentre incentivado a considerar cada una de sus decisiones y revele sus verdaderas preferencias honestamente en el marco de EUT y QHD.

2.4. Características de la muestra

En principio, las características del encuestado que se requirieron en este estudio fueron que tuviera entre 20 y 50 años y estuviera trabajando en ese momento en un empleo formal o lo hubiera estado haciendo hasta hace poco –no más de 3 meses–, o estuviera estudiando para poder acceder a ese empleo formal.²⁰ Asimismo, si es que trabajaba, tenía que hacerlo en instituciones privadas o empresas públicas –distintas de las instituciones de administración pública– y, en caso de que su posición en su ocupación principal fuera la de empleador, tendría que ser asalariado para poder ser encuestado.²¹ Aparte de estos requerimientos,

20 Para identificar si el individuo era un trabajador formal, se usaron sus respuestas relacionadas con su seguro de salud y con el número de empleados de la empresa en la que trabajaba.

21 Si el individuo reportaba que su posición en su ocupación principal era la de empleador, ya no se usaban sus respuestas sobre seguro de salud y número de empleados, sino el hecho de que fuera asalariado o no.

no hubo otras características que restringieran la muestra, salvo el hecho de que el muestreo se hizo en las ciudades de Lima Metropolitana, Arequipa, Trujillo y Tarapoto. Un 70,3% de los encuestados vivían en Lima Metropolitana; un 10,4%, en Arequipa; un 11,6%, en Trujillo; y un 7,7%, en Tarapoto.

Resumimos algunas características sociodemográficas de la muestra de estudio en el cuadro 2.3. En la primera columna, pueden verse las medias y las desviaciones estándar de las características de la muestra total; en la segunda, se reporta lo mismo solo para la muestra de Lima; en la tercera, solo para la muestra de las tres otras provincias; y en la cuarta, se reporta la diferencia de medias entre la muestra de Lima y la de las otras provincias. Vale la pena hacer notar que un 20% de los individuos se encontraban desempleados al momento de la encuesta, un 64% contaban con educación superior, ligeramente menos de la mitad eran hombres y la edad promedio era de casi 33 años. También es interesante observar que un 90% de los encuestados esperaba que su situación económica mejorara en un plazo de un año y que aquellos que estaban empleados creían que la probabilidad de perder su empleo era muy baja. Sobre esto último, les pedimos que, en una escala del 1 al 10, indicaran cuán probable era que fueran despedidos o no se les renovara el contrato en un plazo de seis meses, y la respuesta promedio fue 2. Finalmente, encontramos que ambas muestras –Lima y las demás provincias– fueron muy similares en términos de las características resumidas en la tabla. De hecho, solo resalta que el porcentaje de individuos con educación superior en las tres provincias sea 7 puntos porcentuales mayor, y lo mismo para el indicador de optimismo. En cambio, en Lima la propensión de tener acceso a internet es 10 puntos porcentuales mayor y los trabajadores limeños suelen trabajar alrededor de 3 horas semanales adicionales.

Cuadro 2.3
Características de la muestra

Variables	(1)	(2)	(3)	Diferencia
	Total	Lima	Provincias	(2)-(3)
Desempleado	0,20 (0,40)	0,20 (0,40)	0,21 (0,41)	0,01 (0,02)
Educación superior	0,64 (0,48)	0,62 (0,49)	0,69 (0,46)	0,07* (0,03)
Sexo (hombre = 1)	0,48 (0,50)	0,48 (0,50)	0,48 (0,50)	-0,00 (0,03)
Sin liquidez	0,41 (0,49)	0,40 (0,49)	0,44 (0,50)	0,04 (0,03)
Edad	32,86 (9,57)	33,10 (9,57)	32,30 (9,55)	-0,79 (0,58)
Casado	0,21 (0,41)	0,20 (0,40)	0,23 (0,42)	0,04 (0,03)
Optimista	0,90 (0,30)	0,88 (0,32)	0,95 (0,22)	0,07*** (0,02)
Seguro privado de salud	0,12 (0,32)	0,12 (0,33)	0,11 (0,31)	-0,01 (0,02)
AFP	0,63 (0,48)	0,63 (0,48)	0,64 (0,48)	0,01 (0,03)
Número de personas en el hogar	4,78 (1,94)	4,75 (1,95)	4,83 (1,90)	0,08 (0,12)
Tiene carro	0,15 (0,35)	0,15 (0,36)	0,13 (0,34)	-0,02 (0,02)
Tiene internet	0,59 (0,49)	0,62 (0,49)	0,52 (0,50)	-0,10*** (0,03)
Vivienda de material noble	0,93 (0,25)	0,93 (0,25)	0,93 (0,25)	0,00 (0,02)
Ingreso laboral neto	1626,35 (1200,37)	1650,03 (1291,29)	1573,10 (965,53)	-76,93 (84,75)
Horas de trabajo	47,32 (11,53)	48,18 (11,83)	45,28 (10,53)	-2,89*** (0,84)

Variables	(1)	(2)	(3)	Diferencia
	Total	Lima	Provincias	(2)-(3)
Probabilidad subjetiva de despido	2,00 (1,78)	2,07 (1,81)	1,82 (1,70)	-0,25 (0,13)
Número de individuos	1298	913	385	1298

Nota: Se reporta la media y la desviación estándar entre paréntesis. En la última columna, se reportan las diferencias y el error estándar resultante de una diferencia de medias entre los datos de Lima y provincias. La variable “sin liquidez” es una variable dicotómica igual que 1 para aquellos que reportan tener apenas efectivo para solventarse hasta fin de mes. La variable “optimista” es igual que 1 para quienes esperan que su situación económica mejore en un año. La probabilidad subjetiva de despido es un indicador entre 1 y 10, que representa la probabilidad que cree el individuo de que perderá su empleo durante los próximos seis meses.

3. MODELOS ESTOCÁSTICOS Y ESTRATEGIAS EMPÍRICAS

Hasta este capítulo, se han tratado las teorías de decisión desde una perspectiva determinística; es decir, que tratan la decisión como un proceso exento de error o inconsistencia y, por ende, completamente predecible. Nos referimos a las teorías dominantes en la literatura económica que hemos tratado en el capítulo 1, que son EUT y QHD.²² Sin embargo, las elecciones bajo riesgo e intertemporales no solo tienen un componente determinístico, sino también uno estocástico.²³ Es decir, toda decisión que un individuo tome va a ser influenciada por cuestiones aleatorias, en menor o mayor medida. Algunas fuentes de aleatoriedad son (i) preferencias que fluctúan aleatoriamente, (ii) señales ruidosas, (iii) falta de atención, (iv) gusto por aleatorizar en ciertas personas y (v) equivocaciones al momento de decidir.

La evidencia experimental es clara al mostrar que las decisiones individuales incluyen cierta naturaleza estocástica. Por ejemplo, en el estudio de Agranov y Ortoleva (2017), el 90% de los sujetos escogió diferentes loterías al ser presentado con tres repeticiones distantes del mismo problema de decisión y, cuando las repeticiones fueron seguidas,

22 Tras la crítica a EUT de Allais (1953), la iniciativa de investigación teórica y empírica se centró en proponer modelos de decisión alternativos entre los que destacan la teoría prospectiva, la teoría de la utilidad dependiente del rango y la teoría del arrepentimiento.

23 Dos referencias clásicas del estudio de preferencias usando decisiones discretas son McFadden (2001) y Train (2009).

71% escogieron de manera inconsistente.²⁴ En nuestra muestra, podemos construir dos indicadores de inconsistencia para tener una idea del componente estocástico en nuestros datos. El primer indicador es la proporción de sujetos que revelan una reversión de preferencias; es decir, aquellos que cambian a la alternativa *A* una vez que ya han elegido la *B* en dicha lista. Esto es inconsistente ya que, bajo nuestros supuestos de agente racional, mientras más preguntas se hagan en la lista hay mayores incentivos para elegir la opción *B*. El segundo indicador es la proporción de sujetos que eligen una alternativa estocásticamente dominada. En nuestro diseño experimental, esto implica elegir la opción *A* en la última decisión de alguna lista de tareas de riesgo. Notemos que hacer esto es irracional ya que, cuando el pago mayor es certero, la opción *B* siempre es mayor que la opción *A*. Así, pues, la proporción de sujetos que revierten sus preferencias al menos una vez es 10,8% y la proporción de sujetos que eligen alternativas estocásticamente dominadas al menos una vez es 12,9%; la proporción que comete alguna de ambas fuentes de inconsistencia es 20,7%.^{25 26 27} Estos son indicadores importantes que representan límites inferiores en la inconsistencia o naturaleza estocástica de las preferencias en los sujetos de nuestra muestra.²⁸

24 Este fenómeno ocurre en intervalos de tiempo de días (Tversky, 1969; Hey y Orme, 1994) y también en intervalos de tiempo más cortos (Starmer y Sugden, 1989; Camerer, 1989; Ballinger y Wilcox, 1997).

25 En específico, las proporciones de tareas de riesgo y tiempo que presentan reversión de preferencias son 4,6% y 3,4%, respectivamente.

26 La submuestra de individuos que presenta al menos una reversión de preferencias es similar a la submuestra que no presenta ninguna. Solo observamos dos diferencias estadísticamente significativas mediante una prueba de medias en la que usamos las mismas variables que en el cuadro 2.3. Los que no presentan ninguna reversión de preferencias son mayores por 1,74 años y reportan un indicador de probabilidad subjetiva de despido menor por un punto (sobre 10).

27 En esta misma línea, Galarza (2009) realiza un experimento similar al nuestro en el Perú para medir actitudes ante el riesgo y encuentra que un 52% de los sujetos revierten sus preferencias.

28 Por ejemplo, si somos más exigentes, podríamos señalar que otra señal de inconsistencia es que los sujetos revelen preferencias distintas en diferentes listas, aunque esto también

La ubicuidad de este fenómeno ha impulsado a la teoría económica a incorporar elementos estocásticos para intermediar entre las decisiones erráticas de la vida real y el determinismo inherente a nuestras teorías. Existen varias formas de concebir cómo la aleatoriedad afecta la decisión y cada una resulta en una teoría diferente que afecta los valores estimados de los parámetros de la función de utilidad subyacente de manera sistemática (Loomes y Sugden 1995). Por tanto, hacemos la distinción entre los modelos amparados en teorías determinísticas —que ya introdujimos en el capítulo anterior—, a los que llamamos *modelos determinísticos*, y los modelos que describen el componente estocástico de la decisión, a los que llamamos *modelos estocásticos*, usando la notación de Wilcox (2008).

Un aspecto valioso de los modelos estocásticos es que, al modelar lo que no puede ser explicado —o el residuo—, conforman el puente faltante para armar las estrategias econométricas que nos permiten identificar las preferencias de la muestra de individuos. En particular, con cada modelo estocástico vamos a ser capaces de llegar a una expresión que describa la probabilidad de que el sujeto tome una elección determinada. Una vez que tenemos esta función de probabilidad —o verosimilitud— por cada modelo, el siguiente paso es utilizar técnicas econométricas de máxima verosimilitud para minimizar las desviaciones del modelo e identificar los parámetros de preferencias.

Este capítulo se desarrolla de la siguiente manera. En la sección 3.1, motivamos los modelos estocásticos asumiendo que existe un agente representativo que toma las decisiones, y discutimos algunos elementos centrales en la predicción de decisiones de cada uno.²⁹ El

podría ser explicado por grados de aversión al riesgo o de paciencia heterogéneos según el monto de dinero en juego, las probabilidades o el tiempo de espera en cada lista.

29 Hasta el momento, la literatura no ha establecido nombres convencionales para los modelos estocásticos y pueden ser referidos de manera contradictoria por diferentes investigadores. Por tanto, usamos el argot econométrico también para evitar confusiones.

modelamiento estocástico es de vital importancia para la identificación de las preferencias de los sujetos, pues describe la forma del componente estocástico o término de error que es necesario para plantear las estrategias empíricas. Por eso, al describir cada modelo, también aprovechamos en derivar la función de verosimilitud asociada a este. Luego, en la sección 3.2, extendemos los modelos estocásticos para que tomen en consideración la heterogeneidad observable y no observable de las preferencias en la muestra.

3.1. Modelos estocásticos

Los modelos estocásticos describen la manera en que la aleatoriedad asume un rol en las decisiones que toman los individuos. Las teorías determinísticas que detallamos antes, como EUT y QHD, pueden ser capaces de explicar una parte importante de las decisiones, pero es irreal pensar que van a poder predecirlas totalmente. En esta línea, los modelos estocásticos se encargan de dar forma a aquello que no puede ser explicado. Puede añadirse un componente estocástico en cada una de las fases del proceso de decisión que el sujeto confronta. Dependiendo de qué fase sea, los supuestos de cómo se toma la decisión cambian y, por ende, también el modelo.

Existen tres fases en el proceso de decisión, y podemos asociar cada una de ellas con tres aplicaciones estocásticas que veremos en este estudio. La primera fase corresponde a la identificación, que es cuando el sujeto identifica cuál es su preferencia antes de tomar una decisión, y está relacionada con el modelo estocástico de *parámetros aleatorios*, descrito en la subsección 3.1.1. La segunda fase es la de evaluación, en la que el sujeto calcula las utilidades esperadas de cada una de las loterías entre las que puede escoger y determina cuál le reporta

mayor utilidad. Incorporar el componente estocástico en esta fase se relaciona con la familia de modelos de *utilidad aleatoria*, descrita en la subsección 3.1.2. Finalmente, la tercera fase en la que se puede incorporar la aleatoriedad es la de ejecución, cuando el sujeto implementa el cálculo de utilidad para escoger una opción. Si se añade un componente estocástico en esta fase, podemos relacionarlo con añadir un parámetro de *tiemble*, que se describe en la subsección 3.1.3. El componente estocástico puede entrar en cualquiera de las tres fases y en más de una a la vez; estas no son mutuamente excluyentes. En este estudio, nos limitaremos a añadir un componente de temble a los modelos de parámetros aleatorios o de utilidad aleatoria, el cual detallaremos más adelante.

Modelación estructural

Los modelos estructurales se refieren al tipo de modelos que usamos para estimar los parámetros de una función que presuntamente genera los datos observados. Es decir, están expresados en términos de la función de utilidad del individuo (Moffatt, 2015). Para nuestros fines, esta clase de modelos de probabilidad tiene la forma

$$P(B, A) = F(\cdot), \quad (3.1)$$

donde $P(B, A)$ es la probabilidad de escoger la opción B entre las alternativas A y B . La función de enlace F es creciente en los números reales y está en el intervalo cerrado entre 0 y 1; $F(0) = 0,5$ y $F(x) = 1 - F(-x)$. Por lo general, asumimos que esta es la función normal o la función logística. El contenido de $F(\cdot)$ es una representación de la variable latente de utilidad que determina la decisión. Para este trabajo,

asumimos que las funciones de utilidad latentes tienen forma CRRA y QHD, y que los sujetos maximizan su utilidad según EUT.

3.1.1. Modelos de parámetros aleatorios

El modelo de parámetros aleatorios fue originalmente propuesto por Becker et al. (1963), desarrollado para la elección bajo riesgo por Loomes y Sugden (1995) e implementado por Loomes et al. (2002). Este modelo asume que los parámetros dentro de la función de utilidad del sujeto varían estocásticamente al momento de decidir. A pesar de tener incertidumbre acerca de sus preferencias, los sujetos toman decisiones coherentes con las propiedades generales del modelo determinístico de decisión asumido. El punto clave de este modelo es que las decisiones se basan en realizaciones aleatorias de una distribución de parámetros de preferencias de media m y desviación estándar σ . En el caso de actitudes ante el riesgo, se asume que la variación aleatoria se produce en el coeficiente de aversión al riesgo (r); y en el caso de elección intertemporal, en la función de descuento.

Ilustraremos las características del modelo sobre la base de las preferencias ante el riesgo, pero es sencillo construir el análogo para las preferencias intertemporales. En particular, se empieza por asumir que la probabilidad de que el sujeto elija la alternativa B es igual que la probabilidad de que la utilidad esperada que le genera dicha alternativa sea mayor que la utilidad esperada que le genera la alternativa A ; es decir

$$P(B, A) = P(U(B) > U(A)) = P(U(B) - U(A) > 0). \quad (3.2)$$

Para que esto ocurra, entonces el parámetro de preferencia debe cumplir ciertas condiciones. En el caso de elección ante el riesgo, r debe ser tal que

$$p\bar{x}_B^{1-r} + (1-p)\underline{x}_B^{1-r} - p\bar{x}_A^{1-r} - (1-p)\underline{x}_A^{1-r} > 0, \quad (3.3)$$

donde p es la probabilidad asociada al monto mayor de las loterías; \bar{x}_B y \underline{x}_B son los montos mayor y menor de la lotería de la alternativa riesgosa B , respectivamente; y \bar{x}_A y \underline{x}_A son los montos mayor y menor de la lotería de la alternativa segura A , respectivamente. Si denominamos r^* al valor de r que permite que la desigualdad en (3.3) se cumpla, entonces

$$P(B, A) = P(r < r^*). \quad (3.4)$$

La expresión anterior se interpreta como que el sujeto escogerá la alternativa riesgosa si es que su grado de aversión al riesgo es menor que el que implica indiferencia entre las dos alternativas. El punto clave es que, si suponemos que r es una variable aleatoria, entonces para un individuo particular puede cumplirse que

$$r \sim N(m, \sigma^2). \quad (3.5)$$

Así pues, usando las expresiones (3.4) y (3.5), se tiene que

$$P(B, A) = \Phi\left(\frac{r^* - m}{\sigma}\right), \quad (3.6)$$

donde $\Phi(\cdot)$ es la función de densidad acumulada normal.

Una forma de entender el modelo es que para cada sujeto existe una urna con coeficientes de aversión al riesgo y , al tomar una decisión entre loterías, se realiza uno de los coeficientes de la urna aleatoriamente. En términos prácticos, el modelo asume que los parámetros dentro de la función de utilidad varían estocásticamente al momento de decidir y el individuo utiliza esta realización para calcular la utilidad correspondiente. Por ende, las decisiones que no se puedan

racionalizar bajo ningún valor de los parámetros, condicional a una función de utilidad, tienen probabilidad nula de ser escogidas con este modelo. En otras palabras, el modelo de parámetros aleatorios es incapaz de explicar violaciones de dominancia estocástica y esto puede ser visto como una limitación de este. Ante esto, como sugiere Moffatt (2015), una solución suele ser incorporar un parámetro de tiemble, que será explicado en la subsección 3.1.3.

3.1.2. Modelos de utilidad aleatoria

La otra familia de modelos estocásticos es la de utilidad aleatoria. En estos modelos, se asume que el componente estocástico se incorpora en la fase de *evaluación* de las utilidades de cada lotería, cuando el individuo compara las utilidades que le proveen las alternativas disponibles. Este tipo de modelos de decisión estocástica difieren de los de parámetros aleatorios en cuanto a su interpretación. En los modelos de utilidad aleatoria, el sujeto asume una actitud o preferencia –por ejemplo, el coeficiente de aversión al riesgo o la tasa de descuento– única. Esto implica que los parámetros que estima este tipo de modelos ya no pertenecen a una distribución, sino que son números fijos. Son las utilidades por comparar, y no los parámetros, las que contienen el componente aleatorio; por ello el nombre que reciben estos modelos. Cuando el sujeto se enfrenta a una decisión, evalúa la utilidad que le provee cada lotería directamente, pero su decisión es estocástica, ya que la comparación de utilidad está sujeta a un término de error aditivo.

Modelo de Fechner

En el modelo de Fechner –también conocido como modelo de *utilidad fuerte*–, se introduce un término de error aditivo ε en el diferencial de

utilidades que el individuo evalúa para decidir su elección. En particular, la condición para que el individuo elija la alternativa B es que

$$U(B) - U(A) + \varepsilon > 0, \tag{3.7}$$

donde $\varepsilon \sim N(0, \sigma^2)$.³⁰

Este modelo ha sido nombrado en honor a Fechner (1860), quien lo desarrolló para sus investigaciones en psicofísica, pero fue adaptado al campo de la teoría de decisión por Becker et al. (1963). Este modelo de utilidad aleatoria ha sido utilizado desde los inicios del análisis de la decisión económica en aplicaciones empíricas (Manski, 1977), pero Hey y Orme (1994) lo popularizaron para medir la aversión al riesgo. Desde entonces, ha sido empleado para estimar las preferencias de toda clase de teorías centrales de decisión.

Para derivar la probabilidad de escoger la alternativa B , asumimos una función de enlace normal $\Phi(\cdot)$:

$$P(B, A) = P(U(B) - U(A) + \varepsilon > 0) = P(\varepsilon > U(A) - U(B)) = P\left(\frac{\varepsilon}{\sigma} > \frac{U(A) - U(B)}{\sigma}\right) = 1 - \Phi\left(\frac{U(A) - U(B)}{\sigma}\right) = \Phi\left(\frac{U(B) - U(A)}{\sigma}\right) \tag{3.8}$$

El parámetro σ puede ser entendido como el ruido en la decisión. Si $\sigma = 0$, no hay ruido y la elección es determinística; si $\sigma = \infty$, la elección se determina enteramente por el ruido y ambas alternativas tienen probabilidad 0,5 de ser escogidas (Moffatt, 2015).

30 Esta condición es lo mismo que afirmar que $U(B) + \varepsilon(B) > U(A) + \varepsilon(A)$, de lo cual se entiende con mayor claridad la intención del nombre *utilidad aleatoria*.

Modelo de Luce

Otro modelo de utilidad aleatoria comúnmente empleado es el de Luce –también conocido como el modelo de *utilidad estricta*–. Este modelo estocástico, muy común en la literatura de microeconomía, fue popularizado por Holt y Laury (2002) para estimar el nivel de aversión al riesgo y ha sido usado en estudios influyentes como el de Andersen et al. (2008). En este modelo, la probabilidad de escoger la alternativa B sobre la alternativa A se expresa de la siguiente forma:

$$P(B, A) = \frac{U(B)^{1/\sigma}}{U(A)^{1/\sigma} + U(B)^{1/\sigma}}, \quad (3.9)$$

donde σ es nuevamente un parámetro de precisión, o ruido.

Como han demostrado Luce y Suppes (1965), este modelo es equivalente al de Fechner con una función de enlace *logit* y al reemplazar la diferencia de las utilidades por la diferencia de sus logaritmos naturales.³¹ Este modelo es similar al de Fechner en el sentido de que el término estocástico entra al modelo sumándose al diferencial de utilidades o, en este caso, al diferencial logarítmico de utilidades, como lo llama Wilcox (2008), y por eso forma parte de los modelos de utilidad aleatoria.

A diferencia del modelo de parámetros aleatorios, los modelos de utilidad aleatoria sí admiten violaciones de dominancia estocástica y reversión de preferencias con probabilidad positiva. Esto ocurre porque estos últimos modelos presentan el término estocástico sumado al diferencial de utilidades, por lo que, aunque el valor asignado a una alternativa sea estrictamente menor que el otro, el término estocástico

31 La expresión (3.9) es equivalente a $P(B, A) = \Lambda\left(\frac{\ln(U(B)) - \ln(U(A))}{\sigma}\right)$, donde $\Lambda(\cdot)$ es la función de enlace *logit*.

siempre va a poder suplir este déficit y explicar la decisión del individuo. Dado que estos fenómenos están presentes en nuestros datos, estos últimos modelos son descriptivamente aptos sin necesidad de incluir un parámetro de tiemble.

Apesteagua y Ballester (2018) critican este tipo modelos aplicados a ciertas formas funcionales, como CRRA bajo EUT, para estimar funcionales de preferencias. Los autores argumentan que existe un problema de identificación que implica que el parámetro de preferencia pueda ser subestimado. Por ejemplo, para un σ dado, los incrementos en el parámetro de aversión al riesgo r inicialmente conllevan una mayor probabilidad de escoger la alternativa segura. Ahora, según r se incrementa, la diferencia en la utilidad de las alternativas disminuye, y con ello la probabilidad de elección entre alternativas se acerca a la aleatoriedad –y se aleja de la lotería segura–. Este problema se exagera cuando se estiman modelos con heterogeneidad entre sujetos en las preferencias de riesgo –con σ homogéneo– (Barseghyan et al., 2018); por ejemplo, asumiendo EUT y CRRA, aquellos con mayor aversión al riesgo escogerían entre las loterías de manera más aleatoria.

Modelo de Blavatskyy (o de utilidad más fuerte)

En la tradición de los modelos de utilidad aleatoria, el último modelo que desarrollamos es el propuesto por Blavatskyy (2014), llamado también modelo de *utilidad más fuerte*. Este modelo realiza modificaciones que aplican específicamente para explicar la elección ante incertidumbre. En efecto, el modelo de Blavatskyy modifica al de Fechner truncando el término de error ξ superior e inferiormente, de la siguiente manera:

$$\xi \leq | U(A \vee B) - U(A \wedge B) |, \quad (3.10)$$

donde $A \vee B$ es la menor cota superior y $A \wedge B$ es la mayor cota inferior de las loterías A y B en términos de dominancia estocástica de primer orden.³² Puede expresarse $\xi = \varepsilon^* [U(A \vee B) - U(A \wedge B)]$, donde $\varepsilon \sim N(0, \sigma^2)$; de esta forma, puede verse que ε es un término de error heterocedástico que varía según la decisión que confronta el sujeto.

Al truncar el error de esta forma, el modelo sigue racionalizando el fenómeno de cambio de preferencias, pero ya no admite violaciones de dominancia estocástica, que sí admiten los modelos de Fechner y Luce. Ante esto, dado que en nuestros datos sí se producen casos de violación de dominancia estocástica, será necesario incluir un parámetro de tiemble, como se verá en la próxima sección. Un punto a favor de esta modificación es que resguarda a los modelos de utilidad aleatoria con error homocedástico —descritos anteriormente— de la crítica de Apesteagua y Ballester (2018): el truncar el término de error heterocedástico no permite que adquiera valores que aumenten la diferencia cardinal de las utilidades entre loterías mientras el parámetro de aversión al riesgo r se incrementa.

Dada esta modificación, la probabilidad de escoger la lotería B se expresa como:

$$P(B, A) = \begin{cases} \Phi \left(\frac{U(B) - U(A)}{\sigma [U(A \vee B) - U(A \wedge B)]} \right), & \text{si } U(A \vee B) \neq U(A \wedge B) \\ 0.5, & \text{si } U(A \vee B) = U(A \wedge B) \end{cases} \quad (3.11)$$

donde $\Phi(\cdot)$ es la función de enlace normal.

32 Sea (\bar{x}, \underline{x}) una lotería cuyo premio mayor sea \bar{x} y el menor sea \underline{x} , si un sujeto se enfrenta a una elección entre las loterías $A(x_1, x_2)$ y $B(x_3, x_4)$ que pagan los premios con las mismas probabilidades, la menor cota superior es la lotería (x_3, x_2) y la mayor cota superior es la lotería (x_1, x_4) .

3.1.3. *Parámetro de tiemhle*

Incluir un parámetro de tiemhle permite racionalizar la posibilidad de que, por una variedad de razones, el individuo tome una decisión aleatoriamente. Es concebible que una persona no haya entendido el problema al que se enfrenta o tenga un lapsus de concentración al tomar una decisión. Este lapsus o *tiemhle* resulta en que la toma de decisión no esté conectada con la preferencia *verdadera*, ya sea asumiendo que existe una preferencia única o que sea una realización aleatoria de una distribución. Este tipo de error se puede modelar asumiendo que, para todas las decisiones a las que se enfrenta un sujeto, este elige de manera aleatoria entre las alternativas con probabilidad $w \in [0,1]$ o busca escoger la alternativa que le reporte mayor utilidad con la probabilidad complementaria $1-w$.

Siguiendo a Loomes et al. (2002), la probabilidad de escoger la alternativa B se determina por:

$$P(B, A) = (1-w) P(B, A \mid \text{sin tiemhle}) + \frac{w}{2}, \quad (3.12)$$

donde se reemplaza $P(B, A \mid \text{sin tiemhle})$ por alguna de las probabilidades de escoger B sobre A que han sido detalladas en las subsecciones anteriores. Este tipo de error es útil para racionalizar las violaciones estocásticas de primer orden –por ejemplo, escoger la opción segura en la última decisión en el experimento de riesgo–, ya que estas implican grandes divergencias sobre los posibles valores de las funciones de utilidad.³³

33 Es importante mencionar que, en todos los modelos que incluyen un parámetro de tiemhle, forzamos que este parámetro explique las violaciones estocásticas de primer orden en las últimas decisiones del experimento de riesgo.

3.2. Preferencias heterogéneas

3.2.1. Según características no observables

Como también hacen Andersen et al. (2008) y sugiere Moffatt (2015), es posible extender los modelos discutidos antes de tal manera que se permita heterogeneidad entre sujetos según características no observables. En efecto, si se asume que

$$\begin{aligned} r &\sim N(\mu_r, \eta_r^2) \\ \delta &\sim N(\mu_\delta, \eta_\delta^2), \end{aligned} \tag{3.13}$$

estamos permitiendo heterogeneidad explicada por características no observables entre sujetos. Los modelos explicados antes, sin permitir heterogeneidad no observable, pueden ser racionalizados dentro de este como si $\eta_r = \eta_\delta = 0$. Es también pertinente mencionar que las distribuciones supuestas en los modelos homogéneos son en el nivel del individuo, mientras que las distribuciones en esta sección son en el nivel del grupo total de sujetos.³⁴ Es decir, en los modelos anteriores, se asumía que un agente representativo tomaba una decisión con un componente estocástico; ahora, además de este componente estocástico, se permite que exista una gama de preferencias diversas entre los individuos, que es capturada por η_r y η_δ .

Estas estimaciones pueden llevarse a cabo utilizando métodos de máxima verosimilitud simulada, que consisten en simular desviaciones individuales en los parámetros de preferencias con la finalidad de tomar en cuenta las características de las distribuciones de las que

³⁴ Así, pues, para el modelo de parámetros aleatorios es más preciso considerar que $r_{it} \sim N(m_i, \sigma^2)$; y para los modelos de utilidad aleatoria, que $\varepsilon_{it} \sim N(0, \sigma^2)$, donde i hace referencia al individuo y t a la decisión.

proviene estos parámetros. Este mecanismo también puede ser visto como incluir efectos aleatorios en la estimación. El anexo C muestra un ejemplo de cómo se incorporan los efectos aleatorios y las consideraciones matemáticas que están detrás.

3.2.2. Según características observables

Otra forma de modelar la heterogeneidad es permitir que los parámetros de comportamiento que se estiman varíen según características observables. Una forma sencilla y efectiva de hacerlo es permitiendo que estos parámetros sean funciones lineales de un conjunto de variables observadas del individuo. En este sentido, podría especificarse que el parámetro CRRA siga la siguiente especificación:

$$\hat{r} = \hat{r}_0 + Z\hat{\gamma}, \quad (3.14)$$

donde \hat{r}_0 es el estimado de la constante y $\hat{\gamma}$ es un vector de coeficientes que muestra las diferencias en las actitudes hacia el riesgo según las características observables Z de los individuos. Para el caso de los estimados de tiempo, se puede modelar la heterogeneidad de δ y β .

Esta extensión no solo es importante en términos de predicción, sino que también permite asociar ciertas características de la población de estudio con actitudes hacia el riesgo e intertemporales diferenciadas. Conocer estas diferencias es muy importante para la efectividad de las políticas públicas. En específico, las medidas de protección social –cuya efectividad tiene estrecha relación con las preferencias por el riesgo y el consumo intertemporal– pueden ser muy efectivas en ciertos grupos poblacionales y no tanto en otros.

4. RESULTADOS

Hacemos uso de las estrategias detalladas en el capítulo anterior en búsqueda del modelo y los parámetros estructurales que mejor expliquen las decisiones bajo riesgo y tiempo de los trabajadores peruanos. En la siguiente sección, presentamos los estimados de preferencias ante el riesgo; y en la subsiguiente, los estimados de preferencias de elección intertemporal. En este capítulo, nos referimos a modelos de agente representativo como aquellos que asumen que las preferencias son homogéneas para todo el grupo de sujetos –aquellos de la sección 3.1 del capítulo anterior– y a modelos con efectos aleatorios a aquellos que asumen preferencias heterogéneas –explicados en la sección 3.2–.

4.1. Estimación de actitudes ante el riesgo

En esta sección comparamos y discutimos los parámetros estimados utilizando varias especificaciones que asumen que existe una preferencia que caracteriza a todos los sujetos, o modelos de agente representativo. Presentamos los resultados de tres modelos: parámetros aleatorios, Fechner y el de Blavatskyy. Como vimos en el capítulo anterior, los últimos dos modelos pertenecen a la familia de modelos de utilidad aleatoria. Luego, ampliamos los modelos permitiendo que exista heterogeneidad en las preferencias, tanto no observable como observable. A lo largo de esta y la próxima sección, utilizamos criterios

estadísticos para medir el valor de ajuste de los modelos; es decir, su capacidad predictiva a la muestra de estudio.

Una parte íntegra del objetivo de este trabajo es estudiar el rol de los errores estocásticos en los diferentes tipos de modelos. Para esto, primero estimamos la forma más simple de cada tipo de modelo, donde el parámetro de precisión es fijado a 1 y no hay error de tiembla. Progresivamente, incluimos estos errores. Discutimos el efecto de los errores en la magnitud del estimado y el ajuste del modelo a los datos.

4.1.1. Modelos de agente representativo

El cuadro 4.1 muestra los estimados de los tres modelos de agente representativo para riesgo. Los valores del parámetro de aversión al riesgo r que estimamos varían entre 0,775 y 1,134, y promedian 0,997. Según el modelo determinístico que usamos, un r mayor que 0 implica aversión al riesgo, y mientras mayor sea, más averso al riesgo es el sujeto.³⁵ Esto implica que, independientemente del modelo estocástico que elijamos, el trabajador peruano presenta un grado de aversión al riesgo moderado, y algo más alto que el de otras poblaciones estudiadas. Por ejemplo, Andersen et al. (2008), que usan una estrategia muy similar a la nuestra, encuentran que el CRRA promedio es igual a 0,774 para la población danesa, y Kim y Lee (2012) hallan uno igual a 0,8 en Corea del Sur. En países en desarrollo, De Brauw y Eozenou (2014) estiman un CRRA igual a 0,74 sobre una muestra de granjeros en Mozambique; Harrison et al. (2005) encuentran que el CRRA promedio es igual a 0,536 en aldeas rurales de la India, Etiopía y Uganda,

35 Todos los parámetros CRRA que estimamos son estadísticamente mayores que 0 con un nivel de confianza superior del 99%.

Cuadro 4.1
Estimados de riesgo con modelos de agente representativo

Modelo:	Parámetros aleatorios				Fechner			Blavatsky		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
r	0,937 (0,032)	1,085 (0,035)	1,134 (0,042)	1,074 (0,035)	1,092 (0,012)	0,903 (0,011)	0,775 (0,025)	0,911 (0,026)	1,031 (0,029)	1,031 (0,028)
σ			1,731 (0,043)	0,941 (0,031)			2,804 (0,283)	0,953 (0,103)	0,821 (0,019)	0,662 (0,020)
w		0,207 (0,012)		0,219 (0,011)		0,239 (0,015)		0,243 (0,012)	0,161 (0,011)	
LL	-22 799	-20 721	-20 957	-20 719	-22 082	-20 904	-21 523	-20 903	-20 735	-20 735
AIC	45 600	4 1447	41 918	41444	44 167	41 812	43 051	41 813	41 473	41 476
N	38 940	38 940	38 940	38940	38 940	38 940	38 940	38 940	38 940	38 940

Nota: Los errores estándar agrupados a nivel individual están entre paréntesis. Para el caso de las estimaciones con el modelo de parámetros aleatorios, los estimados de r en realidad corresponden a los estimados de m , la media de r . LL hace referencia al valor de la log-verosimilitud; AIC , al criterio de información de Akaike; y N , al número de observaciones.

mientras que Cassar et al. (2017) –en zonas rurales de Tailandia– y Galarza (2009) –en Ica, Perú– hallan coeficientes CRRA promedio de 0,65 y 0,45, respectivamente.³⁶ Gandelman y Hernández-Murillo (2015) asumen la misma función de utilidad que nosotros, aunque usan una estrategia empírica distinta, y encuentran que el valor del CRRA promedio en varios países –desarrollados y no desarrollados– es muy cercano a 1. Al igual que nosotros, hallan que el peruano es más averso al riesgo que el promedio.

Primero, describimos detalladamente los estimados de los modelos de parámetros aleatorios, que corresponden a las especificaciones de la (1) a la (4). Como ya discutimos, estos modelos asumen que el sujeto decide sobre la base de realizaciones aleatorias de un coeficiente de aversión al riesgo, que proviene de una distribución normal. Tenemos que el promedio del coeficiente de aversión al riesgo estimado toma el valor de 0,937 cuando asumimos que la varianza de r , o σ , es igual a 1; y 1,085 al añadir un parámetro de tiemble. Al incluir este último parámetro, encontramos que el sujeto representativo escoge una lotería de manera aleatoria con 20,7% de probabilidad. Al permitir que σ tome cualquier valor, y no necesariamente 1, hallamos un valor de la media del coeficiente de aversión al riesgo igual a 1,134, e igual a 1,074 al incluir la probabilidad de tiemble. Es interesante notar que al permitir que σ sea flexible, estimamos que toma el valor de 1,731, pero al añadir al modelo la probabilidad de que el sujeto elija aleatoriamente, este se reduce a 0,941, cercano a 1. Este es un patrón común en esta sección: al no permitir la posibilidad de tiemble, el valor de σ aumenta, pues absorbe parte del componente estocástico que le correspondería a w .

36 Todos los resultados mencionados corresponden a modelos de utilidad aleatoria con error homocedástico, excepto el de Cassar et al. (2017), que utiliza una regresión por intervalos parecida al modelo de parámetros aleatorios.

Segundo, pasamos a reportar las estimaciones del modelo de Fechner, que corresponden a las especificaciones (5) a la (8). Recordemos que este modelo proviene de la familia de modelos de utilidad aleatoria, los cuales difieren fundamentalmente de los de parámetros aleatorios en cuanto a que el componente estocástico se incluye en la diferencia cardinal entre las utilidades de las loterías. A diferencia del modelo de Blavatsky, se asume que la varianza del componente estocástico, o parámetro de precisión, es homocedástica; es decir, es la misma para todas las decisiones.

Encontramos que el parámetro de aversión al riesgo toma un valor igual a 1,092 si se asume que la varianza del componente estocástico es igual a 1 y que no hay posibilidad de tiemble. Al añadir el parámetro de tiemble, el coeficiente de aversión al riesgo se reduce a 0,903. De acuerdo con esta última especificación, la probabilidad con la que el sujeto representativo experimente un tiemble y escoja una lotería aleatoriamente es de 24%. Al permitir que el parámetro de precisión sea flexible, reportamos un valor de r igual a 0,775 si se asume que no hay posibilidad de tiemble, e igual a 0,911 cuando se asume que sí hay. De manera similar a los estimados de los modelos de parámetros aleatorios, el valor de σ toma el valor de 2,804 al no incluir tiemble y se reduce a 0,953 cuando sí se incluye.³⁷ Al igual que para la familia de modelos anterior, el valor de w no cambia sustancialmente de una especificación a otra.

Tercero, analizamos los estimados del modelo de Blavatsky (2014), que corresponden a las especificaciones (9) y (10). Este es un modelo perteneciente a la familia de utilidad aleatoria que, a diferencia del modelo de Fechner, asume que la varianza estocástica varía de

³⁷ Es importante recalcar que la magnitud de σ de los modelos de utilidad aleatoria no es comparable con la de los modelos de parámetros aleatorios. Esto se debe a que hay una diferencia en la escala del componente estocástico en cada familia de modelos.

acuerdo con los montos de dinero en juego y sus probabilidades asociadas. Con esto en cuenta, es interesante observar que los estimados de r , iguales a 1,031, son similares a los de los modelos anteriores. El parámetro σ ya no se puede interpretar tan fácilmente como antes (véase la ecuación 3.11), así que no es comparable con los parámetros de precisión de los modelos Fechner. Nuevamente, incluir w en la estimación reduce el valor de σ y, más interesante aún, la heterocedasticidad en el componente estocástico aditivo podría estar causando que el valor de r cambie muy poco al incluirse la probabilidad de tiembla. Según el valor de w , los sujetos escogen de manera aleatoria con una probabilidad de 16,1%.

Ajuste de los modelos de agente representativo

En la parte inferior de la tabla de regresiones, exponemos la log-verosimilitud y el estadístico del criterio de información de Akaike (AIC, por sus siglas en inglés). Este criterio estadístico³⁸ es muy usado para comparar el ajuste entre modelos y seleccionar el mejor en términos de predicción y complejidad. En efecto, el AIC, además de valorar cuánto predice el modelo, castiga si es que el modelo estima más parámetros de los que necesita. El modelo preferido es el que tiene el valor mínimo en el AIC.³⁹ Una virtud importante es que los resultados son válidos para comparaciones entre modelos no anidados (Busemeyer y Diederich, 2014). Los diferentes tipos de modelos que se han empleado en este estudio no son anidados, y dada la corrección que

38 $AIC = 2k - 2LL$, donde k es el número de parámetros del modelo y LL es la log-verosimilitud del modelo estimado (Akaike, 1974).

39 Puede interpretarse que el modelo seleccionado por el AIC es aquel más parsimonioso, es decir, aquel que explica bien los datos de la manera más simple.

incluimos para que el término de tiemble capture las violaciones estocásticas, también es el caso entre los modelos con y sin tiemble. En el cuadro 4.2, ordenamos los estadísticos de criterio de AIC de menor a mayor.

Cuadro 4.2
AIC para modelos de riesgo de agente representativo

Pos.	Modelo	Tiemble	σ flexible	Especificación	AIC
1.	Parámetros aleatorios	✓	✓	(4)	41 444
2.	Parámetros aleatorios	✓		(2)	41 447
3.	Blavatsky		✓	(9)	41 473
4.	Blavatsky	✓	✓	(10)	41 476
5.	Fechner	✓		(6)	41 812
6.	Fechner	✓	✓	(8)	41 813
7.	Parámetros aleatorios		✓	(3)	41 918
8.	Fechner		✓	(7)	43 051
9.	Fechner			(5)	44 167
10.	Parámetros aleatorios			(1)	45 600

Nota: En este cuadro, ordenamos los modelos según el valor del criterio de información de Akaike, de menor a mayor. Recordemos que un menor valor indica una mejor valoración de acuerdo con este criterio. La segunda columna indica el nombre del modelo estocástico; la tercera, si es que el modelo permite que se incluya una probabilidad de tiemble w ; la cuarta, si es que el parámetro de precisión es flexible o igual a 1; la quinta, el número de la especificación; y la sexta, el valor del criterio de información.

Observamos que los estadísticos varían entre 41 444 para la especificación (4), la que mejor explica las decisiones, y 45 600 para la especificación (1), la que peor las explica. Es interesante notar que el mejor y el peor modelo son del tipo de parámetros aleatorios, lo que sugiere que su ajuste relativo es producto de los errores estocásticos —o falta de— incorporados a cada uno. En efecto, salvo por el modelo de

Blavatsky, el ajuste de los modelos mejora sustancialmente al incluir una probabilidad de temble. En cambio, el ajuste mejora solo ligeramente al permitir que σ sea flexible. En general, los modelos de parámetros aleatorios con temble son los que mejor predicen, seguidos de cerca por los modelos de Blavatsky –con o sin temble–, y seguidos, finalmente, por los modelos de Fechner y los de parámetros aleatorios sin temble.

Diferencias en el riesgo estimado según listas

Utilizamos el modelo de parámetros aleatorios con parámetro de temble y σ flexible –que es el que predice mejor entre nuestros modelos– para analizar cómo cambian los estimados según la lista utilizada. Recuérdese que utilizamos tres listas de 10 decisiones cada una para medir las preferencias de riesgo. Queremos saber ahora si las tres listas arrojan estimados similares de preferencias. En el cuadro D.1, en el anexo, se muestra que los estimados del parámetro CRRA varían entre las listas 1 y 3 con respecto a la lista 2. En específico, solo usando la lista 1 estimamos un parámetro CRRA igual a 0,788; con la lista 2, un parámetro igual a 1,422; y con la lista 3, un parámetro igual a 0,823. Analizando los intervalos de confianza, resulta claro que no podemos rechazar la igualdad entre los estimados según las listas 1 y 3, pero sí con respecto a la lista 2. Esta diferencia puede deberse a muchas razones: a un efecto orden –i. e., hacer *switching point* en la decisión 8 en las listas 1 y 3 está asociado a un parámetro CRRA entre 0,7 y 1 aproximadamente, pero hacerlo en la lista 2 está asociado a un parámetro CRRA entre 1,4 y 1,73–, a los distintos montos de dinero sobre los que se basan las decisiones en las listas –i. e., la aversión al riesgo podría depender del monto de dinero en juego– o a otras preferencias

que escapan de EUT, que no desarrollamos en este documento. Otro punto importante que se desprende de este cuadro es que σ y w toman valores similares para las tres listas, lo cual parece apuntar a que el nivel de comprensión de las tareas fue similar en todas las listas.

4.1.2. Incorporando heterogeneidad no observable

A diferencia de los modelos de agente representativo, incluir efectos aleatorios permite modelar la heterogeneidad no observable entre sujetos en la población estudiada, distinguiéndola de la variación dentro de cada sujeto capturada por σ y w . Ya que el foco de este análisis es la actitud ante el riesgo, introducimos el parámetro η , que captura la variación en la distribución de actitudes al riesgo. Asumimos que r —en los modelos de utilidad aleatoria— y m —en los modelos de parámetros aleatorios— provienen de una distribución normal, con media μ y desviación estándar η . De los 10 modelos expuestos en la sección anterior, estimamos el de Fechner, el de Blavatsky y el de parámetros aleatorios, todos incluyendo σ flexible, un parámetro de tiemble y la corrección para violaciones estocásticas mencionada anteriormente. Para estas estimaciones, empleamos 75 extracciones de Halton. Mostramos los estimados en el cuadro 4.3.

Con respecto a la estimación del modelo Fechner con tiemble y efectos aleatorios, r se distribuye con media igual a 1,07 y desviación estándar igual a 2,58. Esta desviación estándar significa un alto grado de variabilidad entre las preferencias de los sujetos. En contraste con esta alta variabilidad entre sujetos, el valor que toma el parámetro de precisión σ es muy pequeño, aunque estadísticamente mayor que cero, y sugiere que, más allá de la probabilidad de tiemble, las preferencias se reflejan inequívocamente en las decisiones tomadas. El valor

estimado del parámetro de tiembla, w , indica que los sujetos toman decisiones de manera aleatoria con una probabilidad de 14,9% en promedio.

Cuadro 4.3
Modelos de riesgo, permitiendo heterogeneidad
no observable entre sujetos

Modelo:	Fechner	Blavatsky	Parámetros aleatorios
	(1)	(2)	(3)
μ	1,069 (0,000)	1,081 (0,047)	1,100 (0,049)
η	2,575 (0,001)	1,781 (0,046)	1,769 (0,046)
σ	0,000 (0,000)	0,336 (0,006)	0,500 (0,012)
w	0,149 (0,003)	0,068 (0,003)	0,069 (0,003)
LL	-14 273	-12 925	-12 888
AIC	28 553	25 859	25 783
N	38 940	38 940	38 940

Nota: Errores estándar entre paréntesis. LL hace referencia al valor de la log-verosimilitud; AIC , al criterio de información de Akaike; y N , al número de observaciones.

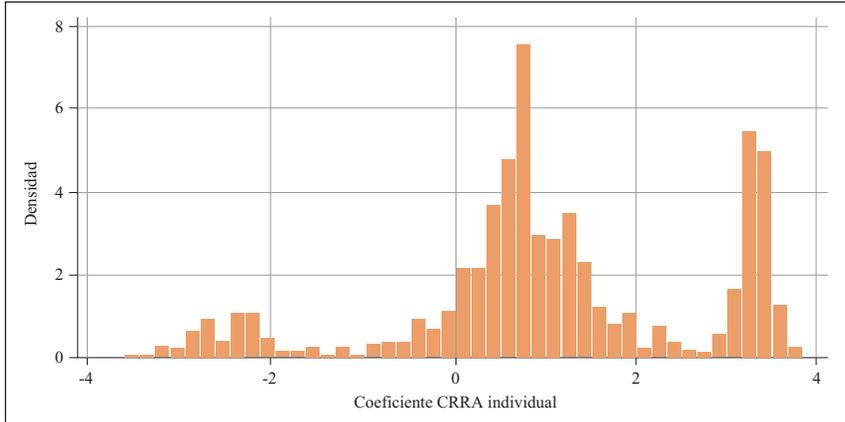
Según el modelo de Blavatsky con tiembla y efectos aleatorios, el parámetro CRRA se distribuye con media 1,08 y desviación estándar igual a 1,78. Aunque no es tan alta como la de Fechner, la desviación estándar η de este modelo también sugiere una alta variabilidad en las preferencias entre sujetos. El parámetro de precisión en este modelo no es tan pequeño como en el modelo de Fechner ($\sigma = 0,336$), aunque sí lo es si lo comparamos con el valor que tomaba en los modelos de

Blavatskyy sin efectos aleatorios. Esto ocurre porque, en los modelos de agente representativo, el parámetro de precisión absorbe también la variabilidad entre sujetos. Para estos modelos con heterogeneidad no observable, σ pasa a ser un indicador de la variabilidad intrasujeto. El valor estimado del parámetro de tiemble –igual a 0,068– también se reduce con respecto a su versión de agente representativo, lo que implica que w absorbe, asimismo, parte de la heterogeneidad entre sujetos.

Finalmente, la estimación del modelo de parámetros aleatorios es cualitativamente muy similar a la del modelo de Blavatskyy. Esto es, el promedio del parámetro CRRA se distribuye con media igual a 1,10 y desviación estándar igual a 1,77. El parámetro de precisión σ estimado es igual a 0,5 e implica que existe variabilidad en las preferencias dentro de cada sujeto al tomar decisiones. Por último, el parámetro de tiemble, igual a 0,069, indica que la probabilidad con la que los sujetos toman decisiones de manera aleatoria es igual a 6,9%.

En estos tres modelos, los efectos aleatorios afectan los valores de algunos parámetros de manera consistente, lo cual sugiere que son importantes para identificar la distribución de preferencias ante el riesgo. Como acabamos de discutir, el parámetro de precisión y la probabilidad de tiemble se reducen considerablemente, lo cual indica que antes absorbían parte de la variabilidad entre sujetos. No obstante, los estimados de aversión al riesgo varían muy poco: el promedio de estos tres modelos, igual a 1,083, es ligeramente superior que el promedio de r para modelos de agente representativo, igual a 0,997. Puede verse la distribución de los coeficientes CRRA individuales, según el modelo de parámetros aleatorios, en la figura 4.1. Notamos que existen dos picos muy marcados, uno alrededor de 1 y el otro alrededor de 3. Según esta distribución, el 85% de los sujetos son aversos al riesgo, y el resto son neutrales y amantes del riesgo. Las distribuciones correspondientes a los otros dos modelos pueden verse en las figuras D.1 y D.2, en los anexos.

Figura 4.1
Distribución de coeficiente de aversión al riesgo,
según el modelo de parámetros aleatorios



Ajuste de los modelos con heterogeneidad no observable

El estadístico de AIC muestra que los modelos con efectos aleatorios son superiores que los de agente representativo por un gran margen. Segundo, todas las pruebas de selección de modelo son consistentes y determinan que los datos son mejor explicados, en primer lugar, por el modelo de parámetros aleatorios con temble ($AIC = 25\ 783$), seguido cercanamente por el modelo de Blavatsky con temble ($AIC = 25\ 859$) y, en último lugar, el modelo Fechner con temble ($AIC = 28\ 553$). Una manera aún más formal de evaluar las diferencias entre los niveles de predicción de cada modelo es mediante las pruebas de Vuong y Clarke. La primera es una prueba paramétrica que se calcula mediante una diferencia estandarizada entre las log-verosimilitudes de los modelos (Vuong, 1989; Loomes et al., 2002). La segunda es una prueba no paramétrica que se calcula mediante la ratio de veces en la

que las verosimilitudes individuales de un modelo son mayores que las del otro (Clarke, 2007). La prueba de Vuong requiere mayores supuestos que la prueba de Clarke, pero es más flexible si se quieren hacer ajustes a la comparación, por lo que realizamos ambas pruebas. Para estos modelos, la prueba de Vuong y la prueba de signos de Clarke, expuestas en el cuadro 4.4, sugieren el mismo orden que el AIC con bastante nivel de significancia: el modelo de parámetros aleatorios es el que mejor predice, y luego sigue el de Blavatsky. Aunque consideramos la heterogeneidad no observable en estos modelos, los resultados de las pruebas sugieren que, nuevamente, el modelo de parámetros aleatorios con tiemble es el mejor para explicar las decisiones ante el riesgo en esta base de datos.

Cuadro 4.4

P-values de un lado para las pruebas de Vuong y Clarke

	Vuong	Clarke
Parámetros aleatorios versus Fechner	0,00	0,00
Parámetros aleatorios versus Blavatsky	0,00	0,00
Fechner versus Blavatsky	1,00	1,00

Nota: El *p-value* corresponde a la hipótesis de que el primer modelo en la comparación no es el mejor entre los dos. Por tanto, mientras más cercano a cero sea el *p-value*, es más beneficioso para el primero.

4.1.3. Incorporando heterogeneidad en características observables

Como se detalló en la sección anterior, la heterogeneidad puede incorporarse no solamente mediante la simulación de características no observables, sino también por medio de la variabilidad en las características observables. Es decir, se puede estimar un modelo de agente

representativo y asumir que, para este caso, el coeficiente de aversión al riesgo depende linealmente de ciertas características observables. En esta ocasión, emplearemos el modelo de agente representativo de parámetros aleatorios con probabilidad de temblor, que es el que presentó mejor ajuste dentro de esta familia de modelos para riesgo.

El cuadro 4.5 muestra los estimados del modelo. En primer lugar, como cabía esperar, según los valores de log-verosimilitud, el nivel de predicción del modelo es mayor que el modelo de parámetros aleatorios de agente representativo, pero esto no ocurre si lo comparamos con el modelo que incorpora efectos aleatorios, lo cual sugiere que la

Cuadro 4.5
Estimados de riesgo permitiendo heterogeneidad
en características observables

Variable	r	σ	w
Constante	0,905*** (0,153)	0,945*** (0,032)	0,217*** (0,011)
Vive en provincia	0,174** (0,080)		
Desempleado	0,022 (0,088)		
Educación superior	-0,034 (0,078)		
Hombre	0,019 (0,071)		
Edad	0,004 (0,004)		

Log-verosimilitud = -20 654

Nota: Errores estándar agrupados a nivel individual entre paréntesis. Esta estimación se realizó utilizando el modelo de parámetros aleatorios con probabilidad de temblor. Los símbolos *, ** y *** representan 90%, 95% y 99% de significancia, respectivamente.

heterogeneidad no observable es más importante que la observable en términos de predicción. Una ventaja de estimar modelos con heterogeneidad observable es que permiten analizar cómo se relaciona el parámetro de interés con ciertas características de los individuos. De manera interesante, no encontramos correlaciones significativas según sexo, educación, edad o condición de desempleo. La única correlación significativa que hallamos es con respecto a la distribución geográfica; en particular, aquellos que viven en provincia —específicamente, en Arequipa, Trujillo y Tarapoto— son más aversos al riesgo. Este dato ofrece pistas acerca de que sobre la aversión al riesgo más podrían pesar factores culturales que factores socioeconómicos; sin embargo, con nuestros datos no podemos afirmar de manera sólida esta suposición ni explorar con más detalle los mecanismos que están detrás.

4.2. Estimados de preferencias intertemporales

Al igual que en la sección de resultados de riesgo, presentamos los estimados de preferencias intertemporales en dos grupos: (i) usando modelos de agente representativo y (ii) incorporando efectos aleatorios. En el grupo de modelos de agente representativo, se han utilizado tres de los modelos estocásticos detallados en el capítulo de metodología: el de parámetros aleatorios, el de Fechner y el de Luce, que se muestran en el cuadro 4.6. No realizamos estimaciones para tiempo usando el modelo de Blavatsky porque este no se puede aplicar si no hay incertidumbre en los pagos. Para el grupo de modelos que incorporan efectos aleatorios, presentamos los resultados de los modelos de parámetros aleatorios y de Luce. Ya no mostramos los estimados usando el modelo de Fechner, puesto que presentó dificultades en la convergencia del estimador de máxima verosimilitud. Además, en esta

sección ya no es necesario hacer correcciones por decisiones estocásticamente dominadas de primer orden, por lo cual los modelos que incluyen una probabilidad de tiemble anidan a los modelos que no, a diferencia de las estimaciones de riesgo.

A diferencia de Andersen et al. (2008), estimamos los parámetros de riesgo y tiempo por separado. Aun así, condicionamos los parámetros intertemporales a las actitudes ante el riesgo previamente estimadas, eliminando la posibilidad de problemas de medición como los discutidos en la sección 2.1.2.

4.2.1. Modelos de agente representativo

Como ya se mencionó brevemente, a diferencia de Andersen et al. (2008), no estimamos las preferencias por riesgo y tiempo en un mismo modelo, sino que primero estimamos los parámetros de riesgo y luego los de tiempo, condicionando la estimación de estos últimos a la curvatura de la utilidad –mediante el CRRA–, ya estimada según cada modelo.⁴⁰ No creemos que esto suponga inconveniente alguno ya que, según EUT y QHD, son los parámetros de preferencia intertemporal los que están sujetos a la curvatura de la utilidad, mientras que la preferencia por riesgo no está sujeta al descuento de utilidades futuras; en otras palabras, la preferencia por riesgo no depende de la preferencia por tiempo, pero sí viceversa.⁴¹ Asimismo, para todas las estimaciones

40 Esto quiere decir que, por ejemplo, para estimar los parámetros intertemporales según el modelo de parámetros aleatorios con tiemble, se ha tenido en cuenta el valor de r igual a 1,074, que es el CRRA que se halló en las estimaciones de riesgo usando dicho modelo con tiemble.

41 A lo largo de este documento, asumimos equivalencia de la utilidad para riesgo y tiempo; es decir, que una misma función de utilidad es considerada para tomar decisiones ante riesgo y certidumbre. Esta visión está en línea con Andersen et al. (2008), pero se opone a la de Andreoni y Sprenger (2012b). Para revisar la discusión sobre este tema, véase Cheung (2016).

se ha incorporado un parámetro de tiemble. A diferencia del caso de riesgo, no hay decisiones de tiempo que dominen estocásticamente a otras, por lo que no se controlará por decisiones dominadas en las estimaciones de tiempo. Así, esto implica que, a diferencia de los resultados de riesgo, los modelos que incluyen parámetro de tiemble estimados a continuación son modelos anidados de los modelos que no incluyen parámetro de tiemble.

Cuadro 4.6
Estimados de tiempo, según modelos de agente representativo

Modelo:	Parámetros aleatorios	Fechner	Luce
	(1)	(2)	(3)
δ	1,952 (0,421)	1,423 (0,030)	1,512 (0,028)
β	1,027 (0,024)	1,003 (0,004)	1,002 (0,004)
σ	1,690 (0,383)	0,078 (0,008)	0,010 (0,001)
w	0,000 (0,000)	0,291 (0,014)	0,291 (0,014)
<i>LL</i>	-16 346	-16 366	-16 367
<i>NN</i>	38 940	38 940	38 940

Nota: Errores estándar agrupados a nivel individual entre paréntesis. Cada una de las estimaciones se llevó a cabo utilizando el valor de r correspondiente a cada modelo estocástico. En efecto, para la estimación de parámetros aleatorios, se impuso $r = 1,074$; para la de Fechner, $r = 0,911$; y para la de Luce, $r = 0,887$. Para el caso de las estimaciones con el modelo de parámetros aleatorios, el estimado de δ en realidad corresponde al estimado de m , la media de δ . *LL* hace referencia al valor de la log-verosimilitud; y *N*, al número de observaciones.

Como se muestra en el cuadro 4.6, el valor de la tasa de descuento (δ) es 195% usando el modelo de parámetros aleatorios, 142%

usando el modelo de Fechner y 151% usando el modelo de Luce. Lo primero que salta a la vista es el gran tamaño de las tres tasas de descuento estimadas, lo cual refleja el alto nivel de impaciencia de los trabajadores peruanos; prefieren consumir en periodos más cercanos, por más de que el beneficio de postergar su consumo sea muy alto. En términos más precisos, estas tasas de descuento sugieren que el trabajador peruano, en el periodo t , valora su consumo un año en el futuro en menos de un 40% ($\approx 1/[1+1,5]$) de lo que lo valora en ese periodo. Sin tener que recurrir a formas funcionales, estos resultados no sorprenden, puesto que alrededor del 55% del total de participantes eligieron siempre la opción temprana en todas las decisiones que se les presentó, lo cual por, sí solo, ya es un indicador importante del nivel de descuento con el que valoran el consumo futuro.⁴² Este nivel de descuento es muy alto y sorprende, ya que, en el ámbito internacional y académico, se suelen encontrar tasas de descuento menores que el 25% (Loewenstein y Prelec, 1992). Existen estudios que encuentran valores que sobrepasan el 100%, pero usando otras estrategias empíricas (Frederick et al., 2002). En contraste con países específicos, Andersen et al. (2008) reportan una tasa de descuento de tan solo 10,1% en la estimación que incluye descuento hiperbólico en Dinamarca. En países en desarrollo, Cassar et al. (2017) hallan una tasa de descuento de 19,6% en zonas rurales de Tailandia; y Callen (2015), una de 24,6% en zonas rurales de Sri Lanka, aunque ninguno de estos

42 Una posible razón detrás del alto valor de las tasas de descuento podría ser una falla del diseño experimental, particularmente que la gente desconfíe de que se le vaya a pagar después si es que elige el pago futuro (y por ende elija siempre la opción más temprana). No obstante, de ser así, los participantes en el tratamiento *front-end delay* (FED) (aquellos que no eligen su consumo futuro contra la inmediatez) estarían exentos de esta sospecha y elegirían la opción temprana de manera menos frecuente que aquellos sin este tratamiento (asumiendo ≤ 1), pero esto no ocurre, lo cual le quita sustento a esta idea. En efecto, 53% de los participantes FED eligen la opción más temprana, mientras que hace lo mismo un 58% de los no FED.

dos últimos considera alguna forma de descuento hiperbólico. Estos estudios utilizan una forma funcional ligeramente distinta de la nuestra; aunque, aun así, permiten vislumbrar una gran diferencia en la impaciencia con respecto a nuestro caso.

Con respecto a la diferencia entre los modelos estocásticos, es interesante notar que ninguna de las tasas de descuento estimadas es distinta estadísticamente de las demás, usando un nivel de confianza del 95%. En particular, sorprende el alto error estándar de la tasa de descuento usando el modelo de parámetros aleatorios –igual a 0,421, aproximadamente un 20% del valor del coeficiente–. En cambio, los mismos errores estándar en los modelos de Fechner y Luce son mucho menores –0,030 y 0,028, respectivamente–, lo cual significa que se hallan tasas de descuento con mayor precisión cuando se usan estos dos últimos modelos; esto no sorprende, debido a la similitud matemática entre ambos. Esta diferencia en la variabilidad estimada de la tasa de descuento es absorbida en los modelos de Fechner y Luce mediante el parámetro de tiembla w –igual a 0,291 en ambos casos–, lo cual no ocurre para el de parámetros aleatorios –en el que toma el valor de 0–. Este ruido en la tasa de descuento –ya sea reflejado en el error estándar de δ o en el valor de w – se debe al alto porcentaje de listas en las que se escoge siempre la opción más temprana –incluso ante tasas de interés muy altas–, lo cual no permite que el modelo de estimación recoja estimados precisos debido a que debe reflejar las decisiones de los individuos con valores de δ extremadamente altos y a aquellos con valores de δ más moderados.

El valor del parámetro de descuento cuasihiperbólico (β) es 1,027 usando el modelo de parámetros aleatorios, 1,003 usando el modelo de Fechner y 1,002 usando el modelo de Luce. Ninguno de estos parámetros es distinto estadísticamente de 1, lo cual sugiere que no existe falta de autocontrol en el consumo e inconsistencia dinámica en el

patrón de consumo de los trabajadores peruanos.⁴³ Para los tres modelos, los estimados de β son estadísticamente similares.⁴⁴ Asimismo, el parámetro estimado usando el modelo de parámetros aleatorios es, de nuevo, el menos preciso –con error estándar de 0,024, a diferencia del error estándar de 0,004 para los modelos de Fechner y Luce–.

Ajuste de los modelos con agente representativo

Con la finalidad de evaluar el nivel de predicción de los modelos estocásticos aplicados a las elecciones intertemporales, volvemos a realizar las pruebas de Vuong y Clarke. En el cuadro 4.7, se observan los resultados de las pruebas. En resumen, según los resultados de las dos pruebas, el modelo de parámetros aleatorios es el que mejor se ajusta

Cuadro 4.7

Resultados de pruebas de Vuong y Clarke para modelos de tiempo

	Vuong		Clarke	
	Estadístico	<i>p-value</i>	Estadístico	<i>p-value</i>
PA versus Fechner	1,74	0,04	24 664	0,00
PA versus Luce	1,84	0,03	24 664	0,00
Fechner versus Luce	2,55	0,01	7086	1,00

Nota: En una comparación “modelo 1 versus modelo 2”, el estadístico de la prueba de Vuong favorece al modelo 1 si es positivo y el *p-value* corresponde a la hipótesis de que el modelo 1 no es el mejor modelo. Para la prueba de Clarke, el estadístico favorece al modelo 1 si es mayor que 19 470 y el *p-value* corresponde a la misma hipótesis.

43 Con un valor de $\beta = 1$, la función de descuento intertemporal pasa a ser una de descuento exponencial.

44 Augenblick, Niederle y Sprenger (2015) también encuentran estimados de β cercanos a 1 cuando lo que está en juego es dinero, mas no cuando es esfuerzo (como en nuestro caso).

a lo observado, pero la diferencia no es clara al comparar los modelos de Fechner y Luce; según la prueba de Vuong, mejor se ajusta el de Fechner, pero según la de Clarke, mejor se ajusta el de Luce.

Dado que las estimaciones de la preferencia por consumo intertemporal están condicionadas a la preferencia por riesgo y que para cada modelo estocástico se han utilizado distintos valores de CRRA, podría surgir la duda de qué parte del ajuste del modelo se deba al valor del CRRA utilizado y no a las bondades del modelo estocástico *per se*. Ante esto, también hemos realizado las estimaciones de tiempo usando un mismo CRRA para los tres modelos estocásticos, que es igual al promedio de los cuatro valores de r estimados para los cuatro modelos de decisión ante el riesgo de que incluyan parámetro de temblor, discutidos en la sección de metodología –incluyendo el modelo de Luce, cuyos estimados no se incluyeron en la sección anterior–. Como se observa en el cuadro D.2, los valores de las tasas de descuento cambian según la variación del CRRA con respecto a la estimación anterior –si el CRRA homogéneo es menor que el CRRA utilizado para los resultados del cuadro 4.6, la tasa de descuento debería aumentar, y viceversa–. Al emplear las pruebas de ajuste mencionadas en el párrafo anterior sobre esta nueva estimación, los resultados siguen siendo los mismos, por lo que descartamos que la superioridad del modelo de parámetros aleatorios se deba al valor del CRRA que asume.

Estimados según listas y tipo de tareas

Antes de continuar con los resultados que incluyen la incorporación de efectos aleatorios al modelo, creemos ilustrativo mostrar también cómo varían nuestros estimados según la lista o el tipo de tarea empleado. Como se explicó en la sección 2.2, cada lista es un grupo de 10 decisiones binarias y, para medir las preferencias intertemporales,

se usaron 3 listas. La diferencia principal entre las listas es la brecha temporal (k) entre el pago temprano y el pago tardío: en la lista 1, esta fue de 3 meses; en la lista 2, 6 meses; y en la lista 3, 12 meses. También mostramos cómo los estimados varían de acuerdo con el tipo de tareas. Como mencionamos en la subsección 2.2.3, dado que un alto porcentaje de sujetos elegía siempre la opción más temprana, decidimos modificar las decisiones con el objetivo de que la opción tardía sea más atractiva. Para esto último, básicamente aumentamos los valores de los pagos tardíos y, así, la tasa de interés efectiva anual asociada más alta que se podía alcanzar en cada lista.

Los cuadros D.3 y D.4, en el anexo, muestran los resultados de las estimaciones desagregando según lista para las tareas iniciales y modificadas, respectivamente. Para las tres listas de las tareas iniciales, encontramos estimados estadísticamente similares tanto para δ como para β . No obstante, al observar los estimados de las tareas modificadas, lo que encontramos es que las tasas de descuento estimadas con las tareas iniciales son muy pequeñas en comparación con las estimadas con las modificadas. Esto es algo que esperábamos debido a que las tareas iniciales no permitían capturar el grado de impaciencia tan alto de los sujetos, mientras que las modificadas ocupan un mayor rango para poder hacerlo. En particular, dentro de las tareas modificadas, la lista asociada a la mayor tasa de interés es la 1, y justamente la tasa de descuento estimada con esta lista es la mayor, igual a 1,546, muy similar a lo estimado para la muestra agregada –igual a 1,512–. Esto último, además, nos permite identificar que es la lista 1 de las tareas modificadas la que está “empujando” los resultados para la muestra total.⁴⁵ Antes de pasar a la siguiente subsección, vale la pena hacer notar

45 Esto se debe a que, aun en las tareas modificadas, más de la mitad de los sujetos –aproximadamente el 58%– siguieron eligiendo la opción temprana siempre hasta el final, lo cual nos lleva a plantear la posibilidad de que incluso el nivel de descuento intertemporal de los agentes sea todavía mayor.

que el valor de β se mantiene estadísticamente similar a 1 para todas las listas y tipo de tareas, lo cual otorga mayor soporte a la idea de que el sesgo hacia el presente no caracteriza las decisiones de consumo del trabajador peruano.

4.2.2. Incorporando heterogeneidad no observable

Al igual que para el caso de riesgo, los modelos pueden asumir que existe heterogeneidad no observable entre los sujetos en sus preferencias. Es decir, puede suponerse que las preferencias bajo las cuales los agentes toman sus decisiones provienen de una distribución por estimar, en lugar de que sea un único parámetro homogéneo el que determine las decisiones para todos los agentes. En este caso particular, estimaremos la distribución de δ asumiendo que se distribuye de manera normal, con media μ_δ y desviación estándar η_δ .

Estimamos dos especificaciones: una en la que imputamos un valor homogéneo de r al modelo, y otra en la que estimamos tanto la heterogeneidad de r como la de δ en el mismo modelo. No estimamos la heterogeneidad de β debido a que nuestro diseño experimental no permite identificar este parámetro de manera individual, ante lo cual estas estimaciones que utilizan efectos aleatorios individuales podrían no identificar correctamente los parámetros. Empleando los resultados de la subsección anterior, imputamos un valor de $\beta = 1$. El único modelo estocástico para el cual pudimos estimar ambas especificaciones es el de Luce, que es el que reportamos más adelante. Esto, de por sí, es un hallazgo interesante. Si bien estas especificaciones representan dificultades adicionales en su estimación —e. g., en ajustar la heterogeneidad a la gran cantidad de sujetos que eligen siempre el pago temprano o los supuestos adicionales en definir la

correlación ρ entre r y δ , el modelo de Luce permite estimar ambas especificaciones y parece ser más flexible respecto a los datos que los demás modelos.

El cuadro 4.8 muestra los resultados de las estimaciones. Para ambas, los valores de la media de δ son similares y giran en torno a 1,1, lo que significa que el trabajador peruano promedio descuenta el valor que le otorga su consumo de un año a otro en aproximadamente 48%. Es interesante que el valor de la media de δ es menor que los δ 's estimados usando los modelos de agente representativo. Esto puede deberse a que, al permitir heterogeneidad no observable, la estimación es menos sensible a los valores extremos generados por el alto porcentaje de individuos que siempre eligen la opción temprana.⁴⁶ Justamente con respecto a la heterogeneidad estimada, las desviaciones estándar sí presentan diferencias entre especificaciones: 1,13 para la especificación en la que solo se estiman las preferencias temporales y 0,77 para la que, además, estima las preferencias por riesgo. La variabilidad de la tasa de descuento es menor que la del coeficiente de aversión al riesgo. Creemos que esto se debe, principalmente, a que una elevada proporción de sujetos –incluso con tasas de interés altísimas– prefirió el pago temprano hasta el final, con lo cual el instrumento que utilizamos para medir la tasa de descuento impuso una cota superior en las preferencias, que se concentró hacia la cola derecha de la distribución.

La figura 4.2 muestra la distribución de δ según la especificación que estima conjuntamente las preferencias de riesgo e intertemporales. Se observa que δ se distribuye con una moda muy concentrada

46 Este argumento toma aún más fuerza si consideramos que nuestra estrategia de estimación recae, en buena parte, en predecir las preferencias que no se han podido capturar –debido a que los sujetos no han realizado *switching points*–, recreando la cola de la distribución que no se observa en los datos.

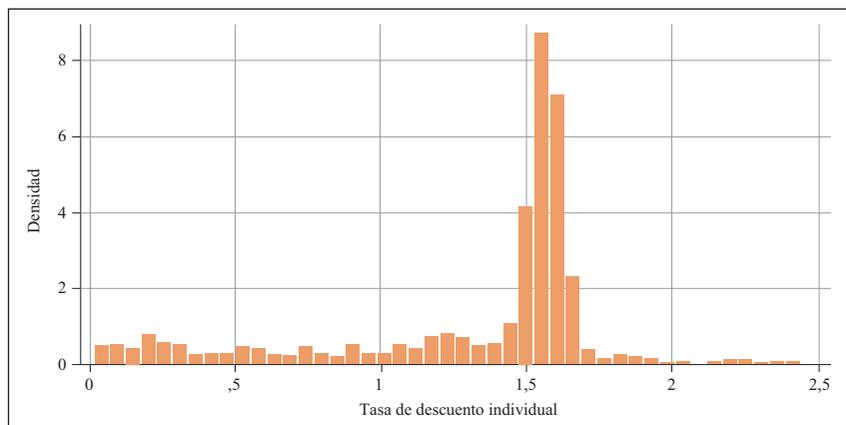
Cuadro 4.8
Estimados de tiempo, según modelos
con heterogeneidad no observable

	Luce	
	Solo tiempo (1)	Tiempo y riesgo (2)
μ_δ	1,163 (0,034)	1,117 (0,044)
η_δ	1,136 (0,037)	0,773 (0,040)
σ tiempo	0,036 (0,001)	0,050 (0,002)
w tiempo	0,020 (0,002)	0,005 (0,001)
μ_r		1,056 (0,027)
η_r		1,548 (0,022)
σ riesgo		0,036 (0,002)
w riesgo		0,032 (0,004)
ρ		-0,009 (0,015)
<i>LL</i>	-9506	-11 418
<i>N</i>	38 940	77 880

Nota: Errores estándar agrupados a nivel individual entre paréntesis. Cada una de las estimaciones se llevó a cabo utilizando el modelo de Luce. Para la especificación (1), se impuso un valor de r igual a 0,887. ρ es la correlación entre r y δ ; para estimarla, asumimos que los choques de r afectan a δ y no al revés. *LL* hace referencia al valor de la log-verosimilitud; y *N*, al número de observaciones.

alrededor de 1,6 y una masa menos densa –aunque constante– para valores menores. De esto podemos confirmar la suposición que habíamos hecho antes: que δ pertenece a una distribución asimétrica a la izquierda y que la media (igual a 1,26) es menor que la mediana (igual a 1,53). Por esta razón, al permitir heterogeneidad, la media de δ resulta diferente del valor de δ para el agente representativo, que se ve mejor reflejado por la mediana. Un patrón similar observamos en la distribución de δ obtenida de la estimación en la que se imputa un valor homogéneo de CRRA, mostrada en la figura D.3 en los anexos.

Figura 4.2
Distribución de la tasa de descuento,
según la especificación conjunta



Ajuste de los modelos con heterogeneidad no observable

Lo último que vale la pena comentar de estos resultados son las log-verosimilitudes calculadas, que ascienden a -9506 usando el modelo que solo estima preferencias temporales y a -11 418 usando el modelo

que estima ambos tipos de preferencias. Ambas cifras son considerablemente menores que las log-verosimilitudes sin incorporar efectos aleatorios, lo cual indica que la heterogeneidad es un factor importante. Vale la pena resaltar que, sobre la base de estas cifras, no podemos decir que una especificación predice mejor que la otra, ya que usan datos distintos. En particular, la especificación conjunta usa los datos de las decisiones de riesgo y tiempo, mientras que la otra solo usa los de las decisiones de tiempo.

4.2.3. Incorporando heterogeneidad en características observables

Al igual que para los estimados de riesgo, también incluimos heterogeneidad en características observables para la estimación de la tasa de descuento y sesgo por el presente. Utilizamos el modelo de parámetros aleatorios con probabilidad de tiemble, que fue el que presentó el mejor ajuste entre los modelos de agente representativo para tiempo, e incorporamos la dependencia lineal de los parámetros con respecto a características observables. Empleamos el mismo conjunto de características que se usó para el caso de riesgo.

El cuadro 4.9 muestra los resultados de esta estimación. Lo primero que salta a la vista es que, según el valor de la log-verosimilitud, el nivel de predicción de este modelo es ligeramente mayor que el de parámetro aleatorio de agente representativo homogéneo, aunque mucho menor que la versión con heterogeneidad no observable. De esto entendemos que la heterogeneidad observable explica menos las decisiones que la no observable.

Luego, una ventaja de este modelo es que permite obtener correlaciones entre los parámetros de comportamiento y las características observables. Así, pues, es interesante notar que la tasa de descuento

es mayor para aquellos que viven en provincia –específicamente en Arequipa, Trujillo y Tarapoto–, lo cual vuelve a traer a colación la posible importancia de factores culturales en las preferencias. También encontramos que la tasa de descuento aumenta con la edad, lo cual podría relacionarse con que, a mayor edad, el periodo de jubilación se acerca y los individuos pasan a pensar menos en el ahorro y más en el consumo. Asimismo, hallamos que ser desempleado tiene una correlación negativa con la tasa de descuento; es decir, que los desempleados valoran más su consumo futuro. Esta relación es importante a la hora de pensar en políticas laborales debido a que, si los desempleados son más pacientes a largo plazo, distribuirán el monto de dinero que reciban de un seguro de desempleo o de la CTS más equitativamente en el rango de tiempo que planean estar sin empleo. De manera contradictoria con esto último, estimamos una correlación negativa del sesgo hacia el presente y la incidencia en el desempleo –lo cual implica que el desempleado se ve tentado a sobreconsumir durante el periodo inmediato–; no obstante, su magnitud es pequeña y su nivel de significancia es menor.

Por último, agregamos a la regresión una variable dicotómica que es igual a 1 si las observaciones corresponden a las tareas de tiempo modificadas (ver sección 2.2.3) para poder mitigar posibles sesgos por el hecho de que aquellos que hicieron estas tareas pudiesen tener características distintas. Como discutimos anteriormente, la tasa de descuento estimada es sustancialmente mayor para los individuos que realizaron las tareas modificadas por las razones que ya expresamos, aunque también encontramos que el sesgo por el presente es ligeramente menor para estas.

Cuadro 4.9
Estimados de tiempo permitiendo heterogeneidad
en características observables

Variable	δ	β	σ	w
Constante	0,438*** (0,084)	1,015*** (0,028)	0,531*** (0,038)	0,000 (0,000)
Vive en provincia	0,080* (0,045)	-0,005 (0,019)		
Desempleado	-0,108** (0,052)	-0,032* (0,019)		
Educación superior	0,017 (0,042)	-0,006 (0,015)		
Hombre	0,034 (0,038)	0,011 (0,013)		
Edad	0,006*** (0,002)	0,000 (0,001)		
Tareas modificadas	0,282*** (0,051)	-0,045*** (0,016)		

Log-verosimilitud = -15 948

Nota: Errores estándar agrupados a nivel individual entre paréntesis. Esta estimación se realizó utilizando el modelo de parámetros aleatorios con probabilidad de tiembla. Por lo mismo, se imputó un valor de r igual a 1,074. Los símbolos *, ** y *** representan 90%, 95% y 99% de significancia, respectivamente.

5. APLICACIÓN DE NUESTROS RESULTADOS: UN EJEMPLO

En esta breve sección, ilustramos cómo podrían usarse nuestros estimados de preferencias de los trabajadores para la evaluación de políticas públicas. En particular, utilizamos como ejemplo una comparación de bienestar entre dos escenarios distintos: uno según el sistema actual en el que el principal mecanismo de protección social en el empleo es la CTS y otro en el que se establece un seguro de desempleo. Para esto, planteamos un modelo muy simple de demanda laboral en el que los trabajadores obtienen utilidad del consumo, cuentan con un coeficiente de aversión al riesgo igual a r y descuentan su futuro con una tasa de descuento δ .

Queremos recalcar que este ejemplo tan solo cumple la función de ilustrar nuestros resultados. No debe ser tomado como una evaluación minuciosa ni completa de cualquiera de estos sistemas de protección. La simpleza del modelo que empleamos, en todo caso, tiene la intención de servir como introducción a modelos más complejos o específicos para entender esta problemática.

5.1. Modelo

Consideramos una economía compuesta por individuos que son aptos para trabajar y buscan hacerlo. Los individuos pueden estar empleados o desempleados. Si el individuo se encuentra empleado, pierde su em-

pleo con probabilidad p_{desp} . En caso de estar desempleado, el individuo no encuentra empleo con probabilidad p_{noemp} , menor que p_{desp} . El valor de p_{desp} de cada individuo se basa en una pregunta hecha en la encuesta, mientras que el de p_{noemp} es simplemente $\sqrt{p_{desp}}$.

Cada periodo de tiempo equivale a un mes, y tras finalizar el mes, cada individuo empleado recibe su salario. De no estar empleado, no recibe su salario, pero puede ser que cuente con los beneficios generados por alguno de los mecanismos de protección que examinamos. El salario siempre es el mismo en todo el historial laboral del individuo y es igual al salario que reportó en la encuesta. Los individuos obtienen utilidad mediante el consumo. Esta función de utilidad tiene la forma CRRA discutida en este documento y se ve afectada por la curvatura determinada por el valor del coeficiente de aversión al riesgo r . También descuentan exponencialmente los flujos de utilidad futura usando una tasa de descuento igual a δ . Para simplificar este ejercicio ilustrativo, y de manera consistente con los resultados de este documento, asumimos que los agentes no se ven afectados por un sesgo hacia el presente.

También simplificamos el ejercicio asumiendo que los sujetos no eligen su nivel de ahorro de manera endógena. El modelo asume que un sujeto empleado consume el 90% de sus ingresos del periodo y ahorra la diferencia. En caso de que esté desempleado, el sujeto intenta consumir lo mismo que consumiría si estuviera empleado. De no contar con los ahorros suficientes, consumirá un menor porcentaje de su salario, dependiendo del nivel de ahorros con el que cuente en ese momento.⁴⁷ Decidimos usar una tasa de ahorro exógena para facilitar

47 En particular, asumimos que los desempleados se rigen por una “regla” para determinar el porcentaje de lo que consumen. Primero, consideran consumir un 80% de su salario; si es que este porcentaje no cumple la regla, entonces intentan con 60% y así hasta el 20%. Si es que el 20% tampoco la cumple, consumen todo su patrimonio. La “regla” de consumo es la siguiente: el nivel de ahorros, más la ayuda económica que reciben por el desempleo, debe ser mayor que el doble del consumo planeado.

la interpretación de este ejercicio de ejemplo. El uso de tasas de ahorro exógenas se remonta hasta el clásico modelo de crecimiento de Solow (1956), aunque, hoy en día, la gran parte de artículos que pretenden modelar las dinámicas de la demanda laboral e incertidumbre en el empleo asumen que el ahorro es endógeno.⁴⁸ Por esto, insistimos en que no deben usarse los resultados de esta comparación entre CTS y seguro de desempleo como evidencia de que una política es mejor que la otra.

5.1.1. Seguro de desempleo

Para el escenario en el que se cuenta con un sistema de seguro de desempleo, asumimos que todos los individuos empleados contribuyen mensualmente a un fondo común desde el cual se repartirán los beneficios a los individuos que se encuentren desempleados y sean elegibles. Para ser elegible al seguro, consideramos que la persona debe haber contribuido al fondo por lo menos cinco meses durante el último año y no debe haber recibido pagos del seguro desde hace más de un año.⁴⁹ En caso de perder el empleo y ser elegible, el sujeto recibe el 80% de su salario tras el primer mes de desempleo. Cada mes consecutivo de desempleo, el sujeto recibe 20 puntos porcentuales menos de su salario, hasta el quinto mes en adelante, cuando el desempleado

48 Para un buen ejemplo de cómo incorporar la elección de ahorro endógena a un modelo de seguro de desempleo, véase Popp (2017). Otros artículos importantes en esta literatura son los de Hansen e Imrohoroğlu (1992), Acemoglu y Shimer (2000), Álvarez y Veracierro (2001), y Engen y Gruber (2001).

49 Estos criterios no son necesariamente los que se suelen utilizar en otros países, pero son cercanos a los que se discutieron cuando el debate sobre esta posibilidad era más mediático. De hecho, más adelante veremos que estos criterios pueden ser muy estrictos y no beneficiar a los trabajadores que tienen empleos más inestables.

deja de recibir beneficios. Asimismo, asumimos que el seguro de desempleo no conlleva gastos administrativos y que las contribuciones esperadas al fondo deben ser iguales a los egresos esperados; es decir, el programa debe ser autosostenible.

De acuerdo con los supuestos del modelo y el conjunto de salarios y probabilidades de despido de los trabajadores encuestados, la tasa de contribución debe ser igual a 2,21% del ingreso mensual de cada sujeto. Esta tasa es condicional a los supuestos específicos de este modelo y no necesariamente es la tasa óptima en caso de implementar dicha política.

5.1.2. Compensación por tiempo de servicios

En el sistema de CTS aplicado en el Perú, los trabajadores reciben un monto similar a la mitad de un salario mensual dos veces al año, el cual es pagado por el empleador. Estos pagos se van acumulando en un fondo individual hasta que la persona pierde su empleo, momento en el cual puede disponer completamente de ese dinero.

Dado que las contribuciones al fondo de la CTS provienen del empleador, podemos pensar que estas, en realidad, son descuentos del salario del trabajador. Es decir, el empleador, en lugar de pagar un mayor salario mensual al trabajador, le descuenta la fracción de su salario que destina al fondo de la CTS. El trabajador asume el costo de la CTS mediante salarios más bajos. En este modelo asumimos que los individuos contribuyen un doceavo de sus salarios a un fondo individual al que solo pueden acceder completamente —e incorporar a sus ahorros— una vez que pierden el empleo. Además, al inicio del periodo de análisis los individuos no tienen fondos disponibles en su cuenta de CTS y los gastos administrativos son nulos.

5.2. Datos

Para simular esta aplicación de nuestros datos, utilizamos las preferencias heterogéneas de riesgo y tiempo estimadas en este documento, las cuales provienen de la segunda especificación del cuadro 4.8. Asimismo, utilizamos los datos de salarios e indicadores de probabilidades subjetivas de despido que recogimos mediante la encuesta que acompaña a los experimentos discutidos en este documento.

Con la finalidad de utilizar la muestra total de nuestros estimados de preferencias, imputamos valores aleatorios de salarios e indicadores de probabilidades subjetivas de despido a aquellos individuos para los que no recogimos estos datos, usando realizaciones aleatorias de las distribuciones que sí observamos.⁵⁰ Luego, para la muestra completa, hacemos algunas transformaciones adicionales al indicador de probabilidad subjetiva de despido para generar la probabilidad “objetiva” de perder el empleo y la probabilidad “objetiva” de encontrar empleo.⁵¹ Según estas probabilidades, la tasa de desempleo es igual a 16,4%, lo cual es consistente con que los trabajadores peruanos sobreestiman sus probabilidades de despido.

50 No contamos con información laboral para toda la muestra debido, principalmente, a que una de nuestras submuestras aún no formaba parte del mercado laboral al momento de la encuesta.

51 Para esto, primero convertimos el indicador –que está enmarcado en un horizonte temporal de seis meses– a una probabilidad subjetiva mensual. Después, añadimos un componente estocástico a cada probabilidad para cada periodo. Finalmente, tomamos dicha probabilidad como la probabilidad de despido –sujeta a estar empleado– y generamos la probabilidad de encontrar empleo –sujeta a estar desempleado– como la mitad de la anterior.

5.3. Resultados

Corrimos 90 simulaciones con 60 periodos de tiempo en cada una. Para cada simulación, calculamos el nivel de bienestar de cada sujeto como la utilidad descontada desde la perspectiva del periodo 0 usando los flujos de consumo futuros. Para cada simulación, se genera una utilidad descontada en el escenario de seguro de desempleo y otra en el de CTS. Luego, calculamos los promedios de las 90 utilidades descontadas para ambos escenarios y los comparamos.

Dado que en nuestro enfoque la utilidad tiene un rol ordinal, mas no necesariamente cardinal, nos limitamos a identificar tan solo si el individuo prefiere el sistema de seguro de desempleo sobre el de CTS; es decir, si $U_{SD} > U_{CTS}$. No prestaremos atención, en este ejemplo, a la magnitud del diferencial entre utilidades para el mismo individuo, ni siquiera porcentual.

Tomando esto en cuenta, encontramos que el 58% de los individuos estarían mejor en un sistema de seguro de desempleo y el resto estarían mejor en el sistema de CTS. Las actitudes ante el riesgo y el tiempo adquieren un rol importante para determinar la preferencia de un sistema sobre otro. De hecho, tenemos que el promedio del coeficiente de aversión al riesgo para aquellos que prefieren el seguro de desempleo es 0,854; y para aquellos que prefieren el sistema de CTS, 1,088. Sobre la tasa de descuento, las diferencias son menos marcadas, lo cual se espera dado que la diferencia entre ambos sistemas recae en la diversificación del consumo y no tanto en la distribución intertemporal de este –más aún ya que hemos asumido una tasa de ahorro exógena–. Para quienes prefieren el seguro de desempleo, su tasa de descuento promedio es 1,236; y para quienes prefieren el sistema de CTS, el promedio es 1,285. En cuanto a las variables de empleo utilizadas en el modelo, tenemos que la probabilidad subjetiva de despido

—transformada a escala mensual— es menor para los que prefieren seguro de desempleo —i. e., 2,6% frente a 4,1%—, aunque el salario es muy similar para ambos grupos.

Cuadro 5.1
Predictores de preferencia por seguro de desempleo
en nuestro ejemplo

Variable	$Pr(U_{SD} > U_{CTS})$	
	(1)	(2)
$\hat{\tau}$	-0,072** (0,030)	0,005 (0,046)
$\hat{\delta}$	-0,105 (0,065)	-0,132 (0,101)
$\ln(w)$	-0,113* (0,065)	-0,111* (0,065)
p_{desp}	-12,223*** (1,496)	-11,285*** (3,369)
$\hat{\tau} \times p_{desp}$		-2,458** (1,000)
$\hat{\delta} \times p_{desp}$		0,727 (2,455)
Constante	1,602*** (0,484)	1,563*** (0,492)

Nota: Errores estándar robustos entre paréntesis. Los símbolos *, ** y *** representan 90%, 95% y 99% de significancia, respectivamente.

La columna (1) del cuadro 5.1 muestra un análisis de regresión *probit* que confirma estas observaciones. El coeficiente de aversión al riesgo y la probabilidad de despido vienen acompañados de coeficientes de signo negativo y significativos, mientras que la tasa de descuento y el salario tienen coeficientes poco o nada significativos. Para

evitar que estas relaciones se vean afectadas por valores extremos o distribuciones específicas, realizamos pruebas no paramétricas, como la prueba de Mann-Whitney y la prueba de igualdad de medianas, y los resultados en el cuadro 5.2 apuntan a que, efectivamente, las distribuciones no son similares, y la aversión al riesgo y la probabilidad de ser despedido son mayores entre quienes prefieren la CTS.

Cuadro 5.2

Comparaciones no paramétricas de las características de los trabajadores según su preferencia por seguro de desempleo

Prueba	r	δ	w	p_{desp}
<i>Mann-Whitney</i>				
$Pr(x_{UCTS} >_{USD} = x_{USD} >_{UCTS})$	0,000	0,034	0,283	0,000
$Pr(x_{UCTS} >_{USD} > x_{USD} >_{UCTS})$	0,696	0,534	0,517	0,642
<i>Prueba de igualdad de medianas</i>				
$Pr(\tilde{x}_{UCTS} >_{USD} = \tilde{x}_{USD} >_{UCTS})$	0,000	0,008	0,273	0,000

Nota: La primera fila muestra la probabilidad de que una extracción aleatoria de la variable para quienes prefieren la CTS sea igual que una extracción aleatoria para quienes prefieren el seguro de desempleo. La segunda fila muestra la probabilidad de que la variable sea mayor para quienes prefieren la CTS sobre los que prefieren el seguro de desempleo. La tercera fila muestra la probabilidad de que las muestras hayan sido extraídas de poblaciones con las mismas medianas.

La columna (2) del cuadro 5.1 también es interesante, pues sugiere que no es la aversión al riesgo *per se* lo que predice negativamente la preferencia por un seguro de desempleo, sino la interacción entre el coeficiente de aversión al riesgo y la probabilidad de ser despedido. Es decir, mientras menos estable sea el empleo, un trabajador averso al riesgo preferiría menos un seguro de desempleo según este modelo. Dicha interacción no ocurre con la tasa de descuento.

Es interesante, entonces, notar que la probabilidad de despido de un trabajador predice negativamente la preferencia por un seguro de desempleo. Esto, en principio, podría parecer contraintuitivo, pues uno creería que, en un seguro de desempleo, quienes son despedidos frecuentemente se benefician sobre quienes no suelen ser despedidos, dado que es un fondo común. No obstante, este resultado contrario sugiere que este no siempre es el caso, y que la atención debe enfocarse también en los criterios de elegibilidad. Si estos criterios son estrictos, entonces las personas con mayor inestabilidad laboral podrían estar incluso peor. En esta misma línea, este efecto negativo del sistema de seguro de desempleo sobre el bienestar de los trabajadores menos inestables se acentúa para los más aversos al riesgo, pues habrá meses en los que su consumo llegaría a ser muy bajo. Por otro lado, los resultados sugieren que no hay una relación marcada entre el nivel de paciencia de los trabajadores y la preferencia por alguno de los sistemas. Esto puede estar asociado al hecho de que, según las características de nuestro modelo, no hay diferencias importantes en cuanto al tiempo que un trabajador tiene que esperar para ser elegible con un seguro de desempleo frente al esquema de CTS, en el cual no hay restricción temporal de contribuciones: si pierde el empleo, puede usar ya sea lo que ahorró el mes anterior o lo que ahorró hace cinco años.

Así, pues, este simple modelo ilustra algunas relaciones importantes que podrían presentarse en el mercado laboral de cumplirse los supuestos que acá mostramos y teniendo en cuenta las preferencias ante el riesgo y el consumo intertemporal. Esto permite tener una idea de la importancia de los parámetros de preferencias que estimamos en este documento, y cómo estos interactúan con otras variables de empleo importantes para determinar qué política es mejor que otra. Por último, recalamos que la finalidad de este ejercicio es ilustrar una

aplicación de nuestros estimados y no convertirse en una evaluación minuciosa o completa de la viabilidad de un seguro de desempleo en el Perú.

6. DISCUSIÓN Y CONCLUSIONES

En este documento, hemos estimado las preferencias ante el riesgo y de consumo intertemporal de los trabajadores peruanos. Nuestras contribuciones pueden dividirse en dos áreas de interés: (i) aquellas relacionadas con las teorías determinísticas de comportamiento de los agentes y (ii) aquellas relacionadas con el componente estocástico en las elecciones que asumen. Esto es, estudiamos cómo los trabajadores toman decisiones considerando el rol de los incentivos observables o “racionales” –como los incentivos económicos– y también el de aquellos factores que no se observan, pero que influyen en la decisión del agente –como la falta de atención o la capacidad cognitiva–.

Sobre la primera área, es preciso primero hacer un recuento de a qué nos referimos con preferencias ante el riesgo y preferencias de consumo intertemporal. Las preferencias ante el riesgo determinan cuán aversa al riesgo es una persona; es decir, cuán dispuesta está a afrontar costos adicionales para reducir la *incertidumbre* en su consumo. En nuestro caso, nos enfocamos en estimar el valor que toma el parámetro de aversión relativa al riesgo (denotado como r). Cuando este parámetro es positivo, se dice que el individuo es averso al riesgo; y cuando es negativo, que es amante del riesgo. Por otro lado, con preferencia de consumo intertemporal nos referimos a la valoración relativa que le asigna el individuo a consumir en una fecha temprana con respecto a consumir en una fecha posterior. En este estudio, nos enfocamos en estimar dos parámetros: la tasa de descuento exponencial (δ) y el

sesgo por el presente (β). El primer parámetro mide cuánto descuenta el individuo el valor de su consumo de un periodo al siguiente –i. e., entre t y $t + 1$ –, mientras que el segundo está asociado a un descuento adicional frente al consumo inmediato –i. e., relativo a $t = 0$ –. En otras palabras, δ representa el descuento *consistente* a través del tiempo; y β , si es menor que 1, representa los *impulsos* (impaciencia) que llevan a una persona a consumir en el presente más de lo que hubiese planeado.

En cuanto a la especificación del componente estocástico, utilizamos varios modelos con supuestos distintos. La primera distinción se establece entre los modelos que asumen agente representativo y aquellos que incorporan heterogeneidades entre los agentes. Los primeros, como su nombre indica, asumen que los parámetros por estimar representan a toda la población de estudio; esto es, que las diferencias entre los sujetos corresponden a componentes estocásticos con un comportamiento homogéneo entre la totalidad de los sujetos. Estos modelos, a su vez, se dividen en dos familias: (i) los modelos de parámetros aleatorios –en los que el término estocástico es un componente aditivo de los parámetros por estimar– y (ii) aquellos de utilidad aleatoria –en los que el término estocástico es un componente aditivo de la función de utilidad del sujeto–. Los modelos de agente representativo pueden modificarse de forma que incorporen heterogeneidades –observables y no observables– entre los agentes. De esta manera, los modelos incorporan el hecho de que las preferencias pueden ser distintas entre sujetos y que estas diferencias no son completamente aleatorias, sino que pueden corresponder a características observables y no observables de los trabajadores.

A continuación, resumimos nuestros hallazgos. En cuanto a la preferencia ante el riesgo, hallamos que el valor del coeficiente de aversión relativa al riesgo (γ) toma valores cercanos a 1. Esto significa que el trabajador peruano muestra un grado de aversión al riesgo moderado

alto, no mucho mayor de lo que se suele encontrar en otras partes del mundo. Además, esto sugiere que utilizar una función de utilidad logarítmica para modelar las decisiones bajo riesgo de un agente representativo es preciso. Como punto de comparación, Andersen et al. (2008) usan una estrategia similar a la nuestra y encuentran que r es igual a 0,774 para la población danesa. En Corea del Sur, Kim et al. (2012) estiman valores cercanos a 0,8 para una muestra representativa del país. En comunidades en desarrollo, los parámetros CRRA estimados oscilan entre 0,45 y 0,74 (Harrison et al. 2005; Galarza, 2009; De Brauw y Eozenou, 2014; Cassar et al., 2017). Gandelman y Hernández-Murillo (2015) asumen la misma función de utilidad que nosotros, aunque usan una estrategia empírica distinta, y encuentran que el valor del CRRA promedio en varios países –desarrollados y no desarrollados– es muy cercano a 1 y que el peruano es más averso al riesgo que el promedio. La mayor aversión al riesgo de países en vías de desarrollo en comparación con países desarrollados puede ser explicada, por lo menos en parte, por el simple hecho de que la pobreza en sí reduce la capacidad de los individuos para asumir riesgo. Haushofer y Fehr (2014) exploran esta hipótesis y llegan a la conclusión de que existe un mecanismo causal que determina que los bajos ingresos incrementan la aversión al riesgo.

Al incorporar a los modelos la heterogeneidad no observable entre agentes, notamos dos aspectos: (i) que existe una muy alta heterogeneidad en las preferencias por riesgo y (ii) que los valores de r se concentran más hacia la cola derecha de la distribución –es decir, hay una concentración importante de sujetos muy aversos al riesgo–. Al permitir heterogeneidad según características observables, hallamos que el coeficiente de aversión al riesgo no se correlaciona de manera significativa con la educación, el sexo, la edad o la condición de desempleo, pero sí lo hace con el lugar de residencia del trabajador. En

particular, es mayor para aquellos que viven fuera de Lima, lo cual lleva a pensar en la importancia de factores culturales.

Con respecto a las preferencias de consumo intertemporal, independientemente del modelo estocástico escogido, encontramos que el trabajador peruano descuenta en gran magnitud su consumo futuro. Usando modelos de agente representativo, las tasas de descuento (δ) estimadas son muy altas y fluctúan entre 150% y 200%. Nuestros resultados sugieren que el valor del consumo de un año en el futuro –v. g., año $t + 1$ – del trabajador peruano equivale aproximadamente al 40% del valor que le atribuye al consumo temprano –v. g., año t –. Sin ir muy lejos, esto se refleja en el hecho de que alrededor del 55% del total de encuestados eligieron siempre el pago temprano en todas las decisiones que se les presentó. Esto sorprende, ya que, en el nivel internacional y académico, se suelen encontrar tasas de descuento menores del 25% (Loewenstein y Prelec, 1992). Existen estudios que encuentran valores que sobrepasan el 100%, pero usando otras estrategias empíricas (Frederick et al., 2002). En contraste con países específicos, Andersen et al. (2008) reportan una tasa de descuento de tan solo 10,1% en la estimación que incluye descuento hiperbólico en Dinamarca. En países en desarrollo, Cassar et al. (2017) hallan una tasa de descuento de 19,6% en zonas rurales de Tailandia; y Callen (2015), una de 24,6% en zonas rurales de Sri Lanka, aunque ninguno de estos dos últimos considera alguna forma de descuento hiperbólico. De manera similar, Haushofer y Fehr (2014) también apoyan la noción de que la pobreza puede inhibir la capacidad del individuo de abstenerse del consumo inmediato a cambio de mayor consumo futuro.

Al incorporar heterogeneidad no observable, las tasas de descuento promedio predichas se reducen aproximadamente a 117% –lo cual sigue implicando un descuento alto–. Esta reducción viene acompañada de una alta heterogeneidad entre agentes, aunque menor que la

heterogeneidad para las preferencias ante el riesgo. En particular, notamos una alta densidad de individuos con preferencias de alrededor de 160%, pero también existe una proporción menos concentrada de individuos con tasas de descuento menores, lo cual explica esta diferencia entre los estimados de agente representativo y baja heterogeneidad no observable. Respecto a su relación con características observables, encontramos que la tasa de descuento es mayor para quienes viven fuera de Lima y mientras más edad tenga la persona; en cambio, es menor para los desempleados, lo cual sugiere que los últimos valoran más su consumo futuro que las personas empleadas. Que los desempleados descuenten menos su consumo futuro es una señal de que podrán administrar más responsablemente la CTS o lo recibido por un seguro de desempleo.

Sobre la existencia de sesgo por el presente, usando modelos de agente representativo no identificamos sesgo por el presente, lo que implicaría que el trabajador descuenta su consumo de manera consistente en todos los periodos. Debido al diseño de nuestro experimento, no podemos analizar la heterogeneidad no observable de esta preferencia en el nivel individual. El que no exista un sesgo significativo por el presente no es poco común y, de hecho, es consistente con hallazgos recientes sobre preferencias temporales e incentivos monetarios (Augenblick et al., 2015; Andreoni et al., 2018).

Con relación al ajuste de los modelos estocásticos, hallamos que el que mejor predice las decisiones de los trabajadores es el modelo de parámetros aleatorios, tanto para las elecciones ante el riesgo como para las de tiempo. Además, los modelos con menores niveles de predicción suelen ser los de utilidad aleatoria con error homocedástico –es decir, los modelos de Fechner y Luce–. No obstante, el de Luce ha sido el más flexible a la hora de utilizar modelos complejos. Encontramos también que es importante controlar en este tipo de modelos por

la probabilidad de que el sujeto pierda la concentración y decida aleatoriamente, más aún en diseños experimentales que permiten la existencia de dominancia estocástica, como el nuestro. Finalmente, como era de suponer, los modelos que toman en cuenta la heterogeneidad de las preferencias presentan un nivel de ajuste mucho mayor que los modelos de agente representativo. De hecho, el mayor incremento de ajuste se produce al incorporar heterogeneidad no observable. Esto nos permite volver a enfatizar el alto grado de heterogeneidad en las preferencias ante el riesgo e intertemporales.

Nuestras contribuciones no son exclusivas al ámbito académico, sino que también son relevantes para la elaboración de políticas públicas. Conocer cuán aversos al riesgo e impacientes son los trabajadores arroja luces sobre la viabilidad de políticas que involucren escenarios no certeros y más de un periodo de tiempo, que suelen ser la gran mayoría si no todas las políticas en el debate público.

Por ejemplo, nuestros resultados nos permiten aproximarnos a la pregunta de si una política de seguro de desempleo –en lugar de la CTS– llevaría a un mayor nivel de bienestar, o si es beneficioso un sistema de pensiones obligatorio. Existen muchas aristas por considerar para poder responder a dichas preguntas, que deben ser cuidadosamente analizadas mediante estudios profundos y específicos. Sin ir muy lejos, mientras más averso al riesgo sea un trabajador, mayor será su preferencia por un seguro en caso de despido –ya sea mediante una CTS o un seguro de desempleo–, pero hay que tener en cuenta que, cuando la decisión de consumo es intertemporal, los trabajadores cuentan con otra fuente de aseguramiento, el ahorro privado, fuertemente vinculado a los tres parámetros de preferencia discutidos en este documento. En este estudio, hallamos que los trabajadores peruanos descuentan en gran medida su futuro y esto sugiere que no ahorran lo suficiente. Por sí solo, esto implicaría que, con un enfoque

paternalista y en términos de bienestar social, es conveniente imponer un esquema de ahorro forzado para los trabajadores mediante sistemas de pensiones o CTS.

En torno a la discusión específica sobre fuentes de aseguramiento frente a despidos, debe tomarse en cuenta lo siguiente: (i) tanto el rol de la CTS –que funciona como un sistema de ahorro forzado durante los primeros años de empleo– como el del seguro de desempleo –cuyos fondos no dependen del ahorro, sino de una cuenta pública– sobre los agentes con sesgo hacia el presente, (ii) los incentivos de los desempleados de buscar trabajo y aceptar ofertas laborales mientras reciben bonificaciones por no trabajar –en caso del seguro de desempleo– y (iii) las dinámicas que se desprenden de la demanda laboral, como cambios en los salarios de equilibrio y número de personal. Algunos estudios –como los de Krusell et al. (2010), Popp (2017), y Acemoglu y Shimer (1999)– ya han abordado estos temas para el caso del seguro de desempleo frente a un escenario sin ningún tipo de aseguramiento. Con los datos adecuados y, dada la creciente cantidad de estudios que relacionan patrones de búsqueda de empleo y preferencias, estamos seguros de que se podrían identificar aplicaciones interesantes para nuestros estimados.

Pocos estudios han explorado cómo decide el trabajador peruano, o el peruano en general. Si bien nuestro trabajo –como el de los otros pocos– constituye un buen punto de partida, es necesario profundizar más en los detalles que caracterizan las elecciones. Por ejemplo, nosotros asumimos que se cumple la hipótesis de utilidad esperada y que el descuento intertemporal es cuasihiperbólico; sin embargo, es importante también conocer el ajuste de otros modelos determinísticos, como la teoría prospectiva o la teoría de utilidad dependiente en rangos. Identificar cómo ponderan las probabilidades los individuos y si utilizan algún punto de referencia al tomar decisiones es relevante

para elaborar políticas públicas en el Perú. La agenda de investigación es, pues, amplia.

Asimismo, creemos que podrían hacerse extensiones o modificaciones al diseño experimental que utilizamos en este estudio, como homogeneizar las decisiones de tiempo para todos los individuos, contextualizar los pagos que ofrecen las loterías a escenarios más vinculados al ámbito laboral, y añadir variaciones que permitan la estimación de ponderación de probabilidades y puntos de referencia. Otras extensiones más precisas podrían involucrar asumir heterogeneidad en los parámetros de precisión dentro de los modelos con efectos aleatorios, lo cual, según Barseghyan et al. (2018), podría aminorar los problemas encontrados por Apestequia y Ballester (2018) en los modelos de utilidad aleatoria, como los de Fechner y Luce. Con la finalidad de evaluar el rol del sesgo por el presente en el bienestar social también podrían estimarse modelos mixtos que asuman que existe más de un proceso generador de datos para el comportamiento de los agentes, variar el *front-end-delay* dentro de las decisiones de cada sujeto para mejorar la estimación de sesgos por el presente individuales, modificar los montos de dinero sobre los que deciden los agentes para capturar si la aversión al riesgo o la impaciencia de los agentes varía según la cantidad en juego, o determinar si los sujetos son ingenuos o sofisticados, empleando la terminología de O'Donoghue y Rabin (1999).

REFERENCIAS BIBLIOGRÁFICAS

- Acemoglu, D., y Shimer, R. (2000). Productivity gains from unemployment insurance. *European Economic Review*, 44(7), 1195-1224. [https://doi.org/10.101/S0014-2921\(00\)00035-0](https://doi.org/10.101/S0014-2921(00)00035-0)
- Acemoglu, D., y Shimer, R. (1999). Efficient unemployment insurance. *Journal of Political Economy*, 107(5), 893-928.
- Agranov, M., y Ortoleva, P. (2017). Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125(1), 40-68. <https://doi.org/10.1086/689774>
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716-723.
- Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine. *Econometrica*, 21(4), 503-546.
- Álvarez, F. y Veracierto, M. (2001). Severance payments in an economy with frictions. *Journal of Monetary Economics*, 47(3), 477-498. [https://doi.org/10.1016/S0304-3932\(01\)00058-7](https://doi.org/10.1016/S0304-3932(01)00058-7)
- Anderhub, V., Güth, W., Gneezy, U. y Sonsino, D. (2001). On the interaction of risk and time preferences: an experimental study. *German Economic Review*, 2(3), 239-253.

- Andersen, S., Harrison, G. W., Lau, M. I. y Rutström, E. E. (2016). Multiattribute utility theory, intertemporal utility and correlation aversion. *International Economic Review*, 59(2), 537-555.
- Andersen, S., Harrison, G. W., Lau, M. I. y Rutström, E. E. (2008). Eliciting risk and time preferences. *Econometrica*, 76(3), 583-618.
- Andreoni, J., Gravert, C., Kuhn, M. A., Saccardo, S. y Yang, Y. (2018). *Arbitrage or narrow bracketing? Experimental tests of money as a primary reward*. NBER Working Paper. NBER. <https://doi.org/10.3386/W25232>
- Andreoni, J., Kuhn, M. A. y Sprenger, C. (2015). Measuring time preferences: a comparison of experimental methods. *Journal of Economic Behavior & Organization*, 116, 451-464.
- Andreoni, J. y Sprenger, C. (2012a). Estimating time preferences from convex budgets. *American Economic Review*, 102(7), 3333-3356.
- Andreoni, J. y Sprenger, C. (2012b). Risk preferences are not time preferences. *American Economic Review*, 102(7), 3357-3376.
- Apestequia, J. y Ballester, M. A. (2018). Monotone stochastic choice models: the case of risk and time preferences. *Journal of Political Economy*, 126(1), 74-106. <https://doi.org/10.1086/695504>
- Augenblick, N., Niederle, M. y Sprenger, C. (2015). Working over time: dynamic inconsistency in real effort tasks. *The Quarterly Journal of Economics*, 130(3), 1067-1115.
- Ballinger, T. P. y Wilcox, N. T. (1997). Decisions, error and heterogeneity. *The Economic Journal*, 107(443), 1090-1105.

- Barseghyan, L., Molinari, F., O'Donoghue, T. y Teitelbaum, J. C. (2018). Estimating risk preferences in the field. *Journal of Economic Literature*, 56(2), 501-564.
- Becker, G. M., DeGroot, M. H. y Marschak, J. (1964). Measuring utility by a single-response sequential method. *Systems Research and Behavioral Science*, 9(3), 226-232.
- Becker, G. M., DeGroot, M. H. y Marschak, J. (1963). Stochastic models of choice behavior. *Behavioral Science*, 8(1), 41-55.
- Bernoulli, D. (1738). Specimen theoriae novae de mensura sortis. *Papers of the Imperial Academy of Sciences in Petersburg*, V, 175-192.
- Blavatskyy, P. R. (2014). Stronger utility. *Theory and Decision*, 76(2), 265-286.
- Böhm-Bawerk, E. v. (1889). *Positive theorie des kapitals*. Wagner. Traducido al inglés (con prefacio y análisis) por W. Smart. Macmillan (1891).
- Busemeyer, J. R. y Diederich, A. (2014). Estimation and testing of computational psychological models. En E. F. P. W. Glimcher (Eds.), *Neuroeconomics* (pp. 49-61). Academic Press.
- Callen, M. (2015). Catastrophes and time preference: evidence from the Indian Ocean earthquake. *Journal of Economic Behavior & Organization*, 118, 199-214.
- Camerer, C. F. (1989). An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty*, 2(1), 61-104.
- Cassar, A., Healy, A. y Von Kessler, C. (2017). Trust, risk, and time preferences after a natural disaster: experimental evidence from Thailand. *World Development*, 94, 90-105.

- Chapman, G. B. (1996). Temporal discounting and utility for health and money. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(3), 771.
- Cheung, S. L. (2016). Recent developments in the experimental elicitation of time preference. *Journal of Behavioral and Experimental Finance*, 11, 1-8.
- Clarke, K. A. (2007). A simple distribution-free test for nonnested model selection. *Political Analysis*, 15(3), 347-363.
- Coller, M. y Williams, M. B. (1999). Eliciting individual discount rates. *Experimental Economics*, 2(2), 107-127.
- De Brauw, A., y Eozenou, P. (2014). Measuring risk attitudes among Mozambican farmers. *Journal of Development Economics*, 111, 61-74.
- Duesenberry, J. (1952). *Income, saving, and the theory of consumer behavior*. Harvard University Press.
- Eckel, C., Johnson, C. y Montmarquette, C. (2002). *Will the working poor invest in human capital? A laboratory experiment*. Working Paper, 2002rp-08. CIRANO Project Reports. CIRANO.
- Engen, E. M. y Gruber, J. (2001). Unemployment insurance and precautionary saving. *Journal of Monetary Economics*, 47(3), 545-579. [https://doi.org/10.1016/S0304-3932\(01\)00051-4](https://doi.org/10.1016/S0304-3932(01)00051-4)
- Epstein, L. G. (1983). Stationary cardinal utility and optimal growth under uncertainty. *Journal of Economic Theory*, 31(1), 133-152.
- Epstein, L. G. y Zin, S. E. (1989). Substitution, risk aversion, and the temporal behavior of consumption and asset returns: a theoretical framework. *Econometrica*, 57(4), 937-969.

- Farmer, R. E. A. (1990). RINCE preferences. *The Quarterly Journal of Economics*, 105(1), 43-60.
- Fechner, G. (1860). *Elemente der psychophysik*. Breitkopf & Härtel.
- Fisher, I. (1930). *The theory of interest, as determined by impatience to spend income and opportunity to invest it*. Macmillan.
- Frederick, S., Loewenstein, G. y O'Donoghue, T. (2002). Time discounting and time preference: a critical review. *Journal of Economic Literature*, 40(2), 351-401.
- Galarza, F. (2009). *Choices under risk in rural Peru*. MPRA Paper 17833. University Library of Munich.
- Gandelman, N. y Hernández-Murillo R. (2015). Risk aversion at the country level. *Federal Reserve Bank of St. Louis Review*, 97(1), 53-66.
- Hansen, G. D. y Imrohoroğlu, A. (1992). The role of unemployment insurance in an economy with liquidity constraints and moral hazard. *Journal of Political Economy*, 100(1), 118-142. <https://doi.org/10.1086/261809>
- Harrison, G. W., Lau, M. I., Rutström, E. E. y Sullivan, M. B. (2005). Eliciting risk and time preferences using field experiments: some methodological issues. En R. M. Isaac y D. A. Norton (Eds.). *Field experiments in economics* (pp. 125-218). Emerald Group Publishing Limited. Research in Experimental Economics, Vol. 10.
- Haushofer, J., y Fehr, E. (2014). On the psychology of poverty. *Science*, 344(6186), 862-867. <https://doi.org/10.1126/science.1232491>

- Hey, J. D. y Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica*, 62(6), 1291-1326.
- Holt, C. A. y Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644-1655.
- Jevons, S. (1871). *The theory of political economy*. Macmillan.
- Kahneman, D. y Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, 47(2), 263-292.
- Kim, Y. I. y Lee, J. (2012). Estimating risk aversion using individual-level survey data. *The Korean Economic Review*, 28(2), 221-239.
- Kirby, K. N. (1997). Bidding on the future: evidence against normative discounting of delayed rewards. *Journal of Experimental Psychology: General*, 126(1), 54-70.
- Knight, F. (1921). *Risk, uncertainty and profit*. Vernon Press Titles in Economics. Vernon Art y Science Inc.
- Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica*, 28(2), 287-309.
- Krusell, P., Mukoyama, T., Rogerson, R. y Şahin, A. (2010). Aggregate labor market outcomes: the roles of choice and chance. *Quantitative Economics*, 1(1), 97-127.
- Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2), 443-478.
- Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *The Economic Journal*, 97(387), 666-684.

- Loewenstein, G. y Prelec, D. (1992). Anomalies in intertemporal choice: evidence and an interpretation. *The Quarterly Journal of Economics*, 107(2), 573-597. <http://www.jstor.org/stable/2118482>
- Loomes, G., Moffatt, P. G. y Sugden, R. (2002). A microeconomic test of alternative stochastic theories of risky choice. *Journal of Risk and Uncertainty*, 24(2), 103-130.
- Loomes, G. y Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review*, 39(3-4), 641-648.
- Luce, R. D. (1959). *Individual choice behavior*. John Wiley & Sons.
- Luce, R. D. y Suppes, P. (1965). Preference, utility, and subjective probability. En R. D. Luce, R. R. Bush y E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 249-410). John Wiley & Sons.
- Manski, C. F. (1977). The structure of random utility models. *Theory and Decision*, 8(3), 229-254. <https://doi.org/10.1007/BF00133443>
- McFadden, D. (2001). Economic choices. *American Economic Review*, 91(3), 351-378.
- Moffatt, P. G. (2015). *Experimentics: econometrics for experimental economics*. Macmillan International Higher Education.
- Moscatti, I. (2017). Expected utility theory and experimental utility measurement, 1950-1985. From confidence to scepticism. *The European Journal of the History of Economic Thought*, 24(6), 1318-1354. <https://doi.org/10.1080/09672567.2017.1378692>

- Myerson, J., y Green, L. (1995). Discounting of delayed rewards: models of individual choice. *Journal of the Experimental Analysis of Behavior*, 64(3), 263-276.
- O'Donoghue, T. y Rabin, M. (1999). Doing it now or later. *American Economic Review*, 89(1), 103-124.
- Phelps, E. S. y Pollak, R. A. (1968). On second-best national saving and game-equilibrium growth. *The Review of Economic Studies*, 35(2), 185-199.
- Pollak, R. A. (1970). Habit formation and dynamic demand functions. *Journal of Political Economy*, 78(4, parte 1), 745-763.
- Popp, A. (2017). Unemployment insurance in a three-state model of the labor market. *Journal of Monetary Economics*, 90, 142-157.
- Rachlin, H., Raineri, A. y Cross, D. (1991). Subjective probability and delay. *Journal of the Experimental Analysis of Behavior*, 55(2), 233-244.
- Rae, J. ([1834] 1990). *The sociological theory of capital: being a complete reprint of the new principles of political economy*. Macmillan.
- Samuelson, P. (1937). A note on measurement of utility. *Review of Economic Studies*, 4(2), 155-161.
- Savage, L. J. (1954). *The foundations of statistics*. Wiley Publications in Statistics.
- Senior, N. W. (1836). *An outline of the science of political economy*. W. Clowes & Sons.
- Solow, R. M. (1956). A contribution to the theory of economic growth. *The Quarterly Journal of Economics*, 70(1), 65-94.
<https://doi.org/10.2307/1884513>

- Starmer, C. y Sugden, R. (1989). Violations of the independence axiom in common ratio problems: an experimental test of some competing hypotheses. *Annals of Operations Research*, 19(1), 79-102.
- Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3), 165-180.
- Tanaka, T., Camerer, C. F. y Nguyen, Q. (2010). Risk and time preferences: linking experimental and household survey data from Vietnam. *American Economic Review*, 100(1), 557-571.
- Thaler, R. (1981). Some empirical evidence on dynamic inconsistency. *Economics Letters*, 8(3), 201-207.
- Train, K. E. (2009). *Discrete choice methods with simulation*, 2.^a Ed. Cambridge Books. Cambridge University Press.
- Tversky, A. (1969). Intransitivity of preferences. *Psychological Review*, 76(1), 31.
- Von Neumann, J. y Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.
- Vuong, Q. H. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica*, 57(2), 307-333.
- Wakker, P. P. (2008). Explaining the characteristics of the power (CRRA) utility family. *Health Economics*, 17(12), 1329-1344.
- Wilcox, N. T. (2008). Stochastic models for binary discrete choice under risk: a critical primer and econometric comparison. En J. C. Cox y G. W. Harrison (Eds.). *Risk aversion in experiments*, Vol. 12 (pp. 197-292). Emerald Group Publishing Limited.

A. Detalles adicionales del procedimiento

Figura A.1
Captura de pantalla de una decisión de tiempo

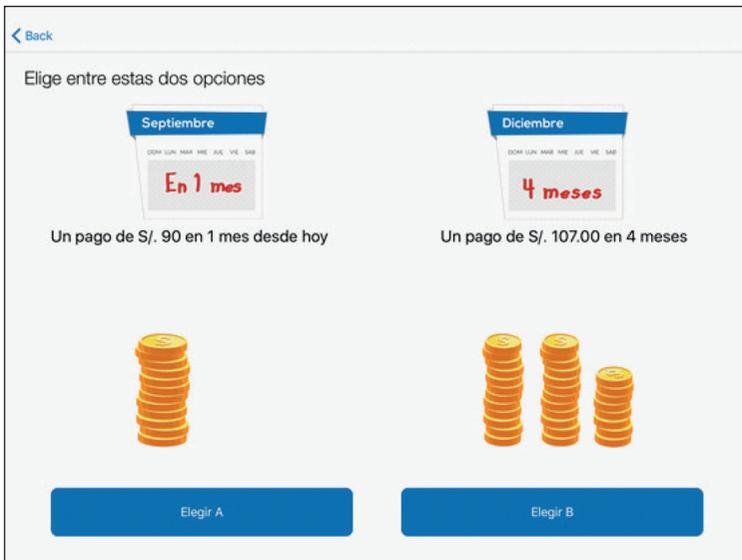


Figura A.2
Captura de pantalla de una decisión de riesgo



Cuadro A.1
Decisiones de tiempo modificadas

Decisión	NFED				FED				TEA (%) s. a. $r=0,8$	Tasa de descuento (%)
	Opción A		Opción B		Opción A		Opción B			
	t	monto	$t+k$	monto	t	monto	$t+k$	monto		
<i>Tarea tiempo 1'</i>										
1	0	50	3	66	1	50	4	66	203,60	24,87
2	0	50	3	75	1	50	4	75	406,25	38,32
3	0	50	3	81	1	50	4	81	588,75	47,10
4	0	50	3	87	1	50	4	87	816,64	55,75
5	0	50	3	91	1	50	4	91	997,20	61,46
6	0	50	3	95	1	50	4	95	1203,21	67,11
7	0	50	3	98	1	50	4	98	1375,79	71,32
8	0	50	3	102	1	50	4	102	1631,89	76,89
9	0	50	3	104	1	50	4	104	1771,77	79,66
10	0	50	3	107	1	50	4	107	1997,27	83,79
<i>Tarea tiempo 2'</i>										
1	0	50	6	62	1	50	7	62	53,76	8,99
2	0	50	6	72	1	50	7	72	107,36	15,70
3	0	50	6	81	1	50	7	81	162,44	21,28
4	0	50	6	89	1	50	7	89	216,84	25,94
5	0	50	6	97	1	50	7	97	276,36	30,35
6	0	50	6	103	1	50	7	103	324,36	33,52
7	0	50	6	110	1	50	7	110	384,00	37,08
8	0	50	6	116	1	50	7	116	438,24	40,02
9	0	50	6	122	1	50	7	122	495,36	42,87
10	0	50	6	127	1	50	7	127	545,16	45,19
<i>Tarea tiempo 3'</i>										
1	0	50	12	63	1	50	13	63	26,00	4,73
2	0	50	12	76	1	50	13	76	52,00	8,73
3	0	50	12	89	1	50	13	89	78,00	12,22
4	0	50	12	102	1	50	13	102	104,00	15,33
5	0	50	12	115	1	50	13	115	130,00	18,13
6	0	50	12	128	1	50	13	128	156,00	20,68
7	0	50	12	141	1	50	13	141	182,00	23,04
8	0	50	12	154	1	50	13	154	208,00	25,23
9	0	50	12	167	1	50	13	167	234,00	27,28
10	0	50	12	180	1	50	13	180	260,00	29,20

B. Rondas de práctica

Con la finalidad de que los sujetos practicara cómo funcionan los juegos, les solicitamos que respondieran a algunas rondas de práctica antes de tomar las decisiones que finalmente usamos para identificar sus preferencias. Tanto para el caso de riesgo como para el de tiempo, se presentaron dos listas: una en la cual lo que estaba en juego era una cantidad de caramelos, que podía variar según la decisión del sujeto;⁵² y otra en la que se jugaba con montos de dinero hipotéticos que, a diferencia de los juegos principales, no iban a ser pagados en caso de que resultara ganador. En total, hubo cuatro listas de práctica: dos para riesgo y dos para tiempo. Todas las decisiones de práctica tenían el mismo formato de las preguntas principales que explicamos en el cuerpo del texto.⁵³

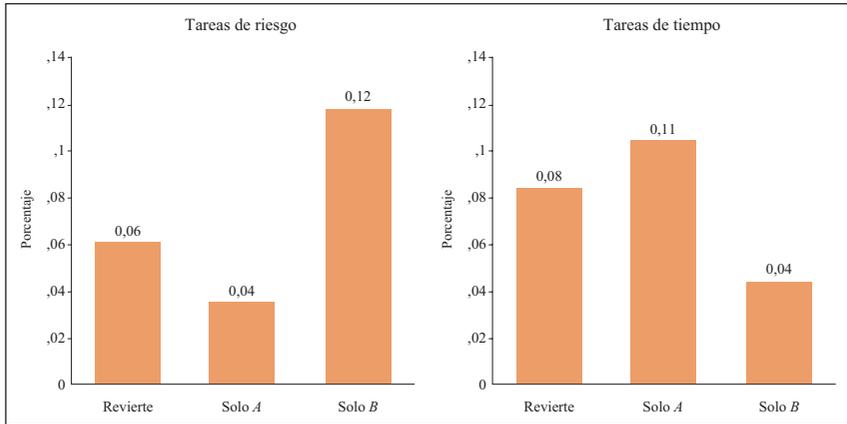
En resumen, las respuestas a las rondas de práctica nos indican que la comprensión de los juegos fue muy positiva en nuestra muestra de sujetos. Como se muestra en la figura B.1, observamos que, en las prácticas de riesgo, solo un 6% revirtió sus preferencias al menos una vez y un 4% eligió solamente la alternativa *A* en ambas listas; estos dos patrones demuestran comportamiento inconsistente o falta de racionalidad por razones que explicamos en el cuerpo principal del documento. Para las prácticas de tiempo, tenemos que solo un 8% revirtió sus preferencias al menos una vez. Cabe resaltar también que estos bajos porcentajes de inconsistencias pueden deberse a que los encuestadores fueron instruidos para corregir a los sujetos si es que

52 En caso de que el sujeto ganara cierta cantidad de caramelos, el encuestador se los entregaba en el momento en que había terminado de responder todas las decisiones de dicha lista.

53 La lista de tiempo con caramelos incluyó solo siete decisiones, a diferencia de todas las demás listas.

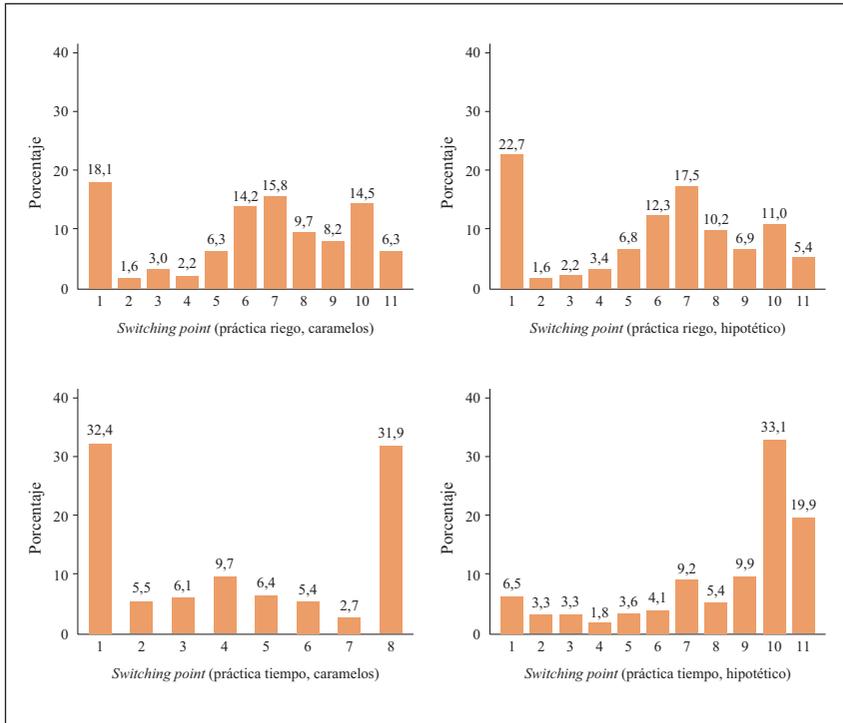
cometían una decisión inconsistente, con la finalidad de hacerles ver su error y de que aprendieran el juego.

Figura B.1
Patrones de racionalidad en tareas de práctica



Luego, en la figura B.2, mostramos las distribuciones de los números de decisión en los que los sujetos cambiaron de la alternativa *A* a la *B* —también conocidos como *switching points*—, para aquellos que no revierten sus preferencias. Para ilustrar la lectura de dichas figuras, en el panel superior izquierdo de la figura tenemos, por ejemplo, que un 15,8% de los sujetos pasaron de elegir continuamente la alternativa *A* a elegir la alternativa *B* en la séptima tarea de caramelos. Para incluir a los sujetos que nunca cambian de elección en cada lista —e. g., solo eligen *A* o solo eligen *B*—, la proporción de los sujetos que solo eligen *B* se muestra como si cambiaran en la decisión número 1 de cada histograma; y la de sujetos que solo escogen *A*, como si cambiaran en la decisión 11 —u 8, para la lista de tiempo con caramelos— de cada histograma.

Figura B.2
Distribución de *switching points* en las cuatro tareas de práctica



Según la distribución de *switching points* en las dos listas de riesgo, vemos que aproximadamente tres cuartas partes de los sujetos cambian de decisión dentro de las 10 decisiones –i. e., identificamos un *switching point*–. Para las tareas de tiempo, vemos que, en la lista con dinero hipotético, la distribución se acumula hacia la derecha; en particular, más de la mitad de los sujetos cambian a elegir el pago tardío en la última decisión o nunca cambian. En la lista con caramelos, también existe un porcentaje considerable que nunca elige el pago tardío, aunque asimismo hay un alto porcentaje que siempre elige recibir el caramelo inmediatamente.

En resumen, las respuestas a las tareas de práctica demuestran que un gran porcentaje no comete inconsistencias, lo cual es un indicador positivo de que entendieron las instrucciones. Además, en las tareas de riesgo, podemos identificar cambios de la alternativa A a la B dentro de las 10 decisiones; en cambio, para las tareas de tiempo, sí observamos porcentajes importantes que se acumulan a los extremos, lo cual es un adelanto de lo que observamos en las respuestas a las tareas de tiempo principales.

C. Implementación de efectos aleatorios

La simulación es utilizada para evaluar integrales que no tienen forma cerrada; es decir, integrales que se deben calcular usando métodos numéricos. Particularmente, este procedimiento es de interés económico en caso de que el analista desee modelar las decisiones de una población con preferencias heterogéneas. En tal caso, la simulación se lleva a cabo para evaluar una integral sobre un término de error aleatorio que aparece en la función de verosimilitud y que, por lo general, no tiene una forma cerrada.

La siguiente expresión muestra el logaritmo de una función de verosimilitud genérica, donde presuntamente se deberá evaluar una integral del tipo mencionado anteriormente para ser maximizada:

$$\log L = \sum_{i=1}^n \ln \int_u \left[\prod_{t=1}^T g(y_{it} | x_{it}; u) \right] f(u) du. \quad (\text{C.1})$$

Aquí, n hace referencia al número de sujetos; y T , al número de decisiones que cada uno tomó. El indicador y_{it} señala el resultado de la decisión t del sujeto i : alternativa A o B . Se asume que las decisiones

dependen de una serie de variables explicativas x_{it} y la variable aleatoria univariante u . Por tanto, u es el efecto aleatorio. Este tiene media 0 y varía entre sujetos, de tal forma que toma un valor único por cada uno y su distribución está dada por $f(u)$.

Considere la integral sobre la variable aleatoria ϵ :

$$I = \int_{-\infty}^{\infty} t(\epsilon) f(\epsilon) d\epsilon, \quad (\text{C.2})$$

donde $f(\epsilon)$ es una función de densidad y $t(\epsilon)$ otra función. La integral de I es $E[t(\epsilon)]$; es decir, puede ser evaluada encontrando el promedio de $t(\epsilon)$ sobre un número de valores de ϵ . La siguiente expresión muestra cómo evaluar I de esta manera:

$$\hat{I} = \frac{I}{R} \sum_{r=1}^R t(\epsilon_r). \quad (\text{C.3})$$

Si la secuencia $\{\epsilon_1 \dots \epsilon_r\}$ es extraída de manera aleatoria y el número de selecciones R es suficientemente grande, esta expresión será una aproximación precisa de la integral de la ecuación (C.2). Para el econometrista, estimar la distribución de u con una estructura de panel donde n sujetos toman T decisiones requiere que se lleve a cabo una serie de selecciones aleatorias sobre la variable ϵ por sujeto.

Las secuencias de Halton

Para el propósito de evaluar integrales, existe una mejor manera de seleccionar instancias de la densidad $f(\cdot)$ que usando un generador de números aleatorios. En específico, hay dos asuntos importantes que deben ser abordados: (i) la cobertura y (ii) la covarianza de la selección.

Con la selección aleatoria independiente, es posible que la selección se enfoque en la cobertura de algunas áreas y no represente bien otras. Un procedimiento que garantice una cobertura homogénea puede aproximar la integral con menos selecciones y mayor precisión.

En cuanto a la covarianza, esta es igual a cero para una selección independiente. Por tanto, procedimientos que garanticen una covarianza negativa entre los números seleccionados pueden reducir la varianza de un estimador como (C.3). Cuando la selección tiene una correlación negativa, la selección de un valor $t(\epsilon)$ mayor que I tenderá a ser seguida por una con un valor menor que I , y por tanto se aproximará con mayor rapidez y precisión a $E[t(\epsilon)]$.

Usando secuencias de Halton se llega a un procedimiento que provee buena cobertura y correlación negativa entre los números seleccionados. Una secuencia de Halton se define en términos de un número primo, p . Por ejemplo, para $p = 2$, la secuencia es:

$$\left\{ \frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{32}, \dots \right\}$$

En la mayoría de los casos, la selección es transformada usando la inversa de la función acumulativa de la distribución normal, $z = \Phi^{-1}(u)$. La transformación escogida dependerá de la distribución de $f(\cdot)$ asumida.

Variación entre sujetos y la función de verosimilitud

Hasta el momento, la exposición de los modelos ha sido llevada a cabo desde el punto de vista de un sujeto, o un sujeto representativo. Ilustraremos lo que sigue tomando el ejemplo para el caso de modelar la decisión ante el riesgo según el modelo de parámetros aleatorios. En

este caso, las probabilidades de la toma de decisión han sido derivadas condicionales a la actitud promedio ante el riesgo, m_i . El supuesto de una población con preferencias ante el riesgo homogéneas es restrictivo en el mejor de los casos, y absurdo en el peor. Por tanto, ahora mostraremos cómo incluir un término de error específico en cada sujeto para capturar las diferencias entre sujetos.

Para lograr nuestro cometido, e introducir la posibilidad de que haya heterogeneidad entre sujetos, debemos permitir que los parámetros de aversión al riesgo varíen entre la población. Entonces, asumimos que la actitud promedio al riesgo, m , proviene de:

$$m \sim N(\mu, \eta^2), \quad (\text{C.4})$$

es decir, m proviene de una distribución normal, con media μ y varianza η .

La siguiente expresión representa la función de log-verosimilitud de la muestra:

$$\log L = \sum_{i=1}^n \ln \left(\int_{-\infty}^{\infty} \prod_{t=1}^T \left[(1-w) \Phi \left(yy_{it} \times \frac{m_i - r_t^*}{\sigma} \right) + \frac{w}{2} \right] f(m; \mu, \eta) dm \right), \quad (\text{C.5})$$

donde el término $f(m; \mu, \eta)$ es la función de densidad normal evaluada en m , yy_{it} toma el valor de -1 o 1 si la elección fue por la alternativa A o B, respectivamente, y los demás parámetros son los ya explicados en el cuerpo del texto. Nótese que esta es la log-verosimilitud particular al modelo de parámetros aleatorios. Podemos asumir, además, que el término $f(m; \mu, \eta)$ tiene la siguiente forma:

$$f(m; \mu, \eta) = \frac{1}{\eta \sqrt{2\pi}} \exp \left[-\frac{(m - \mu)^2}{2\eta^2} \right], \quad (\text{C.6})$$

donde $-\infty < m < \infty$.

Maximizando esta función, estimamos los parámetros μ , σ , η y w . En el caso de este ejemplo, estos parámetros, definidos anteriormente, caracterizan tanto la distribución de preferencias ante el riesgo en la población como el proceso de decisión según el modelo de parámetros aleatorios con tiemble y la teoría de decisión asumida –en nuestro caso, EUT y CRRA–.

D. Resultados adicionales

D.1. Actitudes ante el riesgo

Cuadro D.1
Estimados de riesgo, diferenciado según lista representativa

Modelo:	Parámetros aleatorios		
	Lista 1 (1)	Lista 2 (2)	Lista 3 (3)
r	0,788 (0,029)	1,422 (0,034)	0,823 (0,031)
σ	0,675 (0,024)	0,787 (0,037)	0,736 (0,026)
w	0,234 (0,013)	0,238 (0,013)	0,232 (0,013)
LL	-6696	-6927	-6660
N	12 980	12 980	12 980

Nota: Errores estándar agrupados a nivel individual entre paréntesis. Cada una de las estimaciones se llevó a cabo utilizando el modelo de parámetros aleatorios con parámetro de tiemble y σ flexible. El estimado de r en realidad corresponde al estimado de m , la media de r . LL hace referencia al valor de la log-verosimilitud; y N , al número de observaciones.

Figura D.1
Distribución del coeficiente de aversión al riesgo,
según el modelo de Fechner

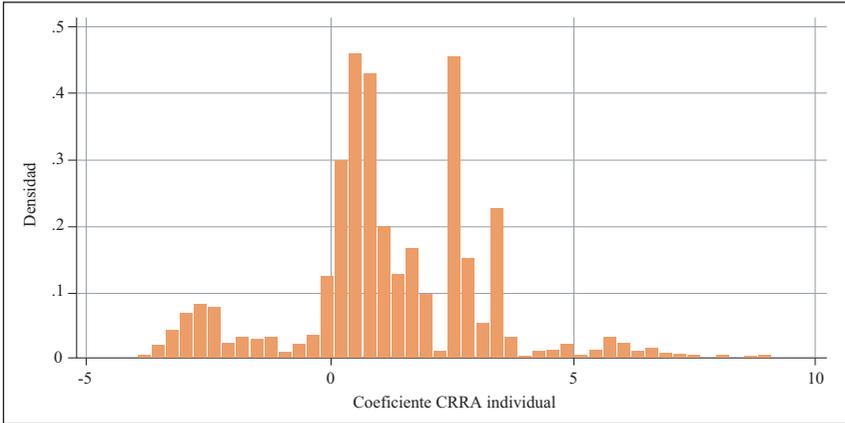
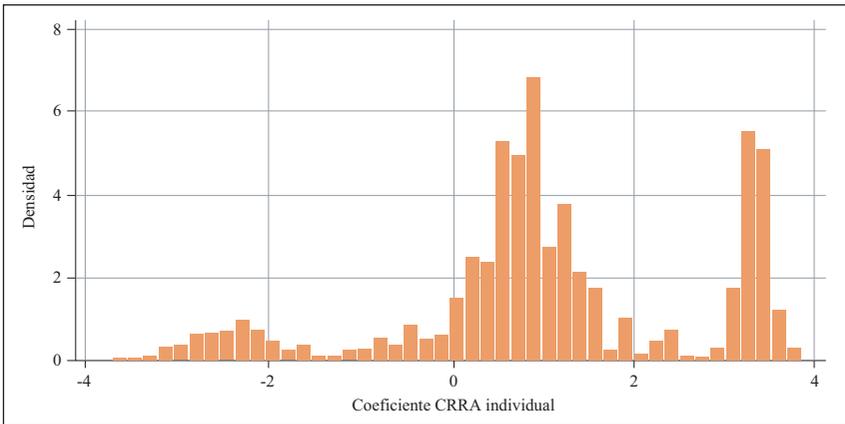


Figura D.2
Distribución del coeficiente de aversión al riesgo,
según el modelo de Blavatsky



D.2. Preferencias intertemporales

Cuadro D.2
Estimados de tiempo usando el mismo CRRA entre modelos

Modelo:	Parámetros aleatorios (1)	Fechner (2)	Luce (3)
δ	2,567 (0,558)	1,169 (0,023)	1,158 (0,020)
β	1,031 (0,029)	1,002 (0,004)	1,001 (0,003)
σ	2,230 (0,508)	0,060 (0,006)	0,008 (0,001)
w	0,000 (0,000)	0,291 (0,014)	0,291 (0,014)
LL	-16 345	-16 366	-16 367
NN	38 940	38 940	38 940

Nota: Errores estándar agrupados a nivel individual entre paréntesis. Cada una de las estimaciones se llevó a cabo utilizando el mismo valor de r , igual a 0,977, que corresponde a la media aritmética de los valores de r estimados de los 10 modelos de agente representativo para decisiones ante el riesgo. Para el caso de las estimaciones con el modelo de parámetros aleatorios, el estimado de δ en realidad corresponde al estimado de m , la media de δ . LL hace referencia al valor de la log-verosimilitud; y N , al número de observaciones.

Cuadro D.3
Estimados de tiempo según lista, para las tareas iniciales

Modelo:	Luce		
	Lista 1 ($k = 3$ meses)	Lista 2 ($k = 6$ meses)	Lista 3 ($k = 12$ meses)
	(1)	(2)	(3)
δ	0.307 (0,025)	0.287 (0,021)	0.260 (0,011)
β	1,004 (0,004)	1,007 (0,006)	1,015 (0,009)
σ	0,015 (0,003)	0,026 (0,005)	0,046 (0,005)
w	0,126 (0,031)	0,097 (0,028)	0,066 (0,019)
LL	-4131	-3844	-3743
NN	9680	9680	9680

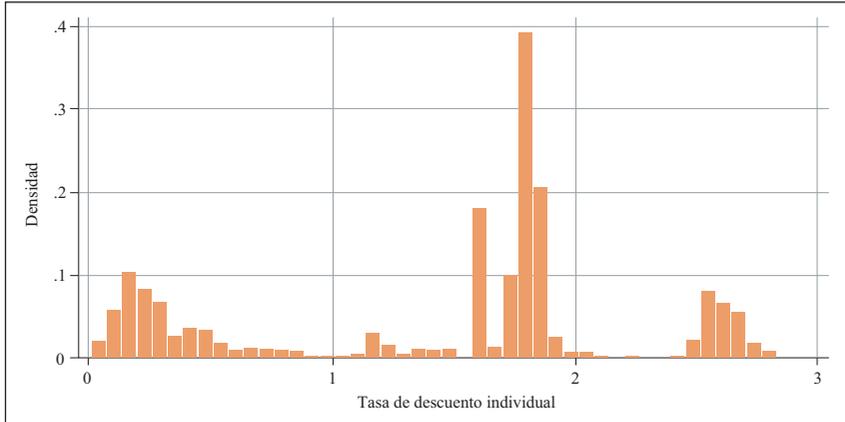
Nota: Errores estándar agrupados a nivel individual entre paréntesis. Cada una de las estimaciones se llevó a cabo utilizando el modelo de Luce con parámetro de tiemle y σ flexible, y un valor de r igual a 0,887. LL hace referencia al valor de la log-verosimilitud; y N , al número de observaciones.

Cuadro D.4
Estimados de tiempo según lista, para las tareas modificadas

Modelo:	Luce		
	Lista 1 ($k = 3$ meses) (1)	Lista 2 ($k = 6$ meses) (2)	Lista 3 ($k = 12$ meses) (3)
δ	1,546 (0,040)	0,820 (0,027)	0,490 (0,017)
β	1,002 (0,005)	1,004 (0,010)	1,003 (0,014)
σ	0,017 (0,002)	0,032 (0,003)	0,046 (0,005)
w	0,142 (0,028)	0,092 (0,023)	0,081 (0,023)
LL	-1366	-1187	-1127
NN	3300	3300	3300

Nota: Errores estándar agrupados a nivel individual entre paréntesis. Cada una de las estimaciones se llevó a cabo utilizando el modelo de Luce con parámetro de tiemble y σ flexible, y un valor de r igual a 0,887. LL hace referencia al valor de la log-verosimilitud; y N , al número de observaciones.

Figura D.3
Distribución de la tasa de descuento,
según la especificación con CRRA imputado



PUBLICACIONES RECIENTES DE GRADE

LIBROS

- 2020 *La dinámica del mercado laboral peruano: creación y destrucción de empleos y flujos de trabajadores*
Miguel Jaramillo y Daniela Campos
- 2019 *Violencia contra las mujeres: la necesidad de un doble plural*
Wilson Hernández (Ed.)
GRADE, CIES y PNUD

DOCUMENTOS DE INVESTIGACIÓN

- 2021 *Reactivando para un mejor país después de la pandemia de COVID-19: una propuesta de diagnóstico de acción para el Perú*
Miguel Jaramillo y Bruno Escobar
Documentos de Investigación, 122
- 2021 *Building back better after the COVID-19 pandemic: a diagnosis and reactivation proposal for Peru*
Miguel Jaramillo y Bruno Escobar
Documentos de Investigación, 121

- 2021 *El efecto de las prácticas parentales en el desarrollo infantil de comunidades urbanas y rurales de la selva peruana*
Juan León, Beatriz Oré, Martín Benavides, Gabriela Conde y Alejandra Miranda
Documentos de Investigación, 120
- 2021 *La educación de niños y niñas con discapacidad en el Perú durante la pandemia del COVID-19*
Gabriela Ramos, Santiago Cueto y Claudia Felipe
Documentos de Investigación, 119
- 2021 *Educación superior en tiempos de pandemia: una aproximación cualitativa a las trayectorias educativas de las y los jóvenes de Niños del Milenio en el Perú*
Vanessa Rojas Arangoitia
Documentos de Investigación, 118
- 2021 *Políticas de protección social y laboral en la República Dominicana*
Hugo Ñopo y Sócrates Barinas
Documentos de Investigación, 117
- 2021 *Políticas de protección social y laboral en el Ecuador*
Hugo Ñopo y Alejandra Peña
Documentos de Investigación, 116
- 2021 *Políticas de protección social y laboral en el Perú: una espiral de buenas intenciones, malos resultados y peores respuestas*
Hugo Ñopo
Documentos de Investigación, 115
- 2021 *What difference do schools make? A mixed methods study in secondary schools in Peru*
Juan León, Gabriela Guerrero, Santiago Cueto y Paul Glewwe
Documentos de Investigación, 114

- 2021 *Estrategias de cuidado infantil en familias vulnerables peruanas: evidencia del estudio cualitativo longitudinal Niños del Milenio*
Vanessa Rojas Arangoitia
Documentos de Investigación, 113
- 2021 *Políticas para combatir la pandemia de COVID-19*
Miguel Jaramillo y Kristian López
Documentos de Investigación, 112
- 2020 *The coronavirus pandemic and its challenges to women's work in Latin America*
Diana Gutiérrez, Guillermina Martín y Hugo Ñopo
Documentos de Investigación, 111
- 2020 *El coronavirus y los retos para el trabajo de las mujeres en América Latina*
Diana Gutiérrez, Guillermina Martín y Hugo Ñopo
Documentos de Investigación, 110
- 2020 *Predictors of school dropout across: Ethiopia, India, Peru and Vietnam*
Santiago Cueto, Claudia Felipe y Juan León
Documentos de Investigación, 109
- 2020 *COVID-19 and external shock: economic impacts and policy options in Peru*
Miguel Jaramillo y Hugo Ñopo
Documentos de Investigación, 108
- 2020 *COVID-19 y shock externo: impactos económicos y opciones de política en el Perú*
Miguel Jaramillo y Hugo Ñopo
Documentos de Investigación, 107

- 2020 *Impactos de la epidemia del coronavirus en el trabajo de las mujeres en el Perú*
Miguel Jaramillo y Hugo Ñopo
Documentos de Investigación, 106
- 2020 *La apuesta por la infraestructura: inversión pública y reproducción de la escasez hídrica en contextos de gran minería en el Perú y Colombia*
Gerardo Damonte, Astrid Ulloa, Catalina Quiroga y Ana Paula López
Documentos de Investigación, 105
- 2020 *Minería y conflictos en torno al control ambiental. La experiencia de monitoreos hídricos en la Argentina, el Perú y Colombia*
Julieta Godfrid, Astrid Ulloa, Gerardo Damonte, Catalina Quiroga y Ana Paula López
Documentos de Investigación, 104
- 2020 *Gobernanzas plurales del agua: formas diversas de concepción, relación, accesos, manejos y derechos del agua en contextos de gran minería en Colombia y el Perú*
Astrid Ulloa, Gerardo Damonte, Catalina Quiroga y Diego Navarro
Documentos de Investigación, 103
- 2020 *Minería, escasez hídrica y la ausencia de una planificación colaborativa*
Gerardo Damonte, Julieta Godfrid y Ana Paula López
Documentos de Investigación, 102

AVANCES DE INVESTIGACIÓN (serie digital)

- 2020 *Entendiendo la paradoja de la maternidad adolescente en Lima Metropolitana: un análisis de los efectos de vecindario en el 2013*
Selene Cueva Madrid
Avances de Investigación, 40
- 2020 *El agua, un anhelo permanente. La minería y sus efectos territoriales sobre el agua en la comunidad afrodescendiente de Patilla, La Guajira, Colombia*
Liza Minely Gaitán Ortiz
Avances de Investigación, 39

Brief de políticas ANÁLISIS & PROPUESTAS

- 2021 *La educación de niños y niñas con discapacidad en el Perú durante la pandemia*
Gabriela Ramos, Santiago Cueto y Claudia Felipe
Análisis & Propuestas, 62
- 2021 *Educación superior en pandemia: una aproximación cualitativa desde Niños del Milenio*
Vanessa Rojas Arangoitia
Análisis & Propuestas, 61
- 2021 *Estrategias de cuidado infantil en familias vulnerables peruanas: evidencia cualitativa de Niños del Milenio*
Vanessa Rojas Arangoitia
Análisis & Propuestas, 60
- 2021 *¿Qué diferencia hacen las escuelas?: un estudio de métodos mixtos en colegios secundarios del Perú*
Juan León, Gabriela Guerrero, Santiago Cueto y Paul Glewwe
Análisis & Propuestas, 59

- 2021 *Estratificación de los retornos a la educación superior en el Perú: el rol de la calidad de la educación y las opciones de carrera*
Alan Sánchez, Marta Favara y Catherine Porter
Análisis & Propuestas, 58
- 2021 *Do more school resources increase learning outcomes?: evidence from an extended school-day reform*
Jorge M. Agüero, Marta Favara, Catherine Porter y Alan Sánchez
Análisis & Propuestas, 57
- 2020 *Acceso a dispositivos y habilidades digitales de dos cohortes en el Perú*
Santiago Cueto, Juan León y Claudia Felipe
Análisis & Propuestas, 56
- 2020 *Perú: construcción de infraestructuras: no perder las oportunidades de reducir las desigualdades en el acceso al agua*
Gerardo Damonte, Ana Paula López y Diego Navarro
Análisis & Propuestas, 55
- 2020 *Colombia: producción de desigualdades en las relaciones con el agua: consecuencias de la construcción de infraestructuras*
Astrid Ulloa, Catalina Quiroga, Liza Gaitán y Gerardo Damonte
Análisis & Propuestas, 54
- 2020 *Chile: nuevos atributos en los protocolos de emergencia ambiental a considerar en el ámbito de la minería*
Julieta Godfrid, Gerardo Damonte y Alex Godoy Faúndez
Análisis & Propuestas, 53
- 2020 *Predictores de la deserción escolar en el Perú*
Santiago Cueto, Claudia Felipe y Juan León
Análisis & Propuestas, 52

- 2020 *Argentina: propuestas para mejorar los mecanismos de evaluación ambiental minera*
Julieta Godfrid, Gerardo Damonte, Ana Paula López y Lautaro Clemenceau
Análisis & Propuestas, 51
- 2020 *Perú: propuestas para mejorar los mecanismos de evaluación ambiental minera*
Julieta Godfrid, Gerardo Damonte, Manuel Glave, Ana Paula López y Diego Navarro
Análisis & Propuestas, 50
- 2020 *Colombia: minería y conflictos en torno a la contaminación del agua: la experiencia de monitoreos hídricos comunitarios en La Guajira, Colombia*
Astrid Ulloa, Catalina Quiroga y Gerardo Damonte
Análisis & Propuestas, 49
- 2020 *Colombia: gobernanzas plurales del agua. Derechos al agua en contextos mineros en Perú y Colombia*
Astrid Ulloa, Gerardo Damonte, Catalina Quiroga y Diego Navarro
Análisis & Propuestas, 48
- 2020 *Colombia: gobernanzas plurales del agua. Derechos al agua en contextos mineros en Perú y Colombia*
Astrid Ulloa, Gerardo Damonte, Catalina Quiroga y Diego Navarro
Análisis & Propuestas, 47
- 2020 *Perú: la planificación colaborativa como solución a la escasez hídrica en contextos de minería a gran escala*
Gerardo Damonte, Julieta Godfrid, Manuel Glave, Ana Paula López y Diego Navarro
Análisis & Propuestas, 46

2020 *Argentina: la planificación colaborativa como solución a la escasez hídrica en contextos de minería a gran escala*

Gerardo Damonte, Julieta Godfrid y Ana Paula López

Análisis & Propuestas, 45

Véanse estas y otras publicaciones en
<http://www.grade.org.pe/publicaciones/>

*Actitudes ante el riesgo y preferencias intertemporales
de los trabajadores peruanos*

se terminó de editar en el
mes de marzo del 2022.

Grupo de Análisis para el Desarrollo
GRADE

Av. Grau 915, Barranco, Lima 15063

Teléfono: 247 9988

www.grade.org.pe

Todo individuo, al tomar decisiones económicas, debe evaluar distintas opciones con resultados inciertos que afectarán su bienestar en diferentes momentos en el tiempo. A partir del encuentro entre la Economía teórica y la experimental, este documento estima las preferencias ante el riesgo y de consumo intertemporal de los trabajadores peruanos.

Los autores contribuyen de dos maneras: estudian cómo los trabajadores toman decisiones considerando el rol de los incentivos de manera “racional” –solo siendo guiados por el beneficio económico– y también el de los factores conductuales que influyen en las decisiones –como la falta de atención, la capacidad cognitiva y la inestabilidad de las preferencias–.

ISBN: 978-612-4374-49-4



9 786124 374494