

El papel del análisis coste-beneficios en la toma de decisiones sanitarias

Metodología para el cálculo de la disposición a pagar como estimador del valor monetario de una mejora en la salud.

Autor: Daniel Pérez Troncoso

Email: danielperez@ugr.es

Director / Tutor: David Mark Epstein

Programa: Doctorado en CC Económicas y Empresariales



Editor: Universidad de Granada. Tesis Doctorales
Autor: Daniel Pérez Troncoso
ISBN: 978-84-1117-527-2
URI <https://hdl.handle.net/10481/77263>

*A mis padres,
mi hermano y
mis abuelos*

Apellidos	Pérez Troncoso
Nombre	Daniel
Director	David M. Epstein
Correo	danielperez@ugr.es
Programa	Doctorado en Ciencias Económicas y Empresariales
Línea	Economía Pública: recaudación, salud, dependencia, educación y gestión del agua
Fecha	Mayo del 2022

TÍTULO

El papel del análisis coste-beneficios en la toma de decisiones sanitarias. Metodología para el cálculo de la disposición a pagar como estimador del valor monetario de una mejora en la salud

RESUMEN

El análisis coste-beneficios es un método de evaluación que clasifica todos los impactos en ventajas y desventajas para computar el beneficio neto y conocer la idoneidad de una decisión. Este requiere la monetización de todos los impactos, costes y beneficios, y aplicado a la evaluación económica de tecnologías sanitarias supondría valorar la salud de las personas en unidades monetarias. Para comodidad de los sanitarios, y debido al enfoque *extra-bienestarista* de la evaluación económica, la monetización de la salud ha sido evitada haciendo uso del análisis coste-utilidad, el cual es adecuado cuando el decisor es superior al propio paciente. No obstante, cuando el enfoque es más cercano al *bienestarista* clásico, y es el propio paciente sobre el que recaen tanto los costes como los beneficios, es conveniente realizar una evaluación basada en un análisis coste-beneficios. Dentro del esquema coste-beneficios, una forma adecuada de medir los impactos positivos es a través de la disposición a pagar. No obstante, la estimación de la disposición a pagar requiere el empleo de metodologías precisas y validadas. En esta tesis se desarrolla de forma teórica y práctica una metodología conocida como experimentos de elección discreta, un método de preferencias declaradas que estima tanto la utilidad como la disposición a pagar de los encuestados por un producto o servicio hipotético. En orden lógico, esta tesis comienza explicando los fundamentos teóricos de los experimentos de elección discreta de forma detallada. Se empieza explicando qué es un experimento de elección discreta y se terminan desglosando los aspectos teóricos más complejos y novedosos. Posteriormente, se presenta un caso de aplicación a una investigación real llevada a cabo y publicada. Esta investigación se empleó para estimar la disposición a pagar de la población española por un servicio de nutrición personalizada. Más tarde, tanto la teoría como la práctica son recuperadas para realizar una contribución metodológica con el objetivo de reducir el número de muestra requerido para conseguir unas estimaciones fiables. Finalmente, se presenta un programa de código abierto para facilitar las tareas de la comunidad investigadora que emplea experimentos de elección discreta.

CONTENIDOS

<i>CAPÍTULO 1. INTRODUCCIÓN</i>	7
1.1 Evaluación económica y análisis coste-beneficios	9
1.2 Estimación de la DAP. Métodos de preferencias declaradas	11
<i>CAPÍTULO 2. EXPERIMENTOS DE ELECCIÓN DISCRETA</i>	13
2.1 Marco conceptual	15
2.2 Marco teórico	17
2.3 Construcción del experimento de elección discreta	24
2.4 Realización del experimento de elección discreta	40
2.5 Algunos detalles	41
2.6 Análisis del experimento de elección discreta	42
2.7 Conclusiones	52
<i>CAPÍTULO 3. EXPERIMENTOS DE ELECCIÓN DISCRETA</i>	53
3.1 Introducción	55
3.2 Métodos	56
3.3 Resultados	63
3.4 Discusión y conclusiones	69
<i>CAPÍTULO 4. EXPLORANDO MÉTODOS DE OPTIMIZACIÓN DE EXP. DE ELECCIÓN DISCRETA</i>	73
4.1 Contexto	75
4.2 Estrategia secuencial óptima para DCEs	75
4.3 Métodos	76
4.4 Resultados	79
4.5 Discusión y conclusiones	84
<i>CAPÍTULO 5. DISEÑOS EXPERIMENTALES EFICIENTES Y ACCESIBLES</i>	87
5.1 Introducción	89
5.2 Detalles técnicos	90
5.3 Generación de un DCE	93
5.4 Validez interna	100
5.5 Carga computacional	103
5.6 Conclusiones	105
<i>CAPÍTULO 6. CONCLUSIONES</i>	107
<i>APÉNDICE TEMA 2</i>	124
<i>APÉNDICE TEMA 3</i>	127
<i>APÉNDICE TEMA 4</i>	137
<i>APÉNDICE TEMA 5</i>	147

Capítulo 1. Introducción



1.1 EVALUACIÓN ECONÓMICA Y ANÁLISIS COSTE-BENEFICIOS

El análisis coste-beneficios (ACB) es un método de evaluación que clasifica todos los impactos que un proyecto tiene sobre la sociedad en costes (desventajas) y beneficios (ventajas) (Boardman et al., 2012). Este método es generalmente empleado para evaluar la idoneidad de políticas públicas y puede realizarse tanto *ex ante* (con anterioridad a la toma de la decisión) como *ex post* (tras la toma de la decisión).

En un ACB básico, a cada impacto (clasificado en *pro* o *contra*) se le asigna un valor en unidades monetarias. Una vez que cada impacto tenga asignado un valor monetario preciso e insesgado, el decisor podrá obtener el beneficio neto del proyecto y tomar una decisión acorde al resultado (expresión 1). Solo cuando el beneficio neto es positivo, el proyecto supondrá una mejora de la situación de *statu quo*, y deberá llevarse a cabo (Boardman et al., 2012). Sin embargo, la monetización de los impactos no es tarea fácil y está detrás de gran parte del desarrollo metodológico en esta materia (McIntosh et al., 2010).

$$\text{Beneficio neto} = \text{Beneficios} - \text{Costes} \quad (1)$$

En la evaluación económica de las tecnologías sanitarias, los programas¹ también tienen costes y beneficios. Para determinar si, por ejemplo, un tratamiento debe ser financiado por un sistema sanitario público, los sistemas sanitarios nacionales se han dotado de agencias de evaluación de tecnologías sanitarias como la RedETS en España, el NICE en Inglaterra y Gales, el HAS en Francia o el SMC en Escocia. Estas agencias, encabezadas por el NICE, han promocionado el empleo de metodologías como el análisis coste-utilidad (ACU) para informar las decisiones del sistema sanitario (Drummond et al., 2015). Aunque el ACU es conceptualmente similar al ACB, sirve a un propósito concreto. Ya que el cometido principal de los sistemas sanitarios es maximizar la salud de la población dada una restricción presupuestaria, el ACU y los umbrales de coste-efectividad empleados son especialmente útiles para mantener la eficiencia del sistema y no desplazar a tratamientos más coste-efectivos. El ACU ha sido considerado un marco adecuado para tomar decisiones cuando el decisor es una institución, como una agencia de evaluación, pues la mayoría de los costes son costes sanitarios, y la mayoría de los beneficios se reportan en términos clínicos o de salud (Figura 1.1). En particular, el ACU evita explícitamente valorar la salud de las personas en términos monetarios, que para algunos sanitarios ha resultado problemático en términos éticos. Esta metodología escapa del enfoque clásico del *bienestarismo* ('welfarism') y adopta un enfoque que se conoce como *extra-bienestarismo* ('extra-welfarism'). Este enfoque, a diferencia del clásico, acepta el uso de resultados más allá de la utilidad, permite ponderar los efectos a

¹ Entendido como cualquier actuación genética que proporcione algún rendimiento en salud (directo o indirecto) al usuario (desde un entrenamiento o dieta, hasta un electrocardiograma o trasplante).

través de otros principios diferentes a las preferencias de los propios afectados, y realiza una asignación diferente a la propuesta por la economía paretiana (Brouwer et al., 2008). Estas características son adecuadas para el ACU, donde las ponderaciones de los resultados en salud se basan en las preferencias de la población en general (no del paciente) y el decisor, el sistema sanitario, asigna los recursos priorizando un tratamiento sobre otro.

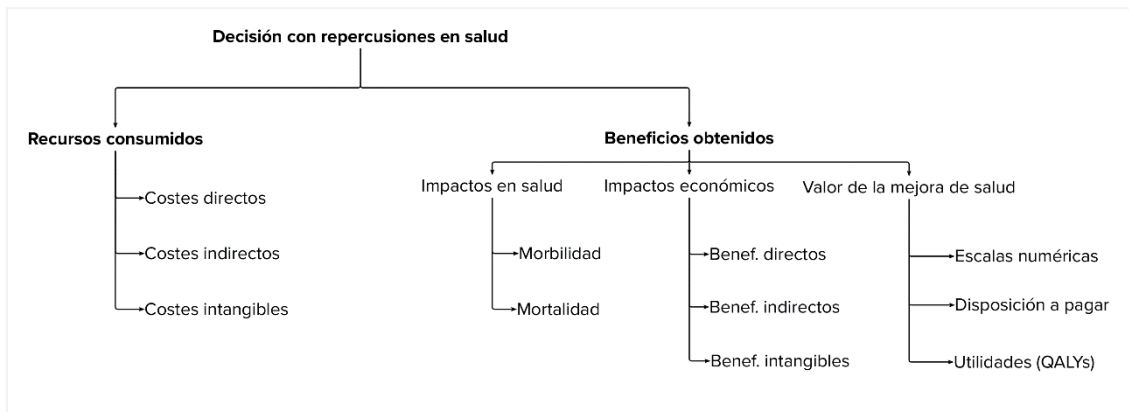


Figura 1.1 Componentes de la evaluación económica. Reproducción de Torrance (1986)

No obstante, el ACU no es ni la única ni (presumiblemente) la mejor alternativa para evaluar todas las decisiones rivales, y este punto puede ser ilustrado con un ejemplo. De entre todas las decisiones que toma una persona al día, se estima que alrededor doscientas están relacionadas con la alimentación (Wansink & Sobal, 2016). Dada la evidencia existente sobre la importancia de la nutrición (McGuire, 2011; Mozaffarian, 2016; Zeevi et al., 2015), tendríamos que aceptar que una gran parte de nuestra salud recae sobre decisiones personales y familiares, como el tipo de dieta, los horarios y hábitos, entre otros, más allá del foco principalmente del ACU empleado por sector sanitario (Epstein, 2021).

Cuando el centro de decisión ya no es una entidad superior sino la propia persona, esta puede sopesar, aunque sea de manera inconsciente, los costes y beneficios que le reporta la toma de la decisión. En este caso, la utilidad que obtienen o pierden los individuos es la principal fuente de la evaluación, las ponderaciones están basadas en sus preferencias y las comparaciones interpersonales no son necesarios, por lo que el marco teórico *bienestarista* ('welfarist') más clásico es aplicable y es factible realizar un ACB (Brouwer et al., 2008).

En el análisis coste-beneficio, la principal dificultad recae en la monetización de los beneficios asociados con la intervención, el servicio o el producto en evaluación, teniendo en cuenta que el beneficio percibido por el individuo no se limite a su propia salud, sino una amplia gama de otros atributos. Una medida teóricamente correcta para estimarlos es la disposición a pagar (DAP), la cual representa la cantidad de unidades monetarias máximas que la persona pagaría para obtener

una determinada mejora (George W Torrance, 1986). El empleo de la DAP en el ACB es conceptualmente ventajoso pues permite comparar los costes de una decisión que pagará el individuo, por ejemplo, el abono anual de un gimnasio, con lo que pagaría por obtener la salud y otras ventajas que le reportará un año de actividad deportiva. No obstante, la estimación de la DAP no está exenta de dificultades prácticas sólo salvables a través de metodología robusta y fundamentada.

1.2 ESTIMACIÓN DE LA DAP: MÉTODOS DE PREFERENCIAS DECLARADAS

A diferencia de los métodos de preferencias reveladas, los cuales permiten la investigación de las preferencias de los consumidores a través de la observación de sus decisiones (Samuelson, 1938), los métodos de preferencias declaradas se basan en encuestas y escenarios hipotéticos para comprender el comportamiento de las personas involucradas y como compensan unas decisiones con otras (Blomquist, 2015). Estos permiten evaluar intervenciones, servicios o productos que todavía no se han comercializado, habilitando a los investigadores a entender la posible demanda del producto y mejorar su diseño. Pese a que los estudios basados en escenarios hipotéticos plantean dudas al respecto de la fiabilidad de sus resultados, y hay autores que piensan que *quien realiza preguntas hipotéticas obtiene respuestas hipotéticas* (Johnston et al., 2017), la metodología hipotética, en especial la que trataremos en profundidad en este trabajo, ha disfrutado de un gran desarrollo y refinamiento durante los últimos años (Haghani, 2021). En los métodos de preferencias declaradas, tanto la forma de preguntar como la manera en la que se presentan los escenarios hipotéticos tienen una gran importancia (Burcharth et al., 2007). Precisamente, estas características han definido los dos grandes tipos de métodos de preferencias declaradas que existen: Valoración Contingente (VC) y modelización de la elección (ME) (Hicks, 2002; Pearce et al., 2006; Willis, 2014).

La Valoración Contingente (VC) es un *término paraguas* bajo el que se clasifican diferentes métodos de encuestación, y cuyo principal cometido es obtener la DAP de los consumidores por diferentes productos o servicios (Hausman, 2012a; Mitchell & Carson, 1989). Dentro de la VC, existen diferentes clasificaciones basadas en el tipo de pregunta (abierta o cerrada) y el número de preguntas que se realizan (una o repetidas) (Klose, 1999). No obstante, como también ocurre con otros métodos de preferencias declaradas, no todos los autores han respetado ni la clasificación ni la nomenclatura, por lo que a veces es difícil saber qué método de VC se ha empleado (Carson & Louviere, 2011; Louviere et al., 2010). Algunos de los métodos de VC más populares son: a) preguntas abiertas, b) elección dicotómica, c) juego de pujas, y d) carta de pago. Pese a la popularidad de la que ha gozado la Valoración Contingente desde finales de los años setenta y principios de los ochenta (Loomis, 1990), sus sesgos y problemas empíricos han

hecho que algunos de sus más grandes promotores hayan pasado *de dudosos a desesperanzados* (Hausman, 2012b) respecto a la VC.

Dentro de los métodos de modelización de la elección, este trabajo está enfocado en un tipo de método del que destaca su robustez metodológica (Louviere et al., 2010), constante evolución (Rose & Bliemer, 2014) y vigencia (Haghani et al., 2021): el Experimento de Elección Discreta (DCE por sus siglas en inglés). Los DCEs proporcionan información de vital importancia cuando no existen datos de mercado, y han sido frecuentemente empleados para disciplinas tan dispares como Valoración Medioambiental, Economía del Transporte, o Economía de la Salud (Haghani et al., 2021). Pese a ser una herramienta circunscrita al ámbito del ACB, tiene un papel clave en el ACU de la Economía de la Salud, pues ha sido ampliamente usada para ponderar la utilidad de cada estado de salud de encuestas como EQ-5D o SF-36 (Ramos-Goñi et al., 2018; Stoolk et al., 2008). Los numerosos métodos de análisis de DCEs nos brindan la oportunidad de estimar tanto las utilidades como las DAPs por los bienes o servicios analizados, además de explorar relaciones de causalidad entre las características sociodemográficas y las elecciones de los participantes.

A continuación, se realiza un desarrollo lógico con el objetivo de ofrecer al lector todos los detalles necesarios para entender las contribuciones prácticas y metodológicas de esta tesis. El segundo capítulo de este trabajo está dedicado a sentar las bases teóricas y prácticas de los experimentos de elección discreta. El tercer capítulo incluye una aplicación de la metodología al ámbito de la nutrición personalizada. El cuarto capítulo contiene una propuesta metodológica. El quinto capítulo presenta un API (application programming interface) de software abierto, libre y gratuito desarrollado por el autor para el uso de la comunidad de investigadores, lo que simplifica considerablemente las consideraciones técnicas asociadas con el diseño de un DCE. En el sexto capítulo se resumen las conclusiones del trabajo.

Capítulo 2. Experimentos de Elección Discreta



2.1 MARCO CONCEPTUAL

2.1.1 ¿Qué es un experimento de elección discreta?

Los experimentos de elección discreta (DCE por sus siglas en inglés) son empleados para obtener información sobre las preferencias de un grupo de personas por productos o servicios hipotéticos. Sus orígenes se remontan a 1927, año en el que se publica la *Ley de juicio comparativo* (Thurstone, 1927) donde se defiende que las elecciones entre dos ejemplares pueden explicarse en función de sus características diferenciales. En 1974, Daniel McFadden aplicó esta misma lógica al análisis económico y dio lugar a la Teoría de la Utilidad Aleatoria (RUT por sus siglas en inglés; *Random Utility Theory*) (McFadden, 1974). La RUT propone que la utilidad que obtiene un individuo de un producto o servicio depende de variables observables (conocidas como atributos) y variables inobservables (como sus características personales) (McFadden, 1994; Train, 2009). Esta definición permitió a McFadden realizar una aproximación de las probabilidades de elección de un producto hipotético sobre otro a través de la estimación de la función de utilidad subyacente usando el modelo *logit* condicional (CL por sus siglas en inglés; *conditional logit*), una regresión logística similar a al modelo multinomial (McFadden, 1974).

En la práctica, en un DCE los encuestados deben responder a varias preguntas. Estas preguntas son conocidas como *choice sets* y consisten en una elección entre varias alternativas de un mismo producto o servicio que solo se diferencian en algunos o todos los niveles de sus atributos (Louviere & Lancsar, 2009). Para ilustrar la explicación, podemos apoyarnos en una versión modificada del ejemplo empleado en Carlsson & Martinsson (2003): supongamos que una persona tiene que elegir entre dos tratamientos con precio y efectividad (atributos) diferentes. Supongamos que la alternativa A tiene un 80% de efectividad y cuesta 100€, mientras que la alternativa B tiene un 90% de efectividad y cuesta 120€ (cada uno de los valores que puede tomar un atributo, por ejemplo 100€ y 120€, es conocido como un nivel). Si la persona investigada escoge la alternativa B, podemos deducir que 1) su utilidad aumenta cuando aumenta el nivel de efectividad y 2) la utilidad de perder 20€ es más pequeña que la de ganar un 10% de efectividad. En un DCE estándar, los encuestados se enfrentan a varias decisiones de este tipo donde las alternativas contienen, generalmente, entre tres y siete atributos (Johnson et al., 2013).

Los DCEs han sido aplicados a numerosas disciplinas como el Ecologismo (Carson & Czajkowski, 2014; Hoyos, 2010; Rakotonarivo et al., 2016), Teoría del Consumidor (Cantillo et al., 2020; Confraria et al., 2017; Güney & Giraldo, 2020), Economía del Transporte (Gundlach et al., 2018; König & Grippenkov, 2020; Monchambert, 2020), y, por supuesto, Economía de la Salud (Bansback et al., 2012; Determann, Lambooi, Steyerberg, et al., 2017; Shah et al., 2015), entre otras. En esta última disciplina, Economía de la Salud, existe una importante tendencia al alza de

esta metodología que comenzó a principios de siglo y que se reforzó a mediados de la década de 2010 (Wang et al., 2021). En este ámbito, los experimentos de elección discreta se han empleado para diferentes cometidos, entre ellos para atraer a sanitarios a zonas rurales (WHO, 2012), para informar decisiones sanitarias de incorporación de tratamientos (Lancsar & Louviere, 2008), para estimar las preferencias de pacientes y/o expertos por diferentes terapias (Soekhai et al., 2019), e incluso para estimar las tarifas de los QALYs² usadas en la actualidad (Devlin et al., 2018; Ramos-Goñi et al., 2018).

2.1.2 Resultados de un experimento de elección discreta

En un DCE se obtienen dos tipos de estimaciones principales. Por un lado, el beneficio obtenido por contar (o no) con cada uno de los niveles de los atributos del producto se conoce como *utilidad* (Hauber et al., 2016). Por ejemplo, la Figura 2.1 es una sencilla representación gráfica de los beneficios de las características del producto. Esta estimación permite realizar una comparación inter e intra-atributo, proporcionando una perspectiva en la que se puede conocer la característica más valorada por las personas encuestadas, la intensidad de esa preferencia en relación con el resto, y la heterogeneidad de las preferencias dentro de la población analizada.

Por otro lado, los experimentos de elección discreta también permiten calcular la disposición a pagar (DAP o WTP por sus siglas en inglés) por cada característica del producto analizado. Esto es posible pues los DCE contienen un atributo precio en el cual se estima la utilidad perdida por cada unidad monetaria gastada, permitiendo compensar la utilidad ganada por cada característica del producto (J. Chang et al., 2017) (se explica más adelante).

2.1.3 Elementos de un Experimento de Elección Discreta

Antes de continuar con la siguiente sección, debe introducirse la terminología que va a emplearse. Los DCE se administran en forma de encuesta y están compuestos por varios elementos: diseño experimental, bloques, choice sets, alternativas, atributos y niveles (Figura 2.2).

- Choice sets (1): los choice sets son las unidades básicas del diseño experimental de un DCE y están compuestos por dos o más alternativas (Street et al., 2005) (la figura 2.2 representa a un choice set).

² Cada estado de salud estimado con el cuestionario EQ-5D-5L representa una utilidad diferente. Para conocer qué estados de salud son más o menos graves que el resto, se emplea un DCE para conocer la utilidad que reporta cada estado de salud. Esta utilidad, es expresada entre cero y uno (donde cero representa a la muerte y uno a un estado de salud perfecta). A un año de vida perfecta se le conoce como QALY (quality-adjusted life-year) o AVAC (año de vida ajustado por la calidad) (Torrance, 1986).

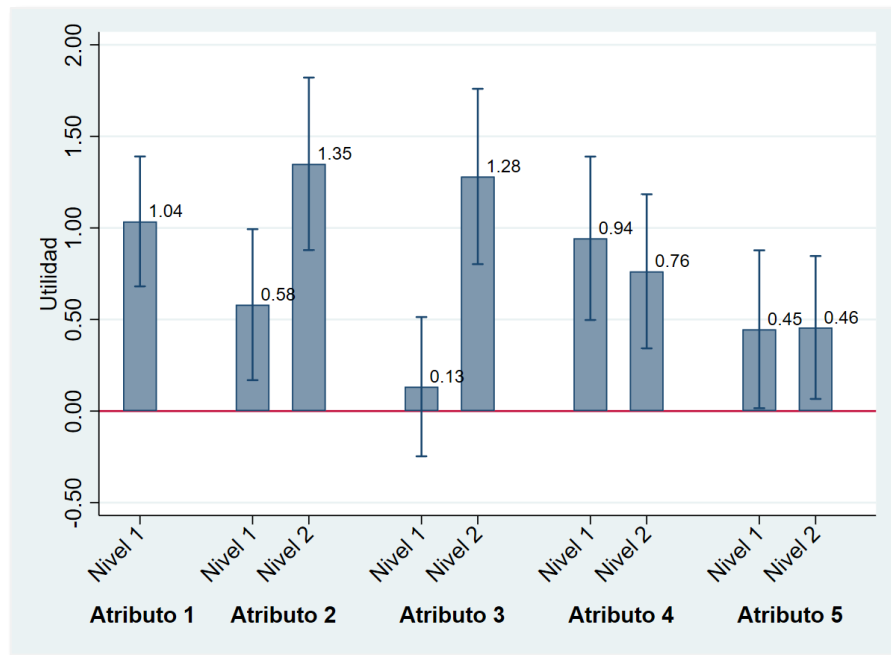


Figura 2.1 Representación de la utilidad aportada por cada atributo y nivel

- Alternativas (3): las alternativas son diferentes configuraciones hipotéticas de un producto o servicio entre las que las personas encuestadas deben elegir. Cada alternativa tiene los mismos atributos, pero se compone de diferentes niveles. Como excepción a este último punto, en muchos DCEs se introducen *alternativas nulas* u *opt-out(s)* que pueden ser escogidas por la persona encuestada si prefiere no hacer ninguna elección.
- Atributos (2): los atributos categorizan las características de los productos. Por ejemplo, cuando decimos que un tratamiento es efectivo en un 90% de los casos, se está hablando de su atributo *efectividad*.
- Niveles (4): los niveles representan los valores que los atributos pueden tomar. Siguiendo el ejemplo anterior, un atributo *efectividad* podría tomar cualquier valor porcentual de 0 a 100. No obstante, los niveles deben ser finitos y especificados con anterioridad.
- Bloques: cuando el número de *choice sets* del experimento supera el número de *choice sets* que un encuestado puede contestar sin sufrir demasiada fatiga, se crean varias versiones de la encuesta con diferentes *bloques* de choice sets. De esta forma, un grupo de encuestados responde a un bloque mientras otro grupo responde otro bloque.

2.2 MARCO TEÓRICO

2.2.1 Fundamentos teóricos: la Teoría de la Utilidad Aleatoria y el logit condicional

- La Teoría de la Utilidad Aleatoria sostiene, como se ha mencionado en la introducción, que los *gustos* de una persona sobre un producto o servicio en particular dependerán de las características del objeto, las características observadas del individuo y del azar.

Mientras que las características del objeto dependen del mismo objeto que se analiza y son observables, el componente de azar es introducido por las características idiosincráticas de la persona que no se observan. A continuación, se formaliza esta idea siguiendo a Train (2009) y Greene (2003).

1 Choice set 1

Atributos	Alternativa A	Alternativa B
Personalización	Nivel 1	Nivel 3
Provisión	Tradicional	Digital
Rutinas deportivas	Incluidas	No
Precio	20 €/mes	30 €/mes
Elección	Elegir la alternativa A <input type="checkbox"/>	Elegir la alternativa B <input type="checkbox"/>

Figura 2.2 Representación de un *choice set* en un DCE

Supongamos que una persona decisora n tiene que escoger entre J alternativas dentro de un mismo *choice set*. Cada una de las alternativas j , supongamos $j = 1, 2$ por simplicidad, le reporta una utilidad (beneficio) U_j . Atendiendo a la RUT, la persona decisora escogerá aquella alternativa que le reporte mayor utilidad. Es decir, la persona n escogerá la alternativa j_1 sí y solo sí $U_{n1} > U_{n2} \forall 1 \neq 2$. Pese a que la decisora conoce la utilidad exacta que le reporta cada alternativa, la parte investigadora sólo observa las elecciones de la decisora, y por tanto las características de la alternativa elegida, llamémoslas x_j (x_1 para el caso de la alternativa elegida), y algunas de las características sociodemográficas de la decisora, s_n . De esta forma, podemos representar la utilidad que obtiene la decisora como $V_{nj} = V(x_j, s_n) \forall j$. A esta parte de la utilidad, V_{nj} , se la conoce como utilidad determinista. No obstante, $V_{nj} \neq U_{nj}$ pues es imposible observar todas las variables que la persona decisora tiene en cuenta para realizar la elección. Por este motivo, la utilidad U_{jn} es comúnmente descompuesta como $U_{nj} = V_{nj} + \varepsilon_{nj}$, donde ε_{nj} representa a todos las variables que afectan a la decisión, pero no están comprendidas en V_{nj} ($\varepsilon_{nj} = U_{nj} - V_{nj}$). Puesto que no se conoce $\varepsilon_{nj} \forall j$, la parte investigadora trata ε_{nj} como una variable aleatoria (he aquí el componente de azar). A partir de lo establecido anteriormente, podemos realizar el siguiente desarrollo (2.1):

$$\begin{aligned}
 Prob_{n1} &= Prob(U_{n1} > U_{n2}) \\
 &= Prob(V_{n1} + \varepsilon_{n1} > V_{n2} + \varepsilon_{n2}) \\
 &= Prob(\varepsilon_{n2} - \varepsilon_{n1} < V_{n1} - V_{n2}).
 \end{aligned}
 \tag{2.1}$$

Por tanto, la probabilidad de que la alternativa j_1 sea escogida es la probabilidad de que $\varepsilon_{n2} - \varepsilon_{n1}$ sea menor a una cantidad que sí podemos observar: $V_{n1} - V_{n2}$. Por tanto, esta probabilidad podrá ser hallada integrando la función de distribución de la variable aleatoria ε_{nj} .

Diferentes modelos realizan diferentes suposiciones sobre la función de distribución de la variable ε_{nj} . En el desarrollo de McFadden (1974), se asume que ε_{ij} es una variable independiente e idénticamente distribuida (iid) que sigue una distribución Gumbel tipo I,

$$f(\varepsilon_{nj}) = e^{-\varepsilon_{nj}} e^{-e^{-\varepsilon_{nj}}}, \quad (2.2)$$

y una función acumulada

$$F(\varepsilon_{nj}) = e^{-e^{-\varepsilon_{nj}}}. \quad (2.3)$$

Una vez realizada esta suposición, sólo queda mencionar que la diferencia entre dos variables iid que siguen una distribución Gumbel I se distribuye de forma logística:

$$\varepsilon_{nji}^* = \varepsilon_{nj} - \varepsilon_{ij}, \quad (2.4)$$

$$F(\varepsilon_{nji}^*) = \frac{e^{\varepsilon_{nji}^*}}{1 + e^{\varepsilon_{nji}^*}}. \quad (2.5)$$

No obstante, esta fórmula sólo será posible si se cumplen los supuestos de iid. Para que se cumplan estos supuestos, el experimento debe estar diseñado de tal forma que la información observable para calcular V_{nj} sea una buena representación de U_{nj} , dejando que ε_{nj} sea sólo *ruido blanco*.

Por otro lado, Train (2009) expone de la siguiente forma la obtención de las probabilidades logit alcanzada por McFadden (1974):

$$\begin{aligned} Prob_{ni} &= Prob(U_{ni} > U_{nj} \forall j \neq i) \\ &= Prob(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \forall j \neq i) \\ &= Prob(\varepsilon_{nj} < \varepsilon_{ni} + V_{ni} - V_{nj} \forall j \neq i) \end{aligned} \quad (2.6)$$

Esta probabilidad tendría que ser calculada como la integral de $P_{ni}|\varepsilon_{ni}$ para todos los valores de ε_{ni} ponderados por su densidad:

$$Prob_{ni} = \int \left(\prod_{j \neq i} e^{-e^{-(\varepsilon_{ni} + V_{ni} - V_{nj})}} \right) e^{-\varepsilon_{ni}} e^{-e^{-\varepsilon_{ni}}} d\varepsilon_{ni}, \quad (2.7)$$

que, tras una manipulación algebraica que puede encontrarse en el Capítulo 3 de Train (2009), se transforma en:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} = \frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}} \quad (2.8)$$

que se trata de la fórmula de probabilidades logit. Este resultado implica que la función de utilidad de la RUT puede ser estimada de forma bastante aproximada y sencilla en cualquier programa estadístico.

Usando una versión simplificada de la investigación llevada a cabo en el Capítulo 3, podemos presentar un ejemplo ilustrativo. Supongamos que existe una persona n que tiene que decidir entre dos servicios de *nutrición personalizada* alternativos. Estos servicios se caracterizan por su nivel de personalización y su formato de provisión. El nivel de personalización puede ser alto o bajo, y el formato de provisión puede ser digital o tradicional. A parte de estas características del producto, la persona decisora también cuenta con características que podrán inclinar la decisión hacia un lado u otro (trabajo, personas a cargo, percepciones sobre la nutrición, habilidad para la cocina, etc.). Si decidimos que estas características son tantas que son imposibles de considerar, podemos introducirlas dentro del término de error ε_n . En este contexto, supongamos que a esta persona se le presentan dos alternativas: la alternativa A y la B. Podemos modelizar la utilidad que n obtiene de la alternativa A como $U_A = \beta_1 P_A + \beta_2 F_A + \varepsilon_A$ (nótese que hemos eliminado el subíndice n pues en este momento nos estamos refiriendo sólo a esta persona), donde P_A es el nivel de personalización de la alternativa A, F_A es el formato de provisión de la alternativa A y β_1, β_2 son los parámetros que tomaran uno u otro valor según sean P y F . Igualmente, la función de utilidad de la alternativa B puede representarse como $U_B = \beta_1 P_B + \beta_2 F_B + \varepsilon_B$. La probabilidad de que n escoja la alternativa A sobre la B puede ser calculada de la siguiente forma:

$$P_A = \frac{e^{\beta_1 P_A + \beta_2 F_A}}{e^{\beta_1 P_A + \beta_2 F_A} + e^{\beta_1 P_B + \beta_2 F_B}} \quad (2.9)$$

que puede resumirse en $P_A = 1/(1 + \exp(V_B - V_A))$. De forma genérica, el resultado 2.9 puede adaptarse a más de alternativas utilizando la fórmula del logit multinomial: $P_1 = e^{V_1} / \sum_{j=1}^J e^{V_j}$.

De forma inversa, se pueden obtener datos de elección de muchos encuestados y estimar estos parámetros a través de una *regresión logística condicional* para conocer la utilidad que cada característica aporta a los individuos. Para ello, hay que establecer explícitamente la conexión entre la aplicación empírica de un DCE y la teoría desarrollada durante esta sección: el *logit condicional* de McFadden (1974). En un DCE contamos con datos de panel, es decir, un número N de individuos responde a las mismas preguntas (*choice sets*). Supongamos que existe un

número K de *choice sets* designados por B_k (B_1, \dots, K). Igualmente, cada *choice set* contiene J_k alternativas indexadas como $j = 1, \dots, J_k$ con sus características x_{jk} . Representemos una función de utilidad media de todos los individuos al escoger la alternativa i en el *choice set* B_k :

$$V_{ik} = \beta' x_{ik}. \quad (2.10)$$

Podemos generalizar desde la Ecuación 2.8 y obtener

$$Prob_{ik} = Prob(x_{ik}|B_k) = \frac{e^{\beta' x_{ik}}}{\sum_{j=1}^{J_k} e^{\beta' x_{jk}}}, \quad (2.11)$$

donde i representa a una de las J_k alternativas (todas las alternativas dentro del *choice set* B_k) y j es el subíndice genérico. El estimador de máxima verosimilitud (MLE) del *logit condicional* maximiza las probabilidades de *obtener los datos observados* por cada *choice set* del que se observa una respuesta (es decir, se realiza la estimación condicionada al choice set). En un DCE, varios individuos responden varios *choice sets*, pero el *logit condicional* no tiene en cuenta la variable identificadora del individuo y su MLE trata cada *choice set* como único. En la Tabla 2.1 (un DCE con dos atributos, '*carac*' y '*precio*', con tres niveles cada uno), la única variable de agrupación que toma el *logit condicional* es la variable *qid* (*choice set*), la cual identifica de forma única a todos los *choice sets* que responden los encuestados (cuya id es diferente, aunque el *choice set* sea igual al que ya respondió otro encuestado).

Hasta ahora, por simplicidad, las variables sociodemográficas han sido ignoradas. En este caso, sería más correcto llamar al modelo presentado anteriormente *logit condicional alternativa-específico*, pues solo incluye variables explicativas de los atributos de las alternativas. No obstante, además de estas variables, McFadden (1974) ya incluye en el componente determinista de la utilidad las variables *caso-específicas*, s_n , que son aquellas variables que describen al individuo. Estas variables se mantienen constantes para el mismo individuo, pero cambiantes entre encuestas y choice sets (respondidos por diferentes personas encuestadas).

$$U_{nj} = v(x_j, s_n)\beta + \varepsilon_{nj} \quad (2.12)$$

$$P_{ni} = \frac{e^{v(x_i, s_n)\beta}}{\sum_{j=1}^J e^{v(x_j, s_n)\beta}} \quad (2.13)$$

En muchas ocasiones, en la literatura de DCEs se habla del *logit multinomial* como sinónimo del logit condicional. No obstante, pese a que ambos modelos podrían parecer lo mismo, su enfoque es distinto y sirven a diferentes propósitos. El logit multinomial explica elecciones multinomiales (por ejemplo, marcas de coches) en función de las características de la persona decisora. No obstante, en el *logit condicional* se explica la posibilidad de elegir un tipo concreto de coche en

función de las características del coche. El enfoque principal del *logit condicional* es el objeto evaluado, mientras que el enfoque del *logit multinomial* es el sujeto evaluador (Hoffman & Duncan, 1988). Es por eso que, mientras sea razonable pensar que las elecciones se hacen en función de las características del producto, motivo por el cual realizamos los DCEs, usaremos el *logit condicional* para su análisis. Además, el *logit condicional* puede también incorporar variables sociodemográficas relevantes (Hoffman & Duncan, 1988).

pid	cs	qid	elec	carac_1	carac_2	precio
1	1	1	1	0	0	10
1	1	1	0	1	0	15
1	2	2	0	1	0	10
1	2	2	1	0	1	20
2	1	3	0	0	0	10
2	1	3	1	1	0	15
2	2	4	1	1	0	10
2	2	4	0	0	1	20

pid: identificador de encuestado/a, cs: identificador de choice set, qid: identificador único de choice set (combinación entre pid y cs), elec: variable binaria que toma el valor 1 cuando se elige la alternativa representada en su fila y 0 cuando no, carac_: característica del producto expresada en forma de variable dummy (en este caso, el atributo carac_ tendría tres niveles pero uno de ellos está omitido por colinealidad), precio: la variable precio en un experimento de elección discreta puede estar expresada de forma dummy o continua. Para calcular la disposición a pagar tiene que estar expresada de forma continua, por lo que en este caso se expresa así.

Tabla 2.1 Ejemplo de base de datos en experimento de elección discreta

2.2.2 Modelos alternativos

Para que el *logit condicional* sea aplicable de forma teóricamente correcta a estudios del comportamiento del consumidor, los términos de error de las alternativas deben ser iid (independientes e idénticamente distribuidos), condición que será infringida (independencia) si los términos de error de dos alternativas están correlacionados (Train, 2009). Esta condición no es trivial pues asegura que las correlaciones que se están estimando sean reales, y no haya *segundas correlaciones* que, de ser tenidas en cuenta, modificarían los resultados substancialmente³. Por otro lado, la idéntica distribución también implica una homogeneidad entre personas encuestadas que no es realista. Así, existe la duda razonable de que los datos que obtenemos de un DCE infrinjan las condiciones iid. Como señala Train (2009), podemos convivir

³ Por ejemplo, supongamos una regresión lineal donde el salario de una persona se ajuste por su nivel de estudios con los datos de toda la población. Supongamos que estimamos la regresión y representamos, en un eje cartesiano, los errores (Y) y el salario (X). Supongamos que observamos que los errores se vuelven más grandes cuanto mayor es el salario. Esto implicaría que existen relaciones de causalidad importantes que no estamos teniendo en cuenta y que son determinantes a la hora de determinar el salario de los más ricos (por ejemplo, la riqueza familiar, el género, el lugar de nacimiento, la edad, etc.).

con esta duda y aceptar las limitaciones del modelo o podemos utilizar otras herramientas (o hacer ambas cosas).

Una alternativa es el *logit mixto* o *logit de parámetros aleatorios* (Mcfadden & Train, 2000). Para presentar el modelo, podemos seguir la explicación intuitiva de Hensher & Greene (2003). La heterogeneidad entre personas encuestadas podría ser abordada a través de la segmentación de la muestra, y el cálculo de un *logit* condicional sobre cada uno de los segmentos de esta muestra. No obstante, el problema es elegir adecuadamente los criterios de segmentación. Una aproximación diferente es asumir que existe variabilidad y aceptar que cada uno de los parámetros del vector β son aleatorios (con media y desviación típica) en lugar de fijos. Supongamos que $\beta \sim N(\mu, \sigma^2)$. En este caso, a través de métodos de simulación, para cada individuo se simula un número R de μ y σ^2 de los que se extraen R β s, y se calculan las probabilidades de elección para todas las β s resultantes. La media de la muestral de β s es tomada como la aproximación al parámetro fijo, aunque este proceso permite conocer si las diferencias entre individuos son o no significativas y cuál es el grado de heterogeneidad.

Por otro lado, existe otro modelo que podemos emplear, el cuál ahonda en la idea de la estratificación mencionada anteriormente. En el *logit de clases latentes*, se asume que las personas encuestadas están clasificadas en Q clases latentes (Greene & Hensher, 2003). De esta forma, a la fórmula de probabilidades *logit* debemos añadir un subíndice más que afectará a los parámetros estimados:

$$Prob_{intq} = \frac{e^{\beta'_q x_{int}}}{\sum_{j=1}^J e^{\beta'_q x_{jnt}}}, \quad (2.14)$$

donde j identifica a las alternativas (donde i es una de ellas), n identifica a la persona que realiza la decisión, t identifica al *choice set*, y q identifica a la clase latente a la que pertenece n . Como puede observarse, cada clase latente contará con preferencias diferentes y, por lo tanto, con parámetros diferentes. En las versiones implementadas en paquetes estadísticos como Stata (StataCorp LLC, 2017), se utiliza el algoritmo *expectation-maximization* (Train, 2008) el cual maximiza las probabilidades de observar las respuestas de los individuos usando un procedimiento similar al *logit condicional*, pero ponderando por la probabilidad de pertenecer a cada una de las clases. Por este motivo, el número de clases debe ser definido previamente. Usualmente se estiman Q modelos de clases latentes con $q = 1, \dots, Q$ y se escoge aquel modelo con mejores medidas de ajuste.

2.3 CONSTRUCCIÓN DEL EXPERIMENTO DE ELECCIÓN DISCRETA

Una vez repasados los fundamentos teóricos en la sección anterior, podemos proceder con la parte más pragmática de la construcción del DCE. No obstante, es imprescindible tener presentes los principios teóricos pues son la base de un buen diseño experimental.

2.3.1 Selección de atributos y niveles

Como vimos en McFadden (1974), la base teórica de los experimentos de elección discreta es la RUT, la cual postula que las elecciones de las personas están condicionadas por la utilidad que aportan las alternativas, dadas sus características y las características de la persona decisora. En un DCE, a las características de las alternativas las conocemos como atributos y las definimos con sus niveles.

Durante la sección anterior, se sembró la idea de que el error, ε_{nj} , de la función de utilidad del individuo n en la alternativa j debía comportarse como ruido blanco. Esto sólo sería posible si, entre otras cosas, se cumplen ciertas características a las que podemos llegar de forma lógica y que han sido recogidas anteriormente en la literatura. Supongamos que la expresión 2.15 representa a la utilidad que queremos estimar:

$$U_j = \beta' x_j + \varepsilon_j \quad (2.15)$$

donde U_j es la utilidad de un determinado conjunto de encuestados por una alternativa j específica, x_j es un vector que incluye, de forma binaria, información sobre los niveles de los atributos que contiene la alternativa j , β' es un vector que contiene cada uno de los coeficientes de las características del vector x_j , y ε_j es el error. Para que el error se comporte de la forma más parecida posible al ruido blanco, x_j debe abarcar a todas las características de la alternativa. Si x_j no abarca todas las características que la persona encuestada tiene en cuenta para tomar la decisión, el error contendrá un componente determinista (y no aleatorio) por lo que no será ruido blanco. Por esto motivo, es importante realizar una buena definición de atributos y niveles que abarque todas las características contempladas al realizar una elección. No obstante, si las alternativas definidas en el DCE contienen demasiados atributos y niveles, es posible que sea demasiado difícil realizar una elección entre alternativas y los encuestados se cansen demasiado rápido. Por este motivo, la mayoría de DCE existentes emplean de 3 a 7 atributos (Johnson et al., 2013).

A continuación, se presenta una lista de requisitos y recomendaciones para la selección de atributos y niveles extraída de la literatura:

- Replicabilidad: como norma general, la selección de atributos y niveles debe estar basada en un proceso de exploración que debe ser presentado como replicable en las mismas condiciones (Johnson et al., 2013).
- Relevancia: los atributos y niveles deben ser relevantes/importantes para la descripción del producto o servicio que se esté evaluando (Lancsar & Louviere, 2008). Para que un atributo o nivel sea relevante, este debe ser una característica generalmente considerada cuando se toman decisiones relacionadas con el producto o servicio evaluado (por ejemplo, número de habitaciones para una casa o efectividad para un tratamiento). Cuando no se incluyen atributos y niveles relevantes, dejamos a la “imaginación” del encuestado suponer cual es el atributo o nivel presente en su defecto.
- Comparabilidad: los atributos deben ser comparables entre sí (Coast et al., 2012). Esto quiere decir que, si por ejemplo estamos evaluando un tratamiento y dos de los atributos son la efectividad y el precio, el encuestado debe ser capaz de realizar intercambios entre precio y efectividad (por ejemplo, pagar un poco más para obtener mayor efectividad, o viceversa). En el momento en el que, por algún motivo, no es posible realizar este intercambio, debemos saber que se ha cometido un error de especificación de atributos y niveles.
- Concreción: los atributos no pueden preguntar directamente sobre la función de utilidad latente que tratamos de estimar (Coast et al., 2012). Es decir, no podemos usar atributos como “nivel de felicidad” o “satisfacción con el producto”, porque se parecerían peligrosamente a estas *puntuaciones latentes* que tratamos de estimar y su interpretación daría lugar a un problema lógico.
- No solapamiento: dos atributos no deben considerar las mismas características para no realizar una doble contabilidad (Johnson et al., 2013).
- No interdependencia: que un atributo pueda tomar cierto nivel no debe depender de otro atributo (pues esto también interferiría con la *comparabilidad*) (Johnson et al., 2013).

Una estrategia que aquí podemos proponer para realizar una selección de atributos y niveles adecuada es: 1) crear una larga lista de posibles atributos y niveles basada en información existente en la literatura publicada. 2) Reorganizar la lista cumpliendo los requisitos anteriormente mencionados a través del descarte de algunos elementos y la combinación de otros. 3) Validar la selección preliminar a través de, por ejemplo, consultas con expertos en los que se puede preguntar si sobra o falta algún atributo o nivel. 4) Realización de un experimento piloto. Finalmente, siempre es recomendable cumplir todos los requisitos propuestos en el checklist de

la Sociedad Internacional para la Farmacoeconomía e Investigación de los Resultados (ISPOR) (Bridges et al., 2011).

Una vez realizada la selección de atributos y niveles, es recomendable reportarlos de forma estándar. Para ello, en la Tabla 2.2 se presenta una selección de atributos y niveles ficticia con la que se tratará en las siguientes secciones.

Con respecto al atributo “Coste” que se ha incluido en la Tabla 2.2, cabe realizar varios comentarios. En primer lugar, el atributo coste es un atributo de fundamental importancia pues servirá para calcular la disposición a pagar por cada uno de los atributos del tratamiento. En el experimento el coste del tratamiento deberá ser presentado como un coste directo para el paciente, pues solo de esta forma se podrá asimilar la ganancia en salud esperada al valor monetario. Esto es posible pues el intercambio (salud-recursos económicos) es evaluado a nivel individual, es decir, es el mismo individuo el que “sacrifica” sus recursos económicos para obtener más salud (que es lo mismo que realiza un servicio de salud pública a nivel macro). Una vez se estime el modelo, podrá calcularse a cuánto dinero equivale una determinada mejora de salud para la población encuestada.

Atributos	Niveles	Definición
Efectividad	90%	El 90% de los pacientes tratados se curan
	70%	El 70% de los pacientes tratados se curan
	50%	El 50% de los pacientes tratados se curan
Reacciones adversas	No graves	Cualquier reacción adversa que no cumpla los anteriores criterios.
	Graves	Cualquier reacción adversa que sea mortal, amenazante para la vida, suponga ingreso, discapacidad, invalidez o malformación congénita.
Incidencia reacciones adversas	Frecuente	Afecta a más del 1% de los pacientes.
	Infrecuente	Afecta del 1% al 0.1% de los pacientes.
	Rara	Afecta a menos del 0.1%
Coste	1000 €	El paciente pagará 1000€ por el tratamiento
	2000 €	El paciente pagará 2000€ por el tratamiento
	3000 €	El paciente pagará 3000€ por el tratamiento
	4000 €	El paciente pagará 4000€ por el tratamiento

Tabla 2.2 Selección de atributos y niveles ficticia para un tratamiento

Supongamos, por ejemplo, que el hipotético tratamiento que se va a evaluar usando la selección de atributos y niveles de la Tabla 2.2 es un medicamento para una enfermedad no grave, que no supone una amenaza para la vida del paciente, pero que en determinadas ocasiones puede

suponer una reducción de la calidad de vida. Esta encuesta podría ser planteada a un grupo de expertos, a la población en general, o a un grupo de pacientes que padecen esta enfermedad (posicionándolos a quien no padezca la enfermedad en la situación de un paciente si la padezca) y revisar las estimaciones para saber cómo de valiosa estiman los encuestados que sería la ganancia en salud concreta. Puesto que en cada uno de los choice sets los encuestados han elegido según sus propios cálculos teniendo en cuenta la efectividad, las reacciones adversas y el coste del tratamiento, la precisión y fiabilidad de la estimación mejorará cuanto más pegados a la realidad sean los escenarios planteados. Esto, solo será posible si se cumplen los requisitos planteados anteriormente para realizar una buena selección de atributos y niveles.

2.3.2 Generación de un diseño experimental: teoría

En la Tabla 2.2 presentamos una selección de atributos y niveles para un diseño experimental $3^2 \cdot 2 \cdot 4$ (dos atributos de tres niveles, un atributo de dos niveles, y un atributo de cuatro niveles). Esta especificación daría lugar a $3^2 \cdot 2 \cdot 4 = 48$ combinaciones de potenciales alternativas. Si estas 48 alternativas fuesen combinadas entre sí en pares de todas las formas posibles para crear choice sets, tendríamos $48 \cdot 47 / 2 = 1128$ potenciales choice sets. Si todos los potenciales choice sets fueran empleados en el DCE, estaríamos ante un diseño factorial completo. Existen varias formas de emplear los diseños factoriales completos. Por ejemplo, los 1128 choice sets podrían dividirse en 94 bloques que contendrían 12 choice sets cada uno. Después, cada uno de estos bloques podría ser contestado por un grupo de encuestados. Otra alternativa sería repartir, por ejemplo, 16 choice sets aleatorios a cada encuestado, hasta que se respondan todos los choice sets del DCE factorial completo. No obstante, normalmente se busca usar un número reducido de versiones de la encuesta (por ejemplo, cuando la encuesta se realiza con papel y lápiz) o se quieren usar sólo aquellos choice sets que aporten más información. La alternativa al diseño factorial completo es el diseño fraccional, es decir, escoger solo alguno de los potenciales choice sets para el diseño. No obstante, la pregunta es ¿cómo decidir cuáles son los choice sets que deben ser empleados en la encuesta?

a) Diseños ortogonales

En el pasado, esta pregunta fue respondida usando teoría de diseños experimentales eficientes desarrollada para modelos de regresión lineal (Rose & Bliemer, 2014). Estas técnicas de construcción están basadas en composiciones ortogonales, las cuales permiten la estimación de todos los efectos principales (efectos de los atributos y sus niveles) con independencia del resto, asumiendo que no existen interacciones relevantes (Addelman, 1962). Estos diseños fueron (y son) muy usados pues evitan que exista colinealidad entre los niveles de los atributos y permiten

la estimación de todos los coeficientes del modelo. Cuando se usan estos diseños experimentales, el procedimiento empleado es el siguiente. En primer lugar, se realiza una selección de atributos y niveles; esta vez usaremos los siguientes tres atributos:

1. Color: rojo (0), verde (1), amarillo (2)
2. Tamaño: grande (0), mediano (1), pequeño (2)
3. Material: metal (0), madera (1), plástico (2)

Color	Tamaño	Material	Set
2	2	0	1
0	1	2	2
2	0	2	3
0	2	1	4
1	2	2	5
2	1	1	6
1	1	0	7
1	0	1	8
0	0	0	9

Tabla 2.3 Diseño ortogonal

Posteriormente, hay que buscar un diseño ortogonal. Esto puede hacerse empleando diseños ortogonales que pueden encontrarse en internet o en software específico. En este caso, emplearemos la funcionalidad “Orthoplan” que puede encontrarse en SPSS (IBM Corp., 2020). Una vez especifiquemos cuantos factores (atributos) y valores (niveles) tiene nuestro diseño, obtendremos el resultado que presentamos en la Tabla 2.3.

Alternativa 1			Alternativa 2			
Color	Tamaño	Material	Color	Tamaño	Material	Tarjeta
2	2	0	0	1	2	1
0	1	2	1	1	0	2
2	0	2	1	0	1	3
0	2	1	2	1	1	4
1	2	2	1	2	2	5
2	1	1	0	2	1	6
1	1	0	0	0	0	7
1	0	1	2	0	2	8
0	0	0	2	2	0	9

Tabla 2.4 Diseño ortogonal 1 (Alternativa 1) + Diseño ortogonal 2 (Alternativa 2)

Cada una de las nueve “tarjetas” de la Tabla 2.3 será un perfil o alternativa de nuestro experimento de elección discreta. Por ejemplo, la primera “tarjeta” hace referencia a una alternativa de color amarillo, de tamaño pequeño y de material metálico.

Para crear un choice set con cada una de las tarjetas, hay varias estrategias. Siguiendo uno de los ejemplos presentados en Street et al. (2005) podemos optar por, por ejemplo, generar un segundo diseño ortogonal y emparejar las tarjetas de los diferentes diseños de dos en dos (Tabla 2.4).

Una alternativa diferente es generar un diseño ortogonal con el doble de atributos con los mismos niveles, donde la primera mitad de atributos designaría a la primera alternativa y la segunda mitad de atributos a la segunda alternativa. En la práctica, esto equivaldría a generar en Orthoplan (IBM Corp., 2020) un diseño ortogonal con seis atributos de tres niveles cada uno. Los resultados de esta estrategia pueden observarse en la Tabla 2.5.

Alternativa 1			Alternativa 2			Tarjeta
Color	Tamaño	Material	Color	Tamaño	Material	
1	1	0	1	1	2	1
0	2	2	1	1	1	2
2	0	0	1	2	2	3
2	0	2	1	0	1	4
1	0	1	2	2	1	5
0	1	0	2	2	1	6
0	0	1	0	1	2	7
1	0	2	2	1	0	8
0	1	2	2	0	2	9
2	2	1	2	0	2	10
1	2	0	0	0	1	11
1	1	1	1	0	0	12
2	1	2	0	2	0	13
2	2	0	2	1	0	14
0	0	0	0	0	0	15
2	1	1	0	1	1	16
0	2	1	1	2	0	17
1	2	2	0	2	2	18

Tabla 2.5 Diseño ortogonal 1 (Alternativa 1) + Diseño ortogonal 2 (Alternativa 2)

El verdadero porqué de la ortogonalidad está basado en los siguientes fundamentos teóricos. En un modelo de regresión lineal, la matriz de varianzas covarianzas (VC) puede ser estimada siguiendo la Expresión 2.16.

$$VC = \sigma^2(X'X)^{-1} \quad (2.16)$$

Puesto que X representa a la matriz de diseño del modelo (equivalente al resultado en la Tabla 2.4 con una codificación específica que se detallará más adelante), generalmente los elementos de la matriz VC se minimizarán cuando X es ortogonal. De esta forma obtenemos un diseño que no nos dará problemas de multicolinealidad y que reducirá las varianzas de las estimaciones (Rose & Bliemer, 2014). No obstante, esto sólo es cierto para modelos lineales, y para analizar DCE empleamos modelos no lineales.

b) El abandono de la ortogonalidad

En cuanto se empezó a asumir que la función objetivo a minimizar no era la representada por la Expresión 2.16, se realizaron nuevas propuestas metodológicas, todas alrededor de la misma idea. Tras identificar la matriz de varianzas-covarianzas asintótica (AVC) del modelo logit condicional, el objetivo de muchos artículos se centró en minimizar los elementos de esta matriz. Entre las diferentes aproximaciones Bunch et al., (1996) popularizó el empleo del D-error (Expresión 2.17).

$$D - error = \det(\Omega_1)^{-\frac{1}{k}} \quad (2.17)$$

La minimización del determinante de la matriz AVC asumiendo un solo encuestado, Ω_1 , (y normalizando por el número de parámetros) implica la minimización de, grosso modo, los errores de las estimaciones que queremos realizar. Por tanto, el diseño experimental será más eficiente cuanto menor sea el D-error. Sin embargo, Ω_N solo puede ser determinada calculando la inversa de la matriz de información de Fisher, la cual es, a su vez, la segunda derivada negativa del logaritmo de verosimilitud del modelo logit condicional dado un valor para los parámetros del modelo (Rose & Bliemer, 2014; Train, 2009). Como puede observarse, la dificultad no sólo estaba en el álgebra, el cual no es un problema siempre que haya alguien capaz de programarlo, sino en el requerimiento de conocer el valor de los parámetros a priori. Dicho de otro modo, para realizar una estimación precisa de los parámetros, necesitamos conocer el valor de los parámetros de antemano. Esta paradoja ha sido superada de diferentes formas. En los siguientes párrafos se presentan algunas de las aproximaciones que se han dado en la literatura a lo largo de los años.

En Bunch et al (1996), se empleó lo que conoceremos como *priors*, un conjunto de parámetros previos que tendrán un valor fijo que será usado para minimizar el D-error (algo así como lo que se piensa que serán los coeficientes cuando se estime el modelo). En este artículo se asume que los *priors* son, en algunos casos, todos iguales a cero o, en otros casos, diferentes a cero. En el artículo de Bunch et al. (1996), más que una investigación aplicada se realiza un ejercicio teórico por lo que no nos detendremos en discutir si los *priors* tienen un valor apropiado o no. Pese a que a estas alturas ya se conocía que los diseños ortogonales no eran apropiados para esta clase

de experimentos, en este artículo se minimiza el D-error dados unos priors, eligiendo entre diseños ortogonales. En Rose y Bliemer (2014) estos diseños son denominados diseños D-eficientes ortogonales localmente óptimos; D-eficientes porque tratan de minimizar el D-error pese a que el procedimiento de búsqueda del diseño óptimo no es exhaustivo; localmente óptimos porque el diseño es optimizado para unos priors fijos; ortogonales porque en la estrategia de búsqueda sólo se consideran diseños ortogonales.

Huber y Zwerina (1996), no obstante, sólo consideraron priors diferentes a cero y realizaron la búsqueda sin restringirla a diseños ortogonales. Puesto que los priors también son fijos, este tipo de diseños pueden ser denominados diseños óptimos locales.

c) Relajación de supuestos restrictivos

En los casos anteriores, se asume un valor previo (prior) conocido para cada parámetro que era usado para buscar el diseño experimental con menor D-error. En la práctica, este procedimiento en general consiste en usar algún algoritmo de búsqueda que calcule el D-error para cada diseño candidato y elija aquel con el más bajo posible. No obstante, en los casos anteriores una mala especificación de los priors (el uso de unos valores equivocados) supondrá una reducción de la eficiencia del modelo. Para mitigar este problema, Sándor y Wedel (2001) propusieron lo que llamaron diseños eficientes bayesianos. Los diseños eficientes bayesianos consisten en tratar a los priors como variables aleatorias con una media y una desviación típica conocidas. De esta forma, el diseño resultante, optimizado para el rango de las distribuciones de los priors, era menos eficiente que un diseño óptimo local pero más robusto a la mala especificación de los priors (Sándor & Wedel, 2001).

d) Otros métodos, algunas conclusiones, y variaciones sobre lo mismo

Hasta ahora, lo que se ha presentado trata de seguir un orden cronológico y comprensible. No obstante, los hallazgos narrados no se han dado de forma lineal y, como advierte Louviere et al. (2010), esta materia ha coexistido junto a un gran grado de confusión entre conceptos y supuestos que no han sido aclarados hasta finales de la década de los 2000. Gran parte de la formalización y organización de la materia se la debemos a dos profesores de Sídney, Michael Bliemer y John Rose, los cuales, aparte de realizar importantes contribuciones metodológicas, han escrito diversos capítulos de manuales desentrañando los errores y aciertos que la literatura sobre DCE ha ido cometiendo prácticamente desde que nació.

En Rose & Bliemer (2014) se desvelan algunos de los motivos de las contradicciones y confusiones que se han producido a lo largo de los años en esta materia. Por ejemplo, en Kuhfeld

(1994) se usa la matriz de VC del modelo lineal para optimizar según una medida llamada la D-eficiencia, la cual no es válida aplicada a modelos no lineales. Similarmente, Huber y Zwerina (Huber & Zwerina, 1996) hablan de que debe existir un equilibrio de utilidad entre las alternativas del choice set para mejorar la eficiencia del diseño, lo cual no es cierto y ha creado mucha confusión pues el artículo se popularizó mucho. Por otro lado, el artículo de 2005 de Street, Burgess y Louviere, “Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments” (Street et al., 2005) presenta ejemplos simples que, como reconocen, son incorrectos y no deben usarse, y después presentan una propuesta de diseños óptimos con priors cero que resulta muy difícil de seguir por lo breve de la explicación, quizás por las restricciones de espacio del artículo (en cualquier caso, esta implementación tampoco es fácil de seguir tras un curso en directo por uno de los autores). Por último, ha existido y aún existe una gran confusión alrededor de la terminología empleada para describir los procedimientos y supuestos de obtención de diseños: optimalidad, eficiencia, D-error, D-eficiencia, D_p -optimalidad, D_z -optimalidad. Aunque no profundizaremos mucho más de lo que ya hemos hecho, estos términos se refieren mayoritariamente a lo mismo, aunque bajo diferentes supuestos.

Contrariando el hallazgo de Zwerina et al., (1996), tanto en Kanninen (2018) como Rose & Bliemer (2004) demostraron finalmente que los diseños no ortogonales superaban a los ortogonales. Posteriormente, estos autores centraron su atención en optimizar el experimento al número de encuestados, Rose & Bliemer (2005), optimizar las varianzas para la estimación de las disposiciones a pagar, Rose et al. (2008), y propusieron métodos como los diseños seriales (o secuenciales) (Bliemer & Rose, 2010a). Esta última aproximación, a la cual se le prestará más atención en los siguientes capítulos, está basada en actualizar los priors y volver a generar un nuevo diseño óptimo cada vez que un encuestado registra una nueva respuesta.

En el siguiente apartado se presentarán formas prácticas de generar diseños experimentales usando algunos de los métodos introducidos.

2.2.4 Generación de un diseño experimental: práctica

Existen numerosas alternativas para la creación de DCE óptimos o eficientes. Entre ellas, merece la pena mencionar el programa Ngene (ChoiceMetrics, 2018) y las macros de SAS (SAS Institute, 2020). Sin embargo, en esta sección van a emplearse alternativas que pueden resultar más familiares como Stata (Statacorp, 2019) y R (R Cran, 2020), específicamente dos paquetes creados por usuarios. A continuación, se presentará el método de optimización que usan estos paquetes y una implementación propia en R.

En primer lugar, en Stata usaremos del paquete 'dcreate'. El paquete 'dcreate' fue creado por Arne Risa Hole en el año 2015 (Hole, 2017) y puede ser descargado desde el repositorio oficial. Aunque no existe un artículo oficial en *The Stata Journal*, desde 2015 el autor proporciona asistencia en *Statalist*: Home > Forums > Forums for Discussing Stata > General > dcreate: a new module for creating efficient designs for discrete choice experiments (Abril de 2014). Además, el autor, el cual es un especialista en la materia y ha publicado diversos paquetes relacionados, proporcionó un seminario en Noruega en 2016 cuyos materiales están disponibles. El paquete está basado en la reducción del D-error⁴ dados unos priors proporcionados por el usuario. El usuario podrá decidir el número de alternativas y choice sets del DCE. Para minimizar el D-error, el paquete usa una versión del algoritmo de Fedorov (Miller & Nguyen, 1994) a través del cual halla un diseño experimental, siendo potenciales candidatos todos aquellos diseños posibles, sin existir restricciones de ortogonalidad. Posteriormente, el autor actualizó el paquete para aceptar diseños bayesianos. En la sección 2.1 del Apéndice 2 puede encontrarse el código de Stata con el que crear el diseño experimental.

La segunda alternativa es usar el paquete 'idefix' en R (Traets et al., 2020), el cual permite generar diseños bayesianos D-eficientes. Además, el paquete contiene diversas herramientas para poner en práctica la encuesta generada. 'idefix' es un paquete más completo que 'dcreate', pero también más difícil de usar. El artículo citado anteriormente, publicado en *Journal of Statistical Software*, contiene toda la información relativa al uso del paquete. En la sección 2.2 del Apéndice 2 puede encontrarse el código de R con el que generar un diseño experimental.

Ambos paquetes pueden ser empleados tanto para generar diseños óptimos como para generar diseños bayesianos. Para comprender las diferencias de forma práctica, a continuación, se formulan los pasos para generar el diseño experimental tanto de forma óptima como bayesiana.

Ejemplo 1. Diseño óptimo (priors fijos)

Para crear un experimento de elección discreta basado en unos priors fijos necesitamos la siguiente información:

⁴ El autor habla indistintamente de aumentar la D-eficiencia o minimizar el D-error. Puesto que Rose y Bliemer (2014) atribuyen la medida de la D-eficiencia a Kuhfeld (1994), artículo que causó confusión por usar la matriz de varianzas-covarianzas del modelo lineal en lugar del modelo no lineal, se ha evitado usar esta nomenclatura en el texto principal. No obstante, a lo largo de la vasta literatura en esta materia, el concepto D-eficiencia es usado en numerosas ocasiones. En algunas ocasiones, como en el caso del paquete 'dcreate' y, en general, los artículos de A. R. Hole, la D-eficiencia se refiere a la inversa del D-error, calculado de forma correcta para un modelo no lineal. En otros casos es difícil saber si cuando se emplea el concepto D-eficiencia los autores se refieren a la D-eficiencia de Kuhfeld (1994) o a la D-eficiencia no lineal. En otros casos, se habla de la D-eficiencia como una medida relativa, donde el D-error del diseño es comparado con el D-error de un diseño que se conoce como el mejor.

- Los atributos y sus niveles
- El número de choice sets
- El número de alternativas
- Si existe o no una alternativa nula
- Los priors

Para este ejemplo usaremos una selección de atributos y niveles $2^2 \cdot 3$, ocho choice sets, dos alternativas por choice set, sin alternativa nula y priors $(1.654, 0.547, -0.53, -1.04)$. Nótese que pese a que el diseño experimental cuenta con un total de $2 + 2 + 3 = 7$ niveles, sólo hemos especificado un total de cuatro priors pues el modelo a calcular solo cuenta con cuatro parámetros. Esta diferencia entre el número de niveles y el número de parámetros del modelo se debe a lo que conocemos como ‘codificación dummy’. Para ilustrar este punto, supongamos que el número de atributos y niveles que hemos especificado se corresponden con la Tabla 2.6.

Atributos	Niveles	Prior
(1) Efectividad	50%	Restringido a cero
	90%	0.491
(2) Efectos secundarios	1% de los casos	Restringido a cero
	0.1% de los casos	0.547
(3) Precio	10€	Restringido a cero
	20€	-0.53
	30€	-1.04

Tabla 2.6 Atributos y niveles del Ejemplo 1

Como puede observarse, uno de los niveles de cada atributo está restringido a cero. De esta forma, el número de parámetros del modelo podrá ser calculado como $(l - k + o)$ donde l hace referencia al número total de niveles, k se refiere al número de atributos, y o toma el valor 1 si hay alternativa nula y 0 si no la hay. En nuestro caso el número de parámetros será $7 - 3 + 0 = 4$.

Ch. set	Alternativa	efect_50	efect_90	esec_1	esec_01	prec_10	prec_20	prec_30
1	1	1	0	0	1	0	1	0
1	2	0	1	1	0	1	0	0

Tabla 2.7 Ejemplo de choice set

Cuando se presenta una matriz de diseño, lo habitual no es hacerlo como en la Tabla 2.5. Generalmente las matrices de diseño son presentadas usando codificación ‘dummy’, es decir, cada uno de los posibles niveles es presentado como una variable binaria que toma el valor cero

o uno dependiendo de si ese nivel se encuentra o no presente en la alternativa. La Tabla 2.7 muestra un ejemplo de choice set con cada nivel especificado como variable dummy.

Ch. set	Alternativa	efect_90	esec_01	prec_20	prec_30
1	1	0	1	1	0
1	2	1	0	0	0

Tabla 2.8 Ejemplo de choice set (2)

Como puede observarse, todos los niveles han sido especificados de forma binaria. Por ejemplo, la primera alternativa tiene una efectividad del 50%, efectos secundarios en el 0.1% de la población y un precio de 20€. No obstante, sabiendo que una alternativa no puede tomar dos niveles del mismo atributo a la vez (por ejemplo, efectividad 50% y efectividad 90%) se pueden conocer todos los niveles de la alternativa si en la Tabla 2.7 eliminamos un nivel por cada atributo (Tabla 2.8).

Ch. set	Alternativa	efect_90	esec_01	prec_20	prec_30
1	1	1	0	0	0
1	2	0	0	0	0
2	1	1	1	1	0
2	2	0	0	0	0
3	1	0	1	0	1
3	2	0	0	0	0
4	1	1	1	0	1
4	2	1	0	0	0
5	1	0	1	0	1
5	2	0	1	0	0
6	1	0	0	0	1
6	2	1	0	1	0
7	1	1	1	0	1
7	2	1	0	1	0
8	1	0	1	0	1
8	2	0	0	0	1

Tabla 2.9 Diseño aleatorio

La Tabla 2.8 nos aporta exactamente la misma información que la Tabla 2.7. Por ejemplo, puesto que existen sólo dos niveles de efectividad, si conocemos que la primera alternativa toma valor 0 en la variable *efec_90* ya sabemos que esa alternativa tiene una efectividad del 90%. Este cambio, además de reducir la cantidad de datos con la que debemos de tratar, es necesario para que no exista multicolinealidad perfecta cuando se estime el modelo. Cada uno de los niveles restantes

en la Tabla 2.8 ($efec_{90}, esec_{01}, prec_{20}, prec_{30}$) se corresponden con cada uno de los priors que se ha especificado para generar el diseño óptimo.

A continuación, el objetivo es minimizar el D-error del diseño experimental. Para ello, el primer paso es generar un diseño aleatorio que cumpla con las características anteriormente propuestas.

A continuación, en la Tabla 2.9 se presenta un diseño aleatorio generado en R.

Para evaluar el D-error del diseño en la Tabla 2.9, empleamos la expresión 2.17,

$$D - error = |\Omega|^{-\frac{1}{k}}, \quad (2.17)$$

teniendo cuenta la expresión 2.18,

$$\Omega = \sum_{s=1}^S X_s^T \cdot (diag(p_s) - p_s p_s^T) \cdot X_s, \quad (2.18)$$

donde k hace referencia al número de parámetros, S es igual al número de choice sets totales, con $s = 1, 2, \dots, S$, y p_s es el vector de probabilidades de elección de las alternativas del choice set s . Para calcular los p_s empleamos la Expresión 2.19,

$$p_{si} = \frac{\exp(X_{si}\beta)}{\sum_{j=1}^J \exp(X_{sj}\beta)}. \quad (2.19)$$

Por ejemplo, $p_1 = (0.839, 0.161)$, por lo que cuando $s = 1$ la expresión 2.18 sería:

$$\begin{aligned} X_s^T \cdot \left(\begin{pmatrix} 0.839 & 0 \\ 0 & 0.161 \end{pmatrix} - (0.839 \quad 0.161) \begin{pmatrix} 0.839 \\ 0.161 \end{pmatrix} \right) \cdot X_s \\ = X_s^T \cdot \begin{pmatrix} 0.5915 & 0.1131 \\ 0.0216 & 0.0041 \end{pmatrix} \cdot X_s, \end{aligned}$$

donde X_s hace referencia al choice set s tal como se encuentra en la Tabla 2.9 (omitiendo la columna choice set y alternativa). Por tanto, el cálculo quedaría como:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0.5915 & 0.1131 \\ 0.0216 & 0.0041 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0.5915 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Una vez este procedimiento sea repetido para todos los choice sets y se realice la suma de las matrices resultantes, obtendríamos

$$\Omega = \begin{pmatrix} 0.3604 & 0.1332 & 0.2256 & -0.0924 \\ 0.1332 & 1.0862 & -0.1167 & 0.7207 \\ 0.2256 & -0.1167 & 0.4756 & -0.3433 \\ -0.0924 & 0.7207 & -0.3423 & 1.0061 \end{pmatrix},$$

y el D-error del diseño aleatorio inicial podría calcularse como

$$D - error = 0.04063^{-\frac{1}{4}} = 2.227.$$

Nótese que el D-error es una medida relativa y solo podrá ser usada para comparar diferentes configuraciones del mismo diseño experimental.

Ch. set	Alternativa	efect_90	esec_01	prec_20	prec_30
1	1	0	0	1	0
1	2	0	1	0	1
2	1	0	1	1	0
2	2	1	0	0	0
3	1	1	0	0	1
3	2	0	1	0	0
4	1	0	1	0	0
4	2	1	0	0	1
5	1	1	0	0	0
5	2	0	1	1	0
6	1	0	0	0	0
6	2	1	1	1	0
7	1	0	1	0	1
7	2	0	0	1	0
8	1	0	0	0	0
8	2	1	1	1	0

Tabla 2.10 Diseño óptimo del Ejemplo 1

A continuación, para reducir el D-error del diseño experimental, pondremos en práctica el algoritmo de Fedorov modificado (Zwerina et al., 1996), el cual es el procedimiento por defecto para encontrar un diseño óptimo. Para ejecutar este algoritmo es necesaria una matriz que contenga el diseño inicial aleatorio, que llamaremos D , y una matriz que contenga todas las alternativas posibles, que llamaremos A . La matriz A contendrá un total de $2 \cdot 2 \cdot 3 = 12$ alternativas.

El algoritmo de Fedorov modificado puede describirse informalmente de la siguiente forma. En la iteración 1, la primera alternativa de D es reemplazada por cada una de las alternativas en A y se evalúa el D-error del diseño en cada reemplazo. Después, aquella alternativa que ha obtenido el menor D-error es sustituida en D . El procedimiento continúa para todas las alternativas del diseño candidato. Una vez se hayan reemplazado todas las alternativas D , comienza la segunda iteración y se repite todo el proceso una vez más. El número de iteraciones puede ser definido por el usuario. Una buena estrategia es comparar el D-error tras cada iteración y detener el algoritmo una vez el D-error de la última iteración sólo mejore en un pequeño porcentaje al de la penúltima iteración.

Tras tres iteraciones del ejemplo anterior, el diseño resultante del procedimiento de optimización, con un D-error de 1.11, puede observarse en la Tabla 2.10. Una vez obtenido el diseño óptimo, este puede ser decodificado y presentado a los encuestados. En la Tabla 2.11 se presenta a modo de ejemplo el primer choice set.

	Alternativa A	Alternativa B
Efectividad	50%	50%
Efectos secundarios	1%	0.1%
Precio	20€	30€
(Escoger una alternativa)	<input type="checkbox"/>	<input type="checkbox"/>

Tabla 2.11. Primer choice sets del diseño óptimo del Ejemplo 1.

Ejemplo 2. Diseño eficiente bayesiano

Para generar un diseño eficiente bayesiano utilizaremos una distribución de priors en lugar de unos priors fijos. En el Ejemplo 1 se emplearon los priors (1.654, 0.547, -0.53, -1.04). Para ilustrar este ejemplo, usaremos los priors del ejemplo anterior como media de una distribución normal multivariante $N = 100$:

$$\begin{pmatrix} 0.9491 & 1.2931 & -2.5772 & -1.6849 \\ 2.6355 & 0.8143 & -2.2872 & -0.0216 \\ \dots & \dots & \dots & \dots \\ 2.5119 & 0.2831 & -1.1815 & -1.1292 \end{pmatrix}$$

A partir de este punto, el procedimiento es el mismo que en el Ejemplo 1, pero el D-error empleado es la media de todos los D-errores calculados para las cien diferentes versiones de los priors. Por ejemplo, supongamos que comenzamos por el diseño aleatorio en la Tabla 2.9. Comenzamos evaluando el D-error del diseño para el primer vector de priors de la distribución multivariante,

(1.654, 0.547, -0.53, -1.04), y repetimos este procedimiento para todos los elementos de la matriz multivariante. A continuación, se presentan los D-errores para todas las filas de priors especificadas en la matriz anterior.

$$P = \begin{pmatrix} 2.2142 \\ 2.8647 \\ \dots \\ 2.5391 \end{pmatrix}$$

La media del vector resultante,

$$\bar{P} = 3.0512,$$

puede usarse para calcular el DB-error como

$$DB - error = 3.0512^{\frac{1}{4}}.$$

Nótese que el DB-error requiere un tiempo de computación que es N veces el tiempo de computación del D-error, siendo N el número de vectores de priors generados (en el ejemplo anterior, 100). Puesto que el algoritmo de optimización necesita conocer el DB-error del diseño en numerosas ocasiones, este procedimiento puede ser muy lento si el cálculo del DB-error implica varios segundos de computación.

Tipo de D-error	Diseño	Sets	Algoritmo	Lenguaje de programación	Tiempo (segundos)
Determinista	2^2	4	Fedorov modificado	R	0.26
Determinista	2^2	4	CEA	R	0.49
Bayesiano	2^2	4	Fedorov modificado	R	17.44
Bayesiano	2^2	4	CEA	R	16.69
Bayesiano	2^2	4	Fedorov modificado	C++	0.12
Bayesiano	2^2	4	CEA	C++	0.27
Bayesiano	$2^2 \cdot 3$	8	Fedorov modificado	R	316.24
Bayesiano	$2^2 \cdot 3$	8	CEA	R	81.04
Bayesiano	$2^2 \cdot 3$	8	Fedorov modificado	C++	15.64
Bayesiano	$2^2 \cdot 3$	8	CEA	C++	2.03

Tabla 2.12. Tiempos de computación de los diseños según sus características.

En la Tabla 2.12 se muestran los tiempos de computación de diseños experimentales óptimos (o óptimos bayesianos) dependiendo de sus características. Como puede observarse, el planteamiento propuesto en la sección anterior, el algoritmo de Fedorov modificado, es el que más tarda, especialmente cuando se incrementa el número de atributos, niveles o choice sets en el modelo experimental. Para diseños medianos o grandes puede observarse como el algoritmo 'CEA' (coordinate exchange algorithm) es mucho más rápido. CEA realiza una búsqueda menos

exhaustiva que Fedorov pues va modificando cada atributo del experimento en lugar de cada alternativa, pero a cambio ofrece una velocidad razonable. No obstante, los tiempos de computación pueden reducirse sustancialmente aún más si usamos lenguajes de programación más rápidos (y directos) que R como, por ejemplo, C++. En el Apéndice 2 Sección 2.3 puede encontrarse un vínculo a un código de R para generar diseños experimentales que minimicen tanto el D-error como el D(B)-error⁵. Cuando se minimiza el D-error se emplea el algoritmo de Fedorov modificado y el cómputo se realiza en R. No obstante, cuando el diseño es bayesiano se emplea el algoritmo CEA y el D(B)-error es computado a través de una conexión entre C++ y R⁶.

2.4 REALIZACIÓN DEL EXPERIMENTO DE ELECCIÓN DISCRETA

Una vez que se conoce el diseño experimental el siguiente paso es llevar a cabo la encuesta. La encuesta puede ser realizada tanto de forma física, con papel y lápiz, o de forma telemática a través de una plataforma de encuesta en internet.

En ambas versiones, varios requisitos deben cumplirse. En primer lugar, la encuesta debe incluir un mínimo de tres secciones: introducción, experimento de elección discreta, y preguntas demográficas. La introducción es de gran importancia pues explica al encuestado qué va a responder y cómo ha de responderlo. La encuesta disponible en Fields et al. (2018) ofrece un buen ejemplo de texto introductorio.

Seguidamente, la persona encuestada debe responder al experimento de elección discreta. En esta sección, los choice sets del DCE deben aparecer en orden aleatorio a cada encuestado para evitar el sesgo de orden. Una medida de seguridad adicional es introducir una pregunta *trampa* con una alternativa dominante. Para ello, sólo tendremos que crear ‘artificialmente’ un choice set donde una de las alternativas sea mejor a la otra en todos los atributos y más barata. De esta forma, si la persona encuestada no escoge la alternativa dominante sabremos que no estaba prestando suficiente atención a la encuesta. Si es una encuesta de ‘papel y lápiz’, el orden aleatorio de las preguntas sólo podrá conseguirse imprimiendo diferentes versiones del cuestionario. Cuando la encuesta es online, normalmente existe una opción para presentar las encuestas en orden aleatorio.

⁵ Todo el código ha sido escrito para esta tesis usando R. El algoritmo de Fedorov fue programado con la asistencia de Arne Risa Hole, quien brindó ayuda a través de email describiendo el funcionamiento del algoritmo. La implementación de CEA se realizó en R, pero fue reemplazada por la del paquete ‘idefix’, la cual estaba programada en C++ y era mucho más rápida.

⁶ El paquete Rcpp permite ejecutar código de C++ pasando información desde R, por lo que ciertos procesos pueden ser acelerados.

Como hemos mencionado anteriormente, el diseño experimental implica decidir qué número de choice sets debe emplearse. Posteriormente daremos algunas claves para decidir qué número de choice sets puede usarse, pero de momento diremos que cuánto mayor sea el número de atributos y niveles empleados más choice sets habrá que usar para calcular todos los parámetros adecuadamente. Una ‘regla de tres’ muy sencilla, que podríamos proponer en esta tesis, podría ser usar doce choice sets en diseños pequeños (2-3 atributos), usar dieciséis choice sets en diseños medianos (4-6 atributos) y veinticuatro o más choice sets para diseños más grandes. Cuando el diseño experimental requiere dieciséis o veinticuatro choice sets, no podemos esperar que cada encuestado responda esta cantidad de choice sets. Teniendo en cuenta que el ‘efecto fatiga’ empieza a hacer efecto a partir de dieciocho choice sets (Mangham et al., 2009), en muchos casos deberemos reducir el número de choice sets que responde cada persona encuestada. Una estrategia común es separar la encuesta en varios bloques. Por ejemplo, supongamos que hemos generado un diseño experimental con dieciséis choice sets pero pensamos que son demasiados para que los responda un solo encuestado. Podemos distribuir aleatoriamente los choice sets entre dos encuestas generando así dos versiones con ocho sets cada una. Estas versiones deben repartirse aleatoriamente entre las personas que vayan a ser encuestadas.

Finalmente, en la sección de preguntas demográficas debemos preguntar por todo aquello que pueda ser significativo para explicar las elecciones de la persona encuestada. Por ejemplo, puesto que los DCE generalmente contienen un atributo precio, preguntar sobre la renta suele ser pertinente. Ya que algunas preguntas pueden hacer dudar al encuestado sobre si desea o no compartir esa información con el encuestador, es recomendable recordar, antes de comenzar esta sección, que la información está completamente anonimizada y que los encuestadores no sabrán (ni requerirán saber) quién ha respondido qué.

2.5 ALGUNOS DETALLES

En este capítulo se ha repasado el marco conceptual, teórico, técnicas de diseño de DCE, y métodos de encuestación. No obstante, es necesario concretar algunos detalles más para ofrecer toda la información necesaria para realizar el diseño experimental.

2.5.1 Número de choice sets

Un DCE requiere al menos el mismo número de choice sets que de parámetros para poder estimar independientemente cada uno de los efectos (Hair et al., 2019). Generalmente, los DCE emplean más choice sets de los estrictamente necesarios para reducir el tamaño de la muestra empleada.

$$n_{par} = l - k + o \quad (2.18)$$

El número de parámetros con los que cuenta un modelo puede calcularse simplemente empleando la expresión 2.18 donde l es el número total de niveles en todo el experimento, k es el número de parámetros y o es una variable binaria que toma el valor 1 si existe una alternativa nula o 0 si no existe.

2.5.2 Muestra requerida

Siempre han existido opiniones contrapuestas en la literatura sobre el tamaño muestral necesario, aunque el artículo de Bekker-Grobb et al. (2015) es el más reciente y el más empleado actualmente para conocer la muestra necesaria. El tamaño muestral necesario depende tanto de los atributos y niveles como del número de choice sets, alternativas, priors y matriz de diseño. La referencia citada anteriormente contiene un código de R que puede emplearse para calcular el tamaño muestral.

2.6 ANÁLISIS DEL EXPERIMENTO DE ELECCIÓN DISCRETA

Como se mencionó anteriormente en el marco teórico, los experimentos de elección discreta se analizan usando modelos de elección discreta. El modelo básico de análisis es la regresión logística condicional, cuyas estimaciones pueden complementarse con el uso de la regresión logística de parámetros aleatorios y la regresión logística de clases latentes. No obstante, antes del análisis debe realizarse un formateo y codificación de los datos.

2.6.1 Reformateo de los datos

Id encuestado	CS 1	CS 2	CS 3	CS 4	Edad	Género
1	A	B	A	A	43	Mujer
2	B	B	A	A	25	Hombre
3	B	A	B	A	74	Mujer
4	B	B	A	B	34	Hombre

Tabla 2.13 Ejemplo de resultados de un DCE. Formato 'ancho'.

Generalmente, los programas de encuestación online, como Google Forms o Qualtrics, permiten exportar los resultados de la encuesta en formatos ampliamente utilizados como Excel. Sin embargo, de forma nativa los datos no están formateados adecuadamente para ser analizados a través de los modelos anteriormente mencionados y deben ser convertidos o reformateados para poder estimar los modelos.

Id encuestado	CS	Alternativa	Elección	Edad	Género
1	1	A	1	43	Mujer
1	1	B	0	43	Mujer
1	2	A	0	43	Mujer
1	2	B	1	43	Mujer
1	3	A	1	43	Mujer
1	3	B	0	43	Mujer
1	4	A	1	43	Mujer
1	4	B	0	43	Mujer
2	1	A	0	25	Hombre
2	1	B	1	25	Hombre
2	2	A	0	25	Hombre
2	2	B	1	25	Hombre
2	3	A	1	25	Hombre
2	3	B	0	25	Hombre
2	4	A	1	25	Hombre
2	4	B	0	25	Hombre
3	1	A	0	74	Mujer
3	1	B	1	74	Mujer
3	2	A	1	74	Mujer
3	2	B	0	74	Mujer
3	3	A	0	74	Mujer
3	3	B	1	74	Mujer
3	4	A	1	74	Mujer
3	4	B	0	74	Mujer
4	1	A	0	34	Hombre
4	1	B	1	34	Hombre
4	2	A	0	34	Hombre
4	2	B	1	34	Hombre
4	3	A	1	34	Hombre
4	3	B	0	34	Hombre
4	4	A	0	34	Hombre
4	4	B	1	34	Hombre

Tabla 2.14 Ejemplo de resultados de un DCE. Formato 'largo.

En la Tabla 2.13 se puede observar una base de datos de un DCE con cuatro encuestados que respondieron a una encuesta con cuatro choice sets. Como puede observarse, las variables 'CS 1', 'CS 2', 'CS 3' y 'CS 4' hacen referencia a cada uno de los choice sets. El valor que toman estas variables es 'A' o 'B', representando a la alternativa escogida por el encuestado. El resto de variables indican algunas características demográficas de la muestra. La base de datos debe ser

transformada de formato ‘ancho’ a formato ‘largo’ para estimar los modelos de elección discreta. En la Tabla 2.14 puede encontrarse la transformación de la Tabla 2.13 a formato ‘largo’.

Id encuestado	CS	Alternativa	Elección	Efectividad	Administración	Precio	Edad	Género
1	1	A	1	0.8	Oral	30	43	Mujer
1	1	B	0	0.6	Inyección	60	43	Mujer
1	2	A	0	0.6	Oral	30	43	Mujer
1	2	B	1	0.8	Inyección	90	43	Mujer
...								
4	3	A	1	0.8	Inyección	60	34	Hombre
4	3	B	0	0.6	Oral	90	34	Hombre
4	4	A	0	0.6	Inyección	90	34	Hombre
4	4	B	1	0.8	Oral	30	34	Hombre

Tabla 2.15 Ejemplo de resultados de un DCE con atributos y niveles. Formato ‘largo’.

Mientras que en la Tabla 2.13 existía una fila por cada encuestado, en la Tabla 2.14 la unidad básica de análisis es la alternativa, por lo que existe una fila por cada alternativa. La variable elección se ha codificado de forma binaria e indica si la alternativa de la fila (A o B) ha sido seleccionada. Nótese que para un DCE de cuatro encuestados con cuatro choice sets de dos alternativas cada uno hay un total de $4 \cdot 4 \cdot 2 = 32$ filas. Nótese también que las características de los individuos (edad y género) permanecen invariantes para cada individuo, pese a que la alternativa varíe.

2.6.2 Codificación de las alternativas

Id enc.	CS	Alt.	Elección	<i>efec60</i> Efectividad = 60%	<i>efec80</i> Efectividad = 80%	<i>admOr</i> Administ = Oral	<i>admIn</i> Administ = Iny	Pr.	Edad	Género
1	1	A	1	0	1	1	0	30	43	Mujer
1	1	B	0	1	0	0	1	60	43	Mujer
1	2	A	0	1	0	1	0	30	43	Mujer
1	2	B	1	0	1	0	1	90	43	Mujer
...										
4	3	A	1	0	1	0	1	60	34	Hombre
4	3	B	0	1	0	1	0	90	34	Hombre
4	4	A	0	1	0	0	1	90	34	Hombre
4	4	B	1	0	1	1	0	30	34	Hombre

Tabla 2.16 Ejemplo de resultados de un DCE con atributos y niveles. Formato ‘largo’.

La Tabla 2.14 tiene el formato correcto, pero carece de información sobre las alternativas que se están eligiendo. Por ejemplo, se conoce que la primera encuestada prefiere la alternativa ‘A’ a la alternativa ‘B’ en el primer choice set, pero esta información no es útil si no se conocen las

diferencias entre 'A' y 'B'. Para ello, cada alternativa debe incluir información sobre sus atributos y niveles, pero estos deben ser introducidos codificados de una forma en concreto. Imaginemos que la encuesta contenía un DCE $2^2 \cdot 3$ (dos atributos de dos niveles y un atributo de tres niveles) donde los atributos eran efectividad (niveles: 60% y 80%), administración (niveles: oral e inyección) y precio (niveles: 30€, 60€ y 90€). Una forma de definir las alternativas en la base de datos sería la presentada en la Tabla 2.15.

En muchas ocasiones, en los DCE los niveles de los atributos están compuestos por variables categóricas, nominales o binarias de las que no podemos asumir continuidad. En otros casos, se emplean variables continuas, como la efectividad en este caso, que nos interesa tratar como categóricas. En general, todos los atributos son codificados como variables binarias, siendo la única excepción el atributo precio, el cual se mantiene como en la Tabla 2.15.

En la Tabla 2.16 los atributos efectividad y administración ha sido codificados de forma binaria. Como cada uno de ellos tenía dos niveles, por cada atributo se han creado dos variables que toman valor uno o cero dependiendo del nivel presente en la alternativa. Por cada atributo, podemos prescindir de uno de los niveles para evitar que exista multicolinealidad perfecta. De esta forma, la base de datos quedaría como en la Tabla 2.17.

Id encue.	CS	Alt.	Elección	<i>efec80</i> Efectividad = 80%	<i>admln</i> Administ = lny	Pr.	Edad	Género
1	1	A	1	1	0	30	43	Mujer
1	1	B	0	0	1	60	43	Mujer
1	2	A	0	0	0	30	43	Mujer
1	2	B	1	1	1	90	43	Mujer
...								
4	3	A	1	1	1	60	34	Hombre
4	3	B	0	0	0	90	34	Hombre
4	4	A	0	0	1	90	34	Hombre
4	4	B	1	1	0	30	34	Hombre

Tabla 2.17 Ejemplo de resultados de un DCE con atributos y niveles. Formato 'largo'.

Nótese que se ha eliminado el nivel '60%' del atributo efectividad y el nivel 'oral' de administración, aunque podría haberse eliminado el otro nivel sin alterar la estimación. Generalmente, de cada atributo se elimina aquel nivel que se considera que aporta menor utilidad al encuestado, los motivos se explican más adelante.

2.6.3 Variaciones y generalización de la codificación

La codificación empleada en la sección anterior es conocida como *dummy coding* o codificación binaria. Existen otras alternativas, entre ellas *effects coding* o codificación por efectos. En la Tabla 2.18 se muestran lado a lado la codificación binaria y por efectos para un atributo hipotético con tres niveles.

	Codificación binaria		Codificación efectos	
	Atrib. 1 Niv. 2	Atrib. 1 Niv. 3	Atrib. 1 Niv. 1	Atrib. 1 Niv. 2
Nivel 1	0	0	1	0
Nivel 2	1	0	0	1
Nivel 3	0	1	-1	-1

Tabla 2.18 Codificación binaria y por efectos.

Como puede observarse, en la codificación binaria se ha omitido el nivel uno mientras que en la codificación por efectos se ha omitido el nivel tres. Como se mencionaba anteriormente, no importa el nivel que se omita, pero existe cierta costumbre en la literatura de omitir uno u otro dependiendo de la literatura.

La codificación binaria es la más fácil de comprender y la que ofrece una interpretación más intuitiva de los coeficientes, por lo que en esta tesis siempre se presupondrá el uso de esta. No obstante, debe saberse que ambas codificaciones obtienen resultados equivalentes y el uso de una u otra no altera las conclusiones que puedan ser extraídas del modelo.

Id encue.	CS	Alt.	Elección	Nula	<i>efec80</i> Efectividad = 80%	<i>admin</i> Administ = Iny	Pr.	Edad	Género
1	1	A	1	0	1	0	30	43	Mujer
1	1	B	0	0	0	1	60	43	Mujer
1	1	C	0	1	0	0	0	43	Mujer
1	2	A	0	0	0	0	30	43	Mujer
1	2	B	1	0	1	1	90	43	Mujer
1	2	C	0	1	0	0	0	43	Mujer
...									
4	3	A	1	0	1	1	60	34	Hombre
4	3	B	0	0	0	0	90	34	Hombre
4	3	C	0	1	0	0	0	34	Hombre
4	4	A	0	0	0	1	90	34	Hombre
4	4	B	1	0	1	0	30	34	Hombre
4	4	C	0	1	0	0	0	34	Hombre

Tabla 2.19 Ejemplo de resultados de un DCE con atributos y niveles. Formato 'largo'.

Por otro lado, si se emplean choice sets con alternativa nula hay varios aspectos que deben tenerse en cuenta al codificar y reformatear. Añadiendo una alternativa nula a cada choice set del ejemplo anterior, la Tabla 2.17 pasaría a ser como la Tabla 2.19.

Ahora cada choice set contiene tres alternativas, 'A', 'B' y 'C', donde 'C' hace referencia a la alternativa nula. Nótese que se ha creado una nueva variable, 'Nula', que toma el valor uno en una alternativa nula y cero en el resto. El resto de niveles tomará el valor cero en la fila de la alternativa nula.

Id encue.	CS	Alt.	Agrup.	Elección	Nula	<i>efec80</i> Efectividad = 80%	<i>admin</i> Administ = Iny	Pr.	Edad	Género
1	1	A	1	1	0	1	0	30	43	Mujer
1	1	B	1	0	0	0	1	60	43	Mujer
1	1	C	1	0	1	0	0	0	43	Mujer
1	2	A	2	0	0	0	0	30	43	Mujer
1	2	B	2	1	0	1	1	90	43	Mujer
1	2	C	2	0	1	0	0	0	43	Mujer
...										
4	3	A	15	1	0	1	1	60	34	Hombre
4	3	B	15	0	0	0	0	90	34	Hombre
4	3	C	15	0	1	0	0	0	34	Hombre
4	4	A	16	0	0	0	1	90	34	Hombre
4	4	B	16	1	0	1	0	30	34	Hombre
4	4	C	16	0	1	0	0	0	34	Hombre

Tabla 2.20 Ejemplo de resultados de un DCE con atributos y niveles. Formato 'largo'.

El DCE que se ha empleado para ejemplificar los puntos anteriormente tratados evalúa las preferencias de unos hipotéticos pacientes sobre un hipotético tratamiento. En este DCE los pacientes tienen la opción de escoger entre dos tratamientos alternativos o no escoger ninguno. La opción de no escoger ningún tratamiento no puede ser considerada como otro tratamiento, es decir, tiene una naturaleza diferente y no se le puede asignar un nivel de efectividad ni un modo de administración. En este caso, el párrafo anterior es cierto y la codificación debe ser realizada como en la Tabla 2.19. Sin embargo, existen casos en los que la alternativa nula es un tratamiento que ya se le está administrando al paciente y sí tiene un nivel de efectividad, un modo de administración y un precio. En estos casos, la 'alternativa nula' debe codificarse como una alternativa más, preferentemente estando compuesta de los niveles de los atributos que se han omitido para evitar la colinealidad. En este caso, no haría falta una variable 'nula' que tome valor uno en el tratamiento considerado nulo. Un buen ejemplo de esta última metodología puede encontrarse en Gundlach (2018), donde se evalúan las preferencias de los ciudadanos por una

ciudad libre de tráfico rodado. En este DCE, todos los atributos incluyen un nivel (el primero) que representa a la situación actual. Estos primeros niveles representantes de la situación de *status quo* son los que componen la alternativa nula. A su vez, estos son los niveles omitidos en cada atributo por lo que la alternativa nula toma valor cero en todos los niveles presentes en la base de datos, pero no existe una variable ‘nula’ como en la Tabla 2.19.

A parte de las variables presentadas en la Tabla 2.19, los modelos de elección discreta requieren una variable de agrupación de choice sets. Esta nueva variable deberá indicar qué alternativas forman parte del mismo choice set, dando una numeración única a cada set siempre que haya sido respondido por un encuestado diferente (sin importar si es el mismo o no que han respondido otros encuestados). En este caso, la variable ‘CS’ de la Tabla 2.19 no cumple este propósito pues repite su numeración para cada encuestado diferente. Un ejemplo puede observarse en la variable ‘Agrup’ de la Tabla 2.20, la cual va incrementándose de tres en tres (número de alternativas) hasta el final de la muestra.

2.6.4 Estimando las preferencias de las personas encuestadas

En primer lugar, se analizarán los resultados a través de la regresión logística condicional. Los datos que se emplearán en este apartado son producto de una simulación de cien respuestas a un experimento de elección discreta con los atributos y niveles empleados en los ejemplos anteriores. Los datos pueden ser recreados a partir del código de R en el apartado 2.4 del Apéndice 2.

Atributos	Niveles	Coficiente	Error estándar	P-valor
Efectividad	60 %	Restringido a cero	-	-
	80 %	0.5662	0.0745	<0.000
Administración	Oral	Restringido a cero	-	-
	Inyección	-0.2631	0.0744	<0.000
Precio	30 €	Restringido a cero	-	-
	60 €	-0.1398	0.0939	0.1367
	90 €	-0.1957	0.1165	0.0929

McFadden Pseudo $R^2 = 0.0676$
Logaritmo de verosimilitud = -517

Tabla 2.21 Resultados de la regresión logística condicional

En la Tabla 2.21 podemos observar los resultados de la regresión logística condicional. Aunque el software estadístico con el que se realizan las estimaciones no incluya las variables omitidas, se han añadido para mejorar la interpretación y comprensión del lector. Como puede observarse, los coeficientes de todas las variables omitidas han sido restringidos a cero. Por otro lado, el resto de coeficientes indican una puntuación que conoceremos como utilidad. Por ejemplo, el nivel 80%

de efectividad otorga 0.57 puntos de utilidad. No obstante, esta puntuación por sí misma no aporta información. La efectividad 80% aporta 0.57 puntos de utilidad sobre el 60% de efectividad, es decir, esa utilidad se atribuye a un aumento del 20% en la efectividad del tratamiento. Esto es así porque la utilidad que otorgan los niveles debe interpretarse de forma relativa al nivel omitido de cada atributo. Otro ejemplo sería el nivel 'Inyección' del atributo 'Administración'. Respecto al nivel base (oral) el modo de administración inyección supone una pérdida de utilidad para el paciente que podemos valorar en 0.26 puntos de utilidad. La pregunta que nos surge ahora es, ¿estaría el paciente dispuesto a aumentar la efectividad aunque el modo de administración sea menos conveniente? En este caso, la respuesta es claramente sí pues gana más utilidad (0.57) de la que pierde (-0.26).

Con respecto a las medidas de ajuste, en los modelos de elección discreta principalmente observamos tanto el logaritmo de verosimilitud como el pseudo R^2 de McFadden. El logaritmo de verosimilitud puede ser usado para comparar los ajustes de diferentes modelos similares. El R^2 de McFadden indica un buen ajuste siempre que se encuentre entre 0.2 y 0.4 (McFadden, 1977). En este caso, el R^2 es muy bajo debido a la procedencia y cuantía de los datos.

Por otro lado, también puede observarse como los niveles del atributo precio restan más utilidad al individuo cuanto mayor es el precio. No obstante, como se planteó anteriormente, es común codificar el atributo precio como una variable continua. En la Tabla 2.22 pueden observarse los resultados con esta codificación.

Atributos	Niveles	Coefficiente	Error estándar	P-valor
Efectividad	60 %	Restringido a cero	-	-
	80 %	0.5662	0.0745	<0.000
Administración	Oral	Restringido a cero	-	-
	Inyección	-0.2631	0.0744	<0.000
Precio	(Vble continua)	-0.0035	0.0019	0.069

McFadden Pseudo $R^2 = 0.0674$
Logaritmo de verosimilitud = -517

Tabla 2.22 Resultados de la regresión logística condicional

En esta ocasión, el atributo 'Precio' codificado como variable continua ofrece un coeficiente de -0.0035 que representa la utilidad perdida por cada euro gastado. Con esta información, la disposición a pagar por, por ejemplo, una efectividad del 80%, podría estimarse como la división negativa entre este coeficiente y el coeficiente del atributo precio (Expresión 2.19).

$$DAP = - \frac{\text{coef. niv. interés}}{\text{coef. atrib. precio}} = - \frac{0.5662}{-0.0035} = 161.77 \quad (2.19)$$

Por un aumento del 20% de efectividad del tratamiento, de media, los pacientes estarían dispuestos a pagar un máximo de 161.77€. Entendido con una lógica diferente, la utilidad que se gana al aumentar un 20% la efectividad y la utilidad que se pierde al gastar 161.77 € es la misma.

Finalmente, pueden estudiarse las variaciones de las probabilidades de elección usando la fórmula de probabilidades logit. A continuación, en la expresión 2.20 se muestra la variación de las probabilidades de elección cuando el modo de administración pasa de ser inyección a oral. Un cambio en los niveles de este atributo supone un 13% más de probabilidades de elegir la alternativa oral *ceteris paribus*.

$$\begin{aligned}
 Prob_i - Prob_o &= \frac{e^{coef_i}}{e^{coef_i} + e^{coef_o}} - \frac{e^{coef_o}}{e^{coef_i} + e^{coef_o}} = 0.4346 - 0.5654 \\
 &= -0.1308
 \end{aligned}
 \tag{2.20}$$

El siguiente modelo, ampliamente empleado en los DCE, es la regresión logística de parámetros aleatorios. Los resultados de esta regresión pueden observarse en la Tabla 2.23. En este modelo, los coeficientes deben ser interpretados de la misma forma que en el logit condicional y pueden realizarse los mismos cálculos y comparaciones. No obstante, este modelo también ofrece información sobre las desviaciones típicas de los parámetros (SD). Aquellas cuyo p-valor indica que son significativamente diferentes de cero nos aportan información sobre el nivel de *consenso* de los encuestados respecto al signo del coeficiente. Es decir, si la SD es significativa, existe un porcentaje significativo de encuestados que opinan lo contrario al resto. Este porcentaje puede ser estimado como $100 \cdot \phi\left(-\frac{\beta_k}{SD_k}\right)$ (Hole et al., 2016). Por ejemplo, en este caso, el 5.2% de los encuestados prefiere un nivel de efectividad menor.

Atributos	Niveles	Coficiente	Error estándar	P-valor
Efectividad	60 %	Restringido a cero	-	-
	80 %	0.5662	0.0745	<0.000
Administración	Oral	Restringido a cero	-	-
	Inyección	-0.2631	0.0744	<0.000
Precio	(Vble continua)	-0.0035	0.0019	0.069
(SD) Efectividad	60%	Restringido a cero	-	-
	80 %	0.6266	0.0826	0.0446
(SD) Administración	Oral	Restringido a cero	-	-
	Inyección	-0.1394	0.0785	0.0756
(SD) Precio	(Vble continua)	-0.0099	0.0020	<0.000

Logaritmo de verosimilitud = -506.18

Tabla 2.23 Resultados de la regresión logística de parámetros aleatorios

Estas diferencias son generalmente debidas a las características sociodemográficas de los encuestados. Una forma de incluir estas variables es a través del logit de clases latentes. Como

se explicó anteriormente, el logit de clases latentes divide a la muestra en un número M de subgrupos y calcula logits condicionales para cada uno de ellos. M debe ser especificado por los analistas, pero es el mismo modelo el que clasifica a cada uno de los encuestados en cada grupo dependiendo de sus respuestas y características. En la práctica, para hallar cuál es el M óptimo, se puede programar un bucle que pruebe todos los M factibles y que muestre cuál es aquel con mejores medidas de ajuste. Los resultados del bucle, el cual se puede revisar en el código proporcionado en 2.4, se presentan en la Tabla 2.24.

Clases	Logaritmo verosimilitud	AIC	BIC	
2	-498.6710	1011.342	1044.134	
3	-496.8096	1015.619	1067.150	
4	-502.5974	1035.195	1105.464	

Tabla 2.24 Medidas de ajuste de la regresión logística de clases latentes para $M = 2, 3, 4$

El modelo de dos clases es el que mejor medidas de ajuste muestra, por lo que estimaremos, de nuevo, la regresión para dos clases latentes. Los resultados se muestran en la Tabla 2.25. Como puede observarse, en la tabla se recopilan los coeficientes de los niveles de los atributos para las dos clases y las covariables de características sociodemográficas.

	Clase 1		Clase 2	
	Coefficiente	SE (p-valor)	Coefficiente	SE (p-valor)
Efectividad	Restringido a cero	-	Restringido a cero	-
	5.4686	301.47 (0.986)	0.0854	1.23 (0.945)
Administración	Restringido a cero	-	Restringido a cero	-
	5.4686	301.41 (0.986)	-1.6464	1.62 (0.309)
Precio	0.0069	0.0303 (0.819)	-0.0296	0.0127 (0.020)
Edad	-0.0076	0.0070 (0.278)	Restringido a cero	-
Género	0.0914	0.3948 (0.817)	Restringido a cero	-
Constante	-0.0597	0.4740 (0.900)	Restringido a cero	-
LL = -510.30				

Tabla 2.25 Regresión logística de clases latentes (2 Clases)

Los coeficientes de los niveles pueden ser interpretados como en el logit condicional, mientras que las covariables de las características sociodemográficas deben ser interpretadas con referencia a la clase de referencia. Por ejemplo, en este caso la Clase 2 es la clase de referencia. Si la edad hubiera sido significativamente negativa, significaría que las personas más jóvenes tienen mayores probabilidades de pertenecer a la Clase 1. Al igual que en el resto de logits, en este caso se puede estimar la disposición a pagar y los cambios de probabilidades para cada clase.

2.7 CONCLUSIONES

En esta sección se han presentado los fundamentos teóricos y aplicaciones prácticas de los experimentos de elección discreta. Se han cubierto, de forma más o menos precisa, la mayoría de conceptos necesarios para llevar a cabo un DCE, pero no todos los existentes. En el futuro, van a publicarse nuevas investigaciones muy esclarecedoras sobre el sesgo hipotético en los DCEs (Haghani et al., 2021a, 2021b), pero habrá que esperar hasta que estos artículos sean revisados por pares para trabajar sobre sus conclusiones. Por otro lado, en los últimos años, diversos autores han planteado numerosas alternativas tanto en modelos de análisis (véase, por ejemplo, el modelo nested logit) como en técnicas de optimización del diseño y encuestación. Algunas de estas novedades serán cubiertas en los capítulos siguientes. El Capítulo 3 presenta una aplicación real de todos los conceptos estudiados anteriormente. El Capítulo 4 propone y pone a prueba una innovación metodológica. El Capítulo 5 no incluye innovaciones metodológicas, pero si propone un novedoso flujo de trabajo en la construcción de DCE.

Capítulo 3. Una aplicación de los Experimentos de Elección Discreta



3.1 INTRODUCCIÓN

En esta sección se presenta una investigación en la que, a través de un experimento de elección discreta, se estimaron las preferencias de los usuarios por un sistema de nutrición personalizada. Esta investigación surge en colaboración con el proyecto Stance4Health, proyecto financiado por el programa Horizonte 2020 de la Unión Europea⁷ y coordinado por la Universidad de Granada. El principal objetivo de Stance4Health es hallar en qué medida unas recomendaciones nutricionales confeccionadas para el usuario, dependiendo de parámetros como su estado de salud y composición de su microbiota, tendrán un efecto significativo en su salud. Por este motivo, en el proyecto se desarrolla una aplicación móvil con la capacidad de ofrecer recomendaciones nutricionales personalizadas para mejorar la actividad de la microbiota intestinal de cada usuario.

A parte de los objetivos finales, uno de los objetivos intermedios del proyecto Stance4Health es investigar cómo los patrones conductuales, las preferencias y las condiciones socioeconómicas influyen sobre los hábitos de alimentación y la disposición del público a emplear las recomendaciones nutricionales personalizadas. Puesto que esta tarea consiste en el cálculo de unas preferencias por un producto sobre el que no existe información de mercado⁸, y del que se puede esperar un beneficio en salud futuro, esta colaboración representaba una oportunidad perfecta para poner en práctica la teoría desarrollada durante los capítulos anteriores. De esta forma, el objetivo de nuestra investigación fue cuantificar las preferencias de los potenciales consumidores del sistema de nutrición personalizada en una fase temprana. El estudio cuantitativo de estas preferencias permite, por un lado, otorgar a los decisores públicos información *a priori* para decidir cómo destinar los recursos públicos en prevención, puesto que para intervenciones nutricionales de este tipo la adherencia (voluntad del usuario a cumplir con las recomendaciones nutricionales) desempeña un papel fundamental. En este sentido, los resultados del ensayo clínico de Stance4Health revelarán los beneficios de la adopción del sistema, pero estos beneficios solo serán generalizables al resto de la población si los potenciales usuarios esperan un beneficio en salud tangible. Por este motivo, las preferencias y expectativas de los usuarios determinarán las posibilidades de éxito del sistema de nutrición personalizada. Por otro lado, nuestra investigación también otorga información para aquellos actores que tengan interés en proveer el servicio desde el sector privado.

⁷ Proyecto Smart Technologies for personalised Nutrition and Consumer Engagement (Grant agreement ID: 816303).

⁸ Pese a que existen alternativas similares en el mercado, el proyecto Stance4Health representa una aproximación científica y basada en la evidencia que considera características relevantes del individuo para ofrecer una recomendación ajustada a sus necesidades. En este sentido, podría considerarse que no existen en el mercado alternativas equivalentes de las que obtener información sobre las preferencias del usuario.

A continuación, se presentan la metodología, resultados y conclusiones de la investigación de forma detallada. La versión *preprint* del artículo puede consultarse en el siguiente enlace (<https://digibug.ugr.es/handle/10481/72306>). El artículo original se encuentra publicado en Appl. Health Econ. Health Policy⁹, y fue fruto del trabajo conjunto con los profesores David Epstein y José Alberto Castañeda García.

3.2 MÉTODOS

La construcción DCE, la recogida de respuestas de los usuarios y el análisis de los datos se llevó a cabo en los pasos que se detallan a continuación. Esta investigación se desarrolló desde finales de 2019 hasta mediados de 2020 y fue finalmente publicada a principios de 2021.

3.2.1 Revisión literaria

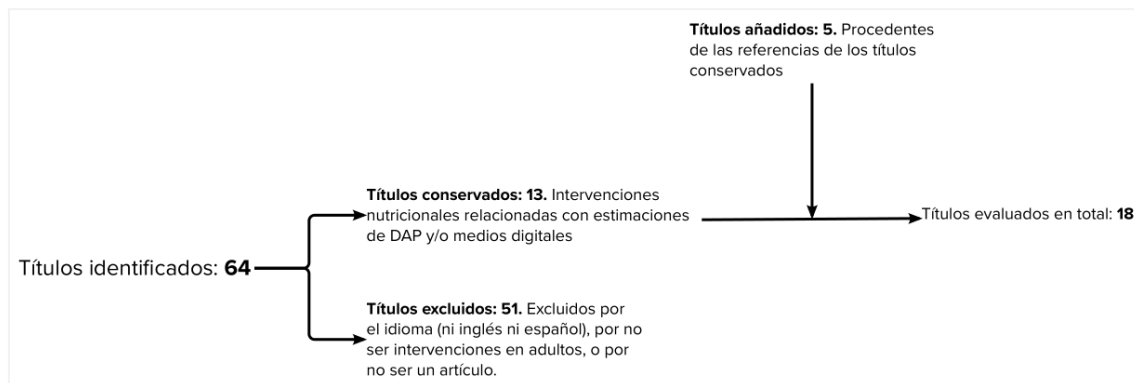


Figura 3.1. Resultados de la revisión de la literatura

En primer lugar, se realizó una breve revisión literaria para conocer el *estado del arte* de la nutrición personalizada y las actitudes de los consumidores ante ella. Para ello, se emplearon dos bases de datos diferentes, Scopus y Pubmed, y dos términos de búsqueda distintos que pueden encontrarse en el Apéndice 3.

La Figura 3.1 resume los resultados de la revisión literaria. De esta se obtuvieron ciertos datos relevantes que se incorporaron a nuestra investigación. A continuación, se mencionan algunos de los más interesantes.

- Existen tres subgrupos poblacionales con diferentes actitudes hacia la nutrición y salud (Ballco & de Magistris, 2019).
- Los pacientes con enfermedades crónicas tienen entre un 50% y 60% de probabilidades de participar en una intervención de estilo de vida (Veldwijk et al., 2013).

⁹ Applied Health Economics and Health Policy es una revista especializada en Economía de la Salud con factor de impacto de 2.561 en 2020 y Q2 en Economía (percentil 72.98) en 2020.

- El precio al usuario puede determinar la actitud de este hacia la nutrición personalizada (Rankin et al., 2018).
- Mayoritariamente, los consumidores están a favor de las políticas (medidas) de promoción de hábitos saludables (Foltz et al., 2012).
- Los usuarios de programas de *telesalud* puntúan la experiencia como positiva y útil (Haugen et al., 2007).
- Se ha demostrado que ciertos programas de *telesalud* han mejorado la salud de sus usuarios (Smith et al., 2016).
- Las actitudes hacia la nutrición personalizada están determinadas por el resultado esperado (Poínhos et al., 2014).
- Existe cierta disposición a pagar por nutrición personalizada en Europa. Programas de supervisión aumentan esta DAP (Fischer et al., 2016).

Estos hallazgos nos permitieron conocer cuál es el estado de la nutrición personalizada en la literatura académica reciente, lo cual nos proporcionó una valiosa información sobre el diseño de las alternativas en el DCE. Por ejemplo, entre las alternativas deberán incluirse esquemas de supervisión (por ejemplo, ofrecido por un/a nutricionista licenciada/o) pues aumentan las probabilidades de que los usuarios quieran pagar por la intervención. La tabla-resumen completa puede revisarse en el Apéndice 3, sección 3.1.

3.2.2 Definición de los atributos y niveles

El siguiente paso consistió en la definición de los atributos y niveles que formarían las alternativas del DCE. Para ello se empleó la información de la revisión literaria anterior, se revisaron las características de las aplicaciones nutricionales para móviles existentes en Google Play Store, y se empleó la investigación de Fischer et al. (2016) como referencia para algunos atributos. Con esta información, se creó una gran lista de características que posteriormente se combinó en atributos y niveles, tal como se sugiere en Bridges et al. (2011).

La revisión de aplicaciones nutricionales se llevó a cabo a través de la búsqueda de los términos “nutrición”, “dieta”, “nutrición personalizada” y “recomendaciones nutricionales” en la aplicación Play Store presente en cualquier *smartphone* Android. Los resultados se ordenaron por puntuación del usuario y número de descargas. Las aplicaciones fueron descargándose en orden y se tomó nota de sus características (funcionalidades y precio). Estas funcionalidades fueron actualizando una lista hasta que las nuevas aplicaciones probadas ya no añadían ninguna funcionalidad nueva y se consideró alcanzada la saturación. A través de este método se analizaron

un total de nueve aplicaciones con quince versiones diferentes y más de veinte funcionalidades. Se puede acceder a la lista completa en la Tabla A3.2. del Apéndice 3.

Por otro lado, también se consideraron como posibles atributos y niveles los tres hipotéticos programas de NP empleados en “Willingness to pay for personalised nutrition across Europe” de Fischer et al. (2016), que consistían en tres intervenciones nutricionales con crecientes niveles de personalización. A continuación, se crearon unos atributos y niveles procedentes de la lista a través del siguiente procedimiento: primero se eliminaron o fusionaron aquellas características duplicadas, se agruparon aquellas características similares con diferencias graduales, se adoptó el contenido restante a una estructura de atributos y niveles, restringiendo el número de elementos totales para no superar lo habitual (Johnson et al., 2013).

3.2.3 Pretests y refinamiento

El anterior procedimiento dio lugar a un diseño, es decir, de cinco atributos donde dos de ellos tenían cuatro niveles (nivel de personalización y precio), y los tres restantes contaban con dos niveles (seguimiento, rutinas de ejercicio, video-tutoriales de recetas). En la Tabla A3.3. del Apéndice 3 pueden observarse en detalle los atributos y niveles empleados. A partir de esta especificación, se construyó un diseño “óptimo”, es decir, se usó un matriz ortogonal generada en SPSS (IBM Corp., 2020) y se siguió el proceso detallado en Street & Louviere (2005) para generar los pares de cada perfil¹⁰. Pese a que los diseños ortogonales no son los más apropiados para experimentos de elección discreta, sí que pueden ser usados cuando no existe información previa, *priors*, para realizar un experimento piloto (Rose & Bliemer, 2014, 2009). De esta forma, se construyó un diseño experimental con dieciséis *choice sets*, con dos alternativas por cada *choice set*. Este diseño experimental fue presentado a una muestra de conveniencia de cuarenta personas y los resultados fueron analizados utilizando una regresión logística condicional (McFadden, 1974) en Stata (StataCorp LLC, 2017). Puede observarse una representación gráfica de los resultados de la regresión logística condicional en la Figura 3.2. Igualmente, tanto los resultados detallados del modelo (para el atributo precio codificado de forma lineal o dummy) como el código para su estimación en R pueden encontrarse en la sección 3.4 del Apéndice 3.

Como puede observarse en la Figura 3.2, la funcionalidad de “recetas” parece poco relevante para los consumidores, por lo que fue descartada para posteriores diseños. Por otro lado, los encuestados manifestaron que el número de *choice sets* era demasiado alto, y en posteriores

¹⁰ El diseño ortogonal generado en SPSS proporciona solo la primera alternativa de cada *choice set*. La segunda alternativa debe ser generada a partir de la primera, siguiendo alguno de los procedimientos detallados en Street & Louviere (2005).

encuestas el diseño fue dividido en dos bloques. Los resultados fueron presentados en VI Taller de Evaluación Económica EEconAES (20 de noviembre de 2019), donde los asistentes realizaron varias sugerencias, entre ellas la inclusión de un atributo con dos niveles que indicara si el servicio sería provisto de forma física (tradicional) o digital (a través de una aplicación móvil). También, algunos de los presentes nos plantearon la posibilidad de incluir un atributo haciendo referencia a suplementos nutricionales que podrían estar presentes en la recomendación. Además, señalaron que los niveles del atributo personalización debían ser reformulados pues, por ejemplo, para una persona con una dieta omnívora y ninguna alergia o intolerancia alimentaria no habría ninguna diferencia entre el nivel uno y el nivel dos.

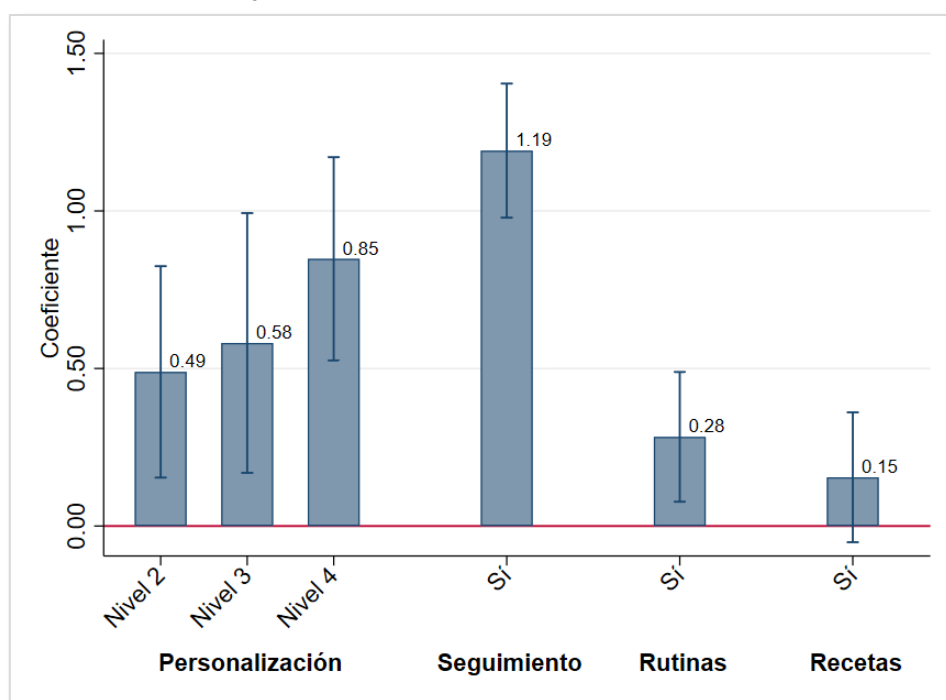


Figura 3.2. Resultados de la regresión logística condicional para el experimento piloto

Por ello, con la ayuda del nutricionista del proyecto Stance4Health, los niveles del atributo personalización fueron reformulados. Además, se incluyó el atributo que designaba el modo de provisión (digital o físico) y un atributo que permitía la inclusión de suplementos en la dieta. Estos cambios dieron lugar a la selección de atributos y niveles que pueden encontrarse en la Tabla A3.6 del Apéndice 3.

Estos atributos y niveles fueron usados para generar un diseño eficiente. Este diseño eficiente se generó usando el paquete “dcreate” (Hole, 2017) para Stata, y los *priors* empleados fueron los coeficientes obtenidos en el experimento piloto. Llegados a este punto, cabe mencionar ciertas salvedades. El experimento piloto incluía un número diferente de atributos y, en ciertos casos - como en el atributo personalización-, la definición de los niveles era ligeramente diferente. La existencia de más o menos atributos no debería ser un problema; de aquellos atributos que no

existían se asumió que los *priors* de sus niveles era igual a cero. No obstante, de aquellos atributos que sí existían y cuya definición de niveles era igual, no se estaría cometiendo un error de especificación pues cada coeficiente es calculado con relación al primer nivel del mismo atributo y no al resto de niveles del resto de atributos del experimento. Es decir, el *prior* de las “rutinas deportivas” representa la utilidad que el usuario obtiene cuando pasa de no tenerlas a tenerlas, por lo que no debería variar según cambien el resto de atributos. Sin embargo, los niveles del atributo personalización varían en este caso, por lo que estaríamos cometiendo un error de especificación si usáramos los *priors* de sus niveles equivalentes en el experimento anterior. No obstante, estos *priors* sí fueron empleados pues, como señalan Rose & Bliemer (2014), la eficiencia del diseño puede ser mayor incluso si los parámetros son aproximados en lugar de exactos.

Atributos	Niveles	
Nivel de personalización (cantidad de información recopilada para diseñar la dieta)	Nivel 1	Hábitos saludables y educación nutricional
	Nivel 2	Dieta personalizada basada en altura, peso y estatura.
	Nivel 3	Dieta personalizada basada en altura, peso, estatura, hábitos deportivos y análisis de sangre
	Nivel 4	Dieta personalizada basada en altura, peso, estatura, hábitos deportivos, análisis de sangre y análisis de la microbiota intestinal.
Formato de provision (cómo se va a prestar el servicio)	Físico	All services are provided in a physical facility that the user needs visit (with a frequency that depends on their personalisation level)
	Digital	All services (except blood and microbiome analyses) are provided through a digital platform via smartphone or computer.
Plan de entrenamiento (personalizado según el usuario)	No	No physical activity training plan.
	Sí	Physical activity training plan (the user gets a daily exercise schedule adapted to their needs).
Precio (cantidad mensual)	10 €	10 € / mes
	20 €	20 € / mes
	30 €	30 € / mes
	40 €	40 € / mes

Tabla 3.1 Selección definitiva de atributos y niveles

El nuevo diseño eficiente generado, fue dividido en dos bloques que fueron presentados aleatoriamente a 52 encuestados de un panel de consumidores¹¹ representativo de la población española. Estos 52 encuestados formaron parte de una prueba preliminar donde se pretendía, en

¹¹ Con posterioridad se detallan las características del panel de consumidores.

primer lugar, comprobar que la encuesta estimaría adecuadamente las preferencias de los consumidores, 2) estimar nuevos *priors* sobre los que basar el experimento definitivo. La encuesta fue programada en Qualtrics (Qualtrics, 2020) y enlazada con el sistema de monitorización de las respuestas del proveedor del panel de consumidores. A esta encuesta se le añadió una “pregunta de validación”, es decir, un *choice set* con una alternativa dominante que el encuestado debía escoger. De esta forma, si la persona encuestada no escogía la alternativa con mayor utilidad podríamos deducir que no estaba prestando atención a la encuesta, que no había entendido la tarea, o que existía algún problema relacionado con la plataforma de la encuesta.

Una vez que se obtuvieron las 52 respuestas, se analizaron los resultados en Stata usando una regresión logística condicional. Estos resultados fueron totalmente inconcluyentes por diferentes razones. Primeramente, los resultados no eran en absoluto significativos y los coeficientes diferían notablemente de los estimados en el ejercicio piloto. Después, las personas encuestadas tardaron muy poco tiempo en responder la encuesta; una media de 9.13 minutos cuando nuestra estimación rondaba los 15 minutos. Por último, el 48% de las personas que respondieron la encuesta fallaron la pregunta de validación.

Para solucionar estos problemas, se tomaron varias medidas. En primer lugar, se contactó con la empresa panelista, la cual nos trasladó que avisaría a los próximos encuestados para que prestasen más atención a la tarea. Otros cambios que se introdujeron fueron 1) la “gamificación” del diseño de los *choice sets*¹², 2) una “pregunta de entrenamiento” en la introducción¹³, 3) cambios en la interfaz y mejor accesibilidad desde teléfonos móviles y 4) simplificación de los *choice sets*. Los *choice sets* fueron simplificados a través de la reducción de los atributos y niveles empleados.

De esta forma, la selección final de atributos y niveles fue la presente en la Tabla 3.1. Se decidió eliminar el atributo “suplementos nutricionales”, pues parecía el menos relevante y más difícil de implementar, y se fusionó el atributo seguimiento con los niveles de personalización, tal como sugirió el nutricionista del proyecto Stance4Health. Tras realizar una nueva prueba con 52 personas encuestadas, los resultados fueron satisfactorios y se procedió con el experimento de elección discreta.

3.2.3 El experimento

¹² Al lado de cada nivel se incluyeron unos iconos que definían la funcionalidad.

¹³ En la introducción, antes de comenzar el experimento de elección discreta, se introdujo una “pregunta de entrenamiento”, un *choice set* donde una de las alternativas dominaba a la otra y, si el encuestado escogía aquella de menor utilidad, se le notificaba y se le pedía que prestase más atención y leyese las instrucciones adecuadamente.

El diseño óptimo (reducción del D-error asumiendo priors no bayesianos) del experimento de elección discreta se realizó en Stata (StataCorp LLC, 2017) empleando el módulo `-dcreate-` diseñado por Hole (2017) y usando los *priors* estimados en la regresión logística condicional del experimento piloto (el código de Stata puede revisarse en la Sección 3.6 del Apéndice 3. Como output de `-dcreate-` se obtuvo una matriz de diseño que representaba la distribución de alternativas y *choice sets* separada en dos bloques. Cada uno de los bloques representaba las preguntas que debía responder el 50% de la muestra. La matriz fue decodificada y se crearon los *choice sets* en Qualtrics. Se configuró la encuesta para que redireccionara aleatoriamente y con un 50% de probabilidades a cada encuestado a uno de los dos bloques. La Tabla 3.2 representa a uno de los *choice sets* empleados en la encuesta. Nótese que cada *choice set* contiene sólo dos alternativas y no existe una opción nula (no adquirir el servicio). Se tomó la decisión de no incluir una alternativa nula puesto que, al tratarse de un panel de consumidores donde los encuestados parecían estar poco comprometidos con la tarea, se temía que la opción nula fuese siempre elegida para finalizar rápidamente el experimento. Por este motivo se omitió la opción nula, pero se preguntó a cada encuestado cuántas de las alternativas elegidas adquiriría con su propio dinero (fuera del experimento). Esta variable se usó para controlar la no disposición a pagar.

La encuesta fue dividida en tres secciones. En la segunda se incluyeron todos los *choice sets* del experimento de elección discreta más un *choice set* de verificación¹⁴, sumando un total de ocho más un *choice set* (ocho *choice sets* provenientes del diseño original con dieciséis *choice sets* dividido en dos bloques + la pregunta de verificación). En la primera sección se introdujo un texto explicativo y la “pregunta de entrenamiento”. Se añadió un temporizador de un minuto que imposibilitaba avanzar hasta que se agotase ese tiempo para obligar a la persona encuestada a leer la introducción. La tercera sección de la encuesta se componía de diez preguntas demográficas: edad, altura, peso, género, educación, ingresos de la unidad familiar, número de miembros en la unidad familiar, horas de deporte a la semana, interés en nutrición saludable y enfermedades crónicas.

La empresa panelista fue elegida entre cinco alternativas por combinar experiencia y precio. A esta empresa se le encargó una muestra de 400 personas representativas de la población española. En este panel, cada vez que una persona responde una encuesta recibe una gratificación de 0.5€. Pese a que la muestra contratada era de 400 personas, la empresa panelista nos concedió tantas pruebas de 50 personas como se necesitasen hasta que la encuesta funcionase adecuadamente.

¹⁴ Un *choice set* con una alternativa dominante para comprobar que los encuestados están prestando atención a las elecciones que realizan.

3.4.2 Análisis

Los resultados de la encuesta fueron analizados empleando tres modelos de elección discreta: regresión logística condicional, regresión logística de parámetros aleatorios, regresión logística de clases latentes. Estos métodos están en línea con la metodología presentada en capítulos anteriores, por lo que se omitirá su desarrollo. Las líneas de código de Stata pueden revisarse en el Apéndice 3 y la sección 3.7.

	Option A		Option B	
Personalization level	Nivel 1: hábitos saludables y educ. nutricional.		Nivel 3: dieta personalizada por nutricionista basada en edad, peso, altura, hábitos deportivos y análisis de sangre	
Delivery format	Digital: servicio provisto vía una app o web.		Físico: servicio provisto en un lugar físico.	
Training plan	No incluido		Plan de entrenamiento diario basado en info. usuario	
Price	20 €/mes		30 €/mes	
Choice	<input type="checkbox"/>		<input type="checkbox"/>	

Tabla 3.2. Ejemplo de choice set empleado en el experimento.

3.3 RESULTADOS

La empresa proveedora del panel envió 2100 invitaciones y un total de 462 personas respondieron a la encuesta (52 en la prueba + 414 del experimento final¹⁵). Entre las personas encuestadas, 33 de ellas no pasaron la tarea de verificación, por lo que fueron descartadas del análisis estadístico. En esta muestra el 52% eran mujeres, la edad media era de 42.8 años, y el índice de grasa corporal medio era de 25.9 kg/m². Mayoritariamente, los encuestados pertenecían a hogares de cuatro miembros y mostraban un alto interés en la nutrición saludable (7.98/10, SC 1.92). La Tabla 3.3 contiene información sobre el resto de características sociodemográficas de la muestra.

Esta muestra se comparó con información sobre la sociedad española para comprobar la representatividad de la muestra. Se descubrió que la edad y los ingresos medios eran representativos mientras que las mujeres con sobrepeso y los hombres con obesidad estaban ligeramente sobrerrepresentados. También el número medio de miembros del hogar era mayor que en la sociedad española.

¹⁵ Las 52 respuestas se corresponden con la última prueba que resultó satisfactoria. Puesto que no se realizaron cambios desde esta versión a la definitiva del DCE, se unieron ambas bases de datos.

3.3.1 Regresión logística condicional

Variable	%	n
N	100	462
Age ^a	42.80	12.80
Height ^a	168.95	9.19
Weight ^a	74.03	15.44
Sex		
<i>Female</i>	53.03	245
<i>Male</i>	46.97	217
Education		
<i>No education</i>	0	0
<i>Primary education</i>	0.65	3
<i>Secondary education</i>	8.44	39
<i>Pre university education</i>	18.18	84
<i>Vocational formation I</i>	12.99	60
<i>Vocational formation II</i>	12.77	59
<i>University degree</i>	35.50	164
<i>Postgraduate</i>	11.47	53
Household income (monthly)		
<i>No income</i>	3.46	16
<i>less than 500€</i>	4.76	22
<i>500€ - 1,000€</i>	10.61	49
<i>1,000€ - 1,700€</i>	28.14	130
<i>1,700€ - 2,300€</i>	19.05	88
<i>2,300€ - 3,000€</i>	16.45	76
<i>3,000€ - 4,000€</i>	11.04	51
<i>4,000€ - 5,000€</i>	5.19	24
<i>more than 5,000€</i>	1.30	6
No. Of household members ^a		
<i>No. of members under 18 years old</i>	0.76	0.88
<i>No. of employed members</i>	1.50	0.81
<i>No. of unemployed members</i>	0.73	0.88
Hours sport per week ^a	3.75	2.89
Interest in healthy nutr. ^a	7.98	1.92
No. of people w/ NCDs	21.65	100

^a Las columnas representan la media y SD en lugar del porcentaje y el número de observaciones

Tabla 3.3. Características sociodemográficas de la muestra

El primer análisis se llevó a cabo con el modelo de elección discreta más básico: la regresión logística condicional. Se decidió especificar la variable precio de forma lineal y no dicotómica (dummy), a diferencia del resto de coeficientes. Esta hipótesis pudo confirmarse cuando se compararon los coeficientes de la variable especificados de forma lineal usando un t-test. La

especificación lineal, como se expuso durante los capítulos anteriores, permite la estimación de la disposición a pagar en los modelos de elección discreta.

El resultado del contraste de hipótesis de linealidad del coeficiente (el precio de la alternativa) puede encontrarse en la Tabla 3.4. Como puede observarse, el p-valor de los tres contrastes indica que no existen diferencias significativas entre los valores de los coeficientes especificados de forma dummy, por lo que asumimos que la pérdida de utilidad que se produce por cada euro de más es lineal.

Diferencias	$\beta_{20\text{€}-10\text{€}}$	$\beta_{30\text{€}-20\text{€}}$	$\beta_{40\text{€}-30\text{€}}$
$\beta_{20\text{€}-10\text{€}}$		0.1277 (p-val.: 0.174)	0.1008 (p-val.: 0.216)
$\beta_{30\text{€}-20\text{€}}$			0.0269 (p-val.: 0.731)
$\beta_{40\text{€}-30\text{€}}$			

Tabla 3.4 Contraste de hipótesis de linealidad del atributo precio

Los resultados de la regresión logística condicional se muestran en la Tabla 3.5. Como puede observarse, todos los coeficientes resultaron significativos excepto el formato de provisión. La utilidad que, de media, obtienen los individuos de los niveles de personalización aumenta cuanto mayor sea el nivel. No obstante, la diferencia de utilidad entre los niveles 3 y 4 no es significativa (p-valor: >0.0000; Apéndice 3, sección 3.7). Esto implica que la inclusión del análisis de microbiota (la principal diferencia existente entre el nivel de personalización 3 y 4) no aporta una utilidad diferencial al encuestado. Por otro lado, el formato de provisión no es significativo, por lo que los individuos son indiferentes entre ambas modalidades del servicio. La inclusión del programa de entrenamiento, no obstante, sí supone un incremento sustancial de la utilidad de los usuarios.

Atributos	Niveles	Coef.	SE (P>z)
Person.	Nivel 2	0.310	0.088 (>0.000)
	Nivel 3	1.202	0.080 (>0.000)
	Nivel 4	1.307	0.094 (>0.000)
Formato	Digital	0.046	0.051 (0.160)
Entrenamiento	Sí	0.831	0.060 (>0.000)
Precio	(Lineal)	-0.056	0.005 (>0.000)

Pseudo R² (ρ^2): 0.1794
L-pseudolikelihood: -1952.0002
AIC: 3916
BIC: 3957.005

Tabla 3.5 Resultados de la regresión logística condicional

Por último, el coeficiente de la variable precio indica la utilidad media que pierden los usuarios cuando el precio del servicio aumenta en un euro. De esta forma, el cociente de la utilidad obtenida

en cualquiera de las características del servicio entre la utilidad perdida por cada euro permite estimar la disposición a pagar de la característica. No obstante, se decidió estimar la disposición a pagar usando los resultados de la regresión logística de parámetros aleatorios ya que ofrecía unas mejores medidas de ajuste que el logit condicional.

3.3.2 Regresión logística de parámetros aleatorios

Los resultados de la regresión logística de parámetros aleatorios son equivalentes al anterior modelo. En la estimación de la Tabla 3.6, además de la información ofrecida en el logit condicional, en logit mixto nos permite conocer si la heterogeneidad de preferencias es significativa (a través de una simulación similar al método *Bootstrap* se estima un número de coeficientes, por ejemplo 1000, y se calcula su desviación estándar, lo cual es reportado en SD ($P > z$)).

De los resultados obtenidos podemos decir que existe una heterogeneidad sustancial en todos los parámetros estimados. Por ejemplo, el 35% de los encuestados no prefiere el nivel de personalización dos al uno, el 45.49% opta por el formato de provisión tradicional, y el 15% no quiere rutinas deportivas en servicio (en la sección 3.7 del Apéndice 3 se puede observar cómo se hicieron estos cálculos).

Atributos	Niveles	Coef.	SE ($P > z$)	SD ($P > z$)
Person.	Nivel 2	0.491	0.152 (0.001)	1.2133 (>0.000)
	Nivel 3	2.295	0.164 (>0.000)	0.4375 (0.07)
	Nivel 4	2.450	0.176 (>0.000)	1.1288 (>0.000)
Formato	Digital	0.160	0.091 (0.079)	1.4117 (>0.000)
Entrenamiento	Sí	1.517	0.116 (>0.000)	1.5210 (>0.000)
Precio	(Lineal)	-0.112	0.010 (>0.000)	0.1161 (>0.000)

Pseudo R^2 (ρ^2): 0.2579
L-pseudolikelihood: -1733.6422
AIC: 3491.284
BIC: 3573.293

Tabla 3.4. Resultados de la regresión logística de parámetros aleatorios

Niveles	WTP (media)	WTP (lim. inferior)	WTP (lim. superior)
Nivel 2	4.4 € ***	1.87 €	6.92 €
Nivel 3	20.56 € ***	17.34 €	23.77 €
Nivel 4	21.94 € ***	18.87 €	25.02 €
Formato digital	1.43 € *	-0.18 €	3.04 €
Rutinas deportivas	13.59 € ***	11.11 €	16.06 €

Tabla 3.5. Disposición a pagar estimada del logit mixto

La estimación de la disposición a pagar se ha recogido en la Tabla 3.7. Como puede observarse, los usuarios están dispuestos a pagar más por aquellos niveles que les reportan más utilidad, por ejemplo, el nivel de personalización 3 y 4. Análogamente, la disposición a pagar por aquellos niveles no significativos es cercana a cero. Además, el uso de los límites superiores e inferiores permite la consideración de diferentes escenarios de acogida del público. Por ejemplo, la disposición a pagar por un servicio de nivel 3, digital, con rutinas deportivas sería de 35.58 euros mensuales de media con un mínimo de 28.11 y un máximo de 42.87 euros.

3.3.3 Regresión logística de clases latentes

Para la estimación de la regresión logística de clases latentes, en primer lugar, se halló el número de clases latentes existentes en la muestra (Apéndice 3, sección 3.7, líneas 18-24). Los resultados del modelo con cuatro clases, aquel con mejores medidas de ajuste, pueden observarse en la Tabla 3.8.

La Clase 1 tiene preferencias débiles pero significativas por los niveles 3 y 4 de personalización. En este caso, el nivel 3 es ligeramente preferido al 4, y es por este motivo por el que no se encontraron diferencias significativas entre los dos niveles en el logit condicional. Por otro lado, la Clase 1 es la que muestra mayores preferencias por el formato digital y menores preferencias por las rutinas deportivas. Estos hallazgos están correlacionados con los mayores niveles de educación y de sedentarismo de la muestra, información que podría explicar el sentido de las preferencias declaradas (obsérvense las medias con relación a las covariables en la Tabla 3.9).

La Clase 2 tiene preferencias significativas por los niveles de personalización más avanzados, pero solo muestran una pequeña disposición a pagar diferencial entre los niveles 3 y 4. Esta clase prefiere que el servicio sea provisto de forma tradicional (con atención física) y tienen preferencias claras por las recomendaciones deportivas (muestran la mayor disposición a pagar por esta característica). Estos hechos coinciden con que es la clase con mayor interés en la nutrición saludable y que realiza bastante ejercicio físico de forma habitual.

La Clase 3 está compuesta por los individuos con mayor disposición a pagar en toda la muestra, especialmente para los niveles 3 y 4 de personalización. Este grupo es indiferente entre la provisión digital y física, y obtiene una moderada utilidad de las rutinas deportivas. Esta gran inelasticidad precio de la demanda se corresponde con aquellos consumidores con mayor renta y mayor número de enfermedades crónicas.

La Clase 4 no está interesada en el servicio de nutrición personalizada y tiene la menor renta de la muestra, mayor edad, y menor interés en la nutrición saludable. No obstante, esta clase realiza

una cantidad considerable de ejercicio físico y tiene el menor nivel de enfermedades crónicas. Estas características (poca renta, salud y suficiente ejercicio físico) pueden explicar los motivos por los cuales estos consumidores no están interesados en este servicio.

Atributos		Niveles % de muestra	Clase 1	Clase 2	Clase 3	Clase 4
<i>Variables alternativa - específicos</i>	<i>Person.</i>	Nivel 2	0.651	1.027***	0.640	0.469**
		Nivel 3	1.597***	2.153***	3.405***	0.489**
		Nivel 4	1.583***	2.272***	3.785***	0.159
	Formato	Digital	1.647***	-0.284**	-0.253	-0.763***
	Ejercicio	Incluido	0.524***	2.514***	0.602***	-0.157
	Precio	(Lineal)	-0.068***	-0.083***	-0.027**	-0.169***
Características			Clase 1	Clase 2	Clase 3	Clase 4
<i>Variables caso-específicas</i>	Ingresos (rango 1 – 7)		0.190	0.128	0.351***	
	Educación (años)		0.207**	-0.096	-0.100	
	Edad		-0.017	-0.019	-0.047***	
	Enfermedades crónicas		0.528	0.982**	1.188***	(Clase de referencia)
	Ejercicio (h. a la semana)		-0.200**	0.049	-0.148***	
	Interés en nutr. Salud.		0.081	0.280**	0.248***	
	Constante		-3.265**	-0.464	0.562	
			Clase 1	Clase 2	Clase 3	Clase 4
WTP (€)	Personaliz. 2		9.59*	12.42*	2.35	2.78**
			(-1.20 to 20.38)	(5.27 to 19.57)	(-14.17 to 18.56)	(0.45 to 5.11)
	Personaliz. 3		23.53***	26.04***	124.85***	2.90**
			(9.78 to 37.29)	(17.40 to 34.68)	(29.55 to 220.16)	(0.29 to 5.51)
	Personaliz. 4		23.32***	27.49***	138.78***	-0.94
			(11.09 to 35.56)	(19.35 to 35.63)	(40.27 to 237.28)	(-1.51 to 3.40)
Formato Digital		24.27***	-3.43***	-9.28	-4.53***	
Rutinas deportivas		(12.47 to 36.07)	(-6.72 to 0.15)	(-24.64 to 6.08)	(-6.15 to 2.90)	
		7.72**	30.42**	22.09*	-0.93	
	(1.02 to 14.41)	(24.68 to 36.17)	(0.68 to 44.85)	(-2.72 to 0.85)		

WTP: disposición a pagar. *: p-value < 0.10, **: p-value < 0.05, ***: p-value < 0.01. Notas: la primera sección muestra los coeficientes respectivos a las características de las alternativas. La segunda sección muestra covariables relativas a las características de los individuos. Estas variables pueden ser interpretadas como variables que incrementan la probabilidad de pertenecer a una clase con respecto a la clase de referencia (en este caso la cuarta). La última sección es la estimación de la disposición a pagar de cada clase por cada variable alternativa-específica. La disposición a pagar se presenta como (media)/(95% CI).

Tabla 3.8. Resultados de la regresión logística de clases latentes

A parte de las características mencionadas, no se han encontrado otras características que expliquen las elecciones poblacionales. Las decisiones no dependen de un componente de género ni IMC (altura y peso), aunque otras características correlacionadas con el IMC (como el ejercicio o el interés en nutrición personalizada) si se han tenido en cuenta.

Clase	Ingresos	Educación	Edad	EC	Ejercicio	Interés
1	4.04	15.53	41.44	0.16	2.64	7.78
2	3.72	13.72	43.55	0.27	4.68	8.43
3	4.21	14.22	39.32	0.27	3.12	8.27
4	3.51	13.81	45.35	0.15	3.80	7.40
Total	3.82	14.16	42.81	0.22	3.75	7.99

Ingresos (en nueve rangos), educación (en años de educación estimados), EC (porcentaje de personas con una o más enfermedades crónicas), ejercicio (media de horas a la semana), interés en nutrición personalizada (media, 0-10)

Tabla 3.9. Medias de las características de la muestra por clase

3.3.4 Análisis de sensibilidad

Se ha realizado un análisis de sensibilidad para estimar las disposiciones a pagar controladas por la probabilidad de adquisición de los consumidores. Para ello, de la muestra se dividió en dos partes: 1) aquellas personas con una probabilidad de adquisición menor al 33% (nótese que al final de la encuesta los encuestados respondieron a la pregunta *¿cuántas de las elecciones realizadas adquirirías en la realidad?*), 2) aquellos con una probabilidad de adquisición mayor o igual al 33%. De esta forma, se descubrió que la disposición a pagar del primer grupo era un 40% menor a la media y un 50% menor a la del otro grupo. Este hallazgo confirma que la probabilidad de adquisición (la variable usada de control) es un predictor fiable de la disposición a pagar.

3.4 DISCUSIÓN Y CONCLUSIONES

Esta investigación sugiere la existencia de un interés generalizado en la nutrición personalizada, materializada principalmente en el nivel de personalización 3 descrito anteriormente. Asimismo, los consumidores también han mostrado un interés significativo por otras características del servicio como las rutinas deportivas.

Al igual que otros estudios han identificado diferentes clases de consumidores, principalmente *partidarios* y *no partidarios* (Determann, Lambooi, Gyrd-Hansen, et al., 2017; Foltz et al., 2012), esta investigación va más allá clasificando a los consumidores en cuatro clases según sus elecciones y controlados por sus características sociodemográficas. De esta forma, ha sido posible descubrir que las actitudes *a priori* de los consumidores dependen de la renta, de su educación, de la edad, de las enfermedades crónicas que padecen, de algunos hábitos como el ejercicio, y del interés que tienen en la nutrición saludable.

Gracias a los resultados del modelo de clases latentes, se ha realizado una clasificación *informal* del tipo de consumidores existentes que podría ser útil para la segmentación y provisión del servicio de forma privada o pública:

- Clase 1: consumidor digital (*tech savvy*): son personas con un nivel de educación alto, hábitos sedentarios, que presentan una gran preferencia por el formato de provisión digital pero que no obtienen mucha utilidad ni de los consejos nutricionales ni de los deportivos. Estas personas serían el público adecuado para una aplicación de nutrición personalizada simple y barata.
- Clase 2: deportistas: esta clase se caracteriza por estar compuesta por personas más interesadas en el deporte que en la nutrición personalizada (son el grupo que declara más horas de deporte a la semana). Para esta clase, la introducción del servicio de atención deportiva parece ser un determinante de la disposición a adherirse a la intervención. Declaran una disposición a pagar media y podrían adoptar el servicio si el coste no fuera un impedimento.
- Clase 3: partidarios de la nutrición personalizada: esta clase de usuarios están interesados en la nutrición personalizada y, debido a que tienen la mayor renta, están generalmente dispuestos a pagar una gran cantidad posiblemente porque esperan unos beneficios en salud mayores al resto de la muestra (son el grupo con mayor porcentaje de enfermedades crónicas).
- Clase 4: no partidarios: este grupo mostró poco interés en cualquier tipo de nutrición personalizada y, por tanto, una bajísima disposición a pagar (lo cual puede ser explicado por la renta media del grupo). Otros motivos que condicionan las preferencias de este grupo son la baja prevalencia de enfermedades crónicas y buenos hábitos de rutinas deportivas.

Algunas de las conclusiones que se alcanzan en este capítulo son compatibles con ciertos hallazgos de la literatura. Por ejemplo, la mejora de la salud y la prevención de enfermedades son fuertes motivantes para la adopción de intervenciones personalizadas (Vallée Marcotte et al., 2018; Veldwijk et al., 2013). Aun así, ninguno de los cuatro grupos muestra una disposición a pagar diferencial relevante entre los niveles de personalización 3 y 4. Puesto que la única diferencia que existía entre estos niveles era el análisis y orientación hacia la microbiota intestinal, podemos declarar que los consumidores tienen poco interés *a priori* en la nutrición orientada hacia la microbiota. Los resultados del proyecto Stance4Health y una labor divulgativa podrían variar esta percepción, pero hasta que esto no ocurra los resultados de un experimento de elección discreta

de estas características mostrarán el poco interés generalizado de la población en este aspecto de la salud digestiva.

Limitaciones

Los resultados encontrados son sólo aplicables a España, aunque existen estudios que señalan similitudes en preferencias nutricionales entre los países europeos (Fischer et al., 2016). Por otro lado, este estudio no ha contemplado aspectos como la privacidad de los datos de los usuarios. Pese a que muchos estudios señalan que la privacidad de los datos es un aspecto de gran importancia para muchos consumidores (Feng-Gu et al., 2021; Fischer et al., 2016; Mustapa et al., 2020; Póinhos et al., 2014; Rankin et al., 2017; Stewart-Knox et al., 2015; Vallée Marcotte et al., 2018), la mayoría de ocasiones fallan al tomar medidas para proteger sus datos (paradoja de la privacidad) (Kokolakis, 2017).

Conclusiones

Este experimento de elección discreta ha demostrado que existe un mercado para un servicio de nutrición personalizada (también con atención deportiva) para un grupo sustancial de consumidores. Mientras que un quinto de los consumidores (Clase 3) está claramente motivado y dispuesto a pagar por él, el resto de grupos adoptarán la intervención sólo si el precio es lo suficientemente bajo. Nuevas investigaciones deben indagar en si el resto de grupos (Clase 1, Clase 2, y, sobre todo, Clase 4) pueden ser atraídos a la intervención o, al contrario, no la requieren o necesitan.

Capítulo 4. Explorando métodos de optimización de Experimentos de Elección Discreta



4.1 CONTEXTO

A continuación, se presentan los resultados de una investigación realizada para encontrar una estrategia secuencial óptima en línea con la propuesta de Bliemer y Rose (2010a). La siguiente sección está dividida en introducción, métodos, resultados y conclusiones. De la introducción se ha omitido todo el contexto, el cual ya no era necesario dado el contenido de los anteriores capítulos. El texto ha sido adaptado y traducido del artículo publicado en *Journal of Choice Modelling*¹⁶ (<https://doi.org/10.1016/j.jocm.2022.100357>).

A continuación, se desarrolla un estudio de simulación en el que se pone a prueba la estrategia secuencial explicada en la introducción de la siguiente sección. La estrategia secuencial original proponía actualizar priors cada vez que una nueva respuesta era registrada en el DCE, pero a continuación se prueba cómo varía la precisión y sesgo de las estimaciones cuando los priors se actualizan cada 5, 10, 20 o más respuestas. Estas simulaciones se realizaron en R, usando un código programado para la ocasión. A continuación, se presentan los detalles de forma más técnica. Las conclusiones, en la última sección, indican que actualizar priors cada cinco encuestados podría ser la mejor alternativa.

4.2 ESTRATEGIA SECUENCIAL ÓPTIMA PARA DCEs

Para resolver el problema de la especificación errónea de los priors, en el capítulo dos se han mencionado dos aproximaciones diferentes. Una de ellas, atribuible a Sándor y Wedel (2001), propone tratar a los priors como variables aleatorias que se distribuyen según una función de distribución. En una segunda aproximación más reciente, propuesta por Bliemer y Rose (2010a), los autores sugieren actualizar los priors y el diseño experimental cada vez que un encuestado responde al DCE. Este método, conocido como método serial o secuencial, puede resumirse en los siguientes pasos:

- a. Todos los priors son iguales a cero.
- b. Dados los priors, encuentra un diseño experimental que minimice el D-error.
- c. Una vez encontrado el diseño, presenta la encuesta a un encuestado.
- d. Mezcla las respuestas del encuestado con las registradas anteriormente (si hay).
- e. Utiliza en logit condicional para estimar la función de utilidad latente.
- f. Reemplaza los priors con los coeficientes obtenidos (si son significativos).
- g. Vuelve al punto b) o finaliza el DCE si no hay más encuestados.

¹⁶ Journal of Choice Modelling es una revista especializada en modelos de elección discreta con factor de impacto de 3.091 en 2020 y Q1 en Economía (percentil 75.93) en 2020.

Aunque es conocido que este método obtiene unas buenas estimaciones, puede ser muy caro y tedioso de realizar incluso si el DCE es realizado en línea pues siempre habrá dos o más encuestados que quieran responder la encuesta simultáneamente (Yu et al., 2011).

En esta investigación, se parte de la hipótesis de que los diseños secuenciales mantienen, en cierta medida, su rendimiento cuando los priors son actualizados con menor frecuencia (por ejemplo, cada 5, 10, 20, o 100 encuestados). En caso de que la hipótesis fuese cierta, la eficiencia del método serial incrementaría debido a que se obtendría el mismo rendimiento consumiendo menos recursos y tiempo (partiendo de la base de que es menos costoso actualizar los priors, y el diseño, cada $n > 1$ encuestados que cada $n = 1$ encuestados). El objetivo de este estudio es probar el rendimiento de estrategias secuenciales que actualicen sus priors menos frecuentemente y compararlas a la propuesta original.

4.3 MÉTODOS

Se empleó el paquete “idefix” (Traets et al., 2020) en R (R Core Team, 2021) para simular los resultados de once estrategias secuenciales diferentes. Se siguió un método Monte Carlo, por lo que cada simulación realizada fue repetida cien veces para confirmar que los resultados no fueron debidos al azar. Cada una de las once estrategias simuladas se diferenciaba del resto en el número de encuestados empleados para actualizar los priors del diseño. Además, cada una de las estrategias fue simulada empleando un rango de 8 a 16 choice sets. Todos los DCE simulados emplearon la misma especificación de atributos y niveles, $2^2 \times 3$ (dos atributos de dos niveles más uno de tres niveles). Las preferencias empleadas para simular las respuestas se corresponden con los coeficientes $\beta_{11} = -2$, β_{21} , $\beta_{31} = 1$, $\beta_{32} = 2$, y se empleó la codificación por efectos para realizar las estimaciones. En este caso, β_{xy} hace referencia al coeficiente “real” del nivel y del atributo x . Estos coeficientes “reales”, que conocemos pues es el valor que usamos para simular las preferencias, serán denominados “coeficientes objetivos” en adelante. Por tanto, una estrategia será mejor cuanto más se acerque $\tilde{\beta}_{xy}$ a β .

Un número sustancial de pruebas se realizó antes de decidir la especificación de atributos y niveles que se emplearía en las simulaciones ($2^2 \times 3$), la cual se eligió como punto intermedio entre el diseño de mayor tamaño posible cuyas simulaciones se llevaran a cabo en un tiempo computacional razonable. Cada una de las simulaciones, definiendo simulación como una determinada estrategia usando un número de choice sets de 8 a 16, fue completada en un rango de tiempo que fue desde los 30 minutos hasta las 26 horas (dependiendo de la complejidad del diseño). Se estima que se tardó alrededor de 300 horas en obtener los resultados de todas las simulaciones en un ordenador con un procesador AMD de 6 núcleos de 3.79 GHz y 16 GB de

RAM. La Tabla 4.1 contiene un resumen de las estrategias empleadas. Como puede observarse, en todos los casos se simularon las respuestas de un total de 200 encuestados.

Id de la simulación	Iteraciones (I)	Encuestados por iteración (n)
1	1	200
2	2	100
3	3	50 (x2) + 100
4	4	50
5	6	20 (x5) + 100
6	10	20
7	11	10 (x10) + 100
8	20	10
9	21	5 (x20) + 100
10	40	5
11	200	1

Tabla 4.1 Iteraciones y encuestados por iteración

En la Figura 4.1 puede observarse una representación gráfica del código de R empleado para automatizar las simulaciones (Sección 4.1 del Apéndice 4). El código consiste en un bucle que crea un diseño D-eficiente basado en los priors del último logit condicional estimado. El código representado en la Figura 4.1 ejecuta el siguiente procedimiento:

- Inicializa el código especificando n (encuestados por iteración), I (número de iteraciones), y S (número de choice sets). Los priors iniciales siempre son cero y los coeficientes objetivo $\beta_{11} = -2$, β_{21} , $\beta_{31} = 1$, $\beta_{32} = 2$.
- Computa un diseño D-eficiente usando el paquete “idefix” (Traets et al., 2020).
- Simula las respuestas de n encuestados para el DCE generado.
- Mezcla las respuestas junto con las obtenidas en anteriores iteraciones (en el caso de que hubiera alguna).
- Estima un logit condicional con la base de datos al completo.
- Los parámetros del logit condicional reemplazan a sus correspondientes priors solo si son significativos ($\alpha = 0.05$).
- Vuelve al punto b). Cuando el procedimiento sea repetido I veces, pasa al punto h).
- Almacena los coeficientes de cada iteración. Borra la base de datos.
- Vuelve al punto a). Finaliza el bucle cuando se haya repetido cien veces.

De esta forma, el código en la sección 4.1 del Apéndice 4 resultó en cuatro hojas de Excel por simulación (una por cada coeficiente simulado) con I filas (donde se puede observar cómo el coeficiente va variando cuando n respuestas son añadidas en cada iteración) y cien columnas (donde el mismo procedimiento se repite cien veces para cumplir con el método Monte Carlo).

Cada estrategia dio lugar a nueve resultados diferentes (8-16 choice sets) con cien estimaciones (Monte Carlo) para cada uno de sus cuatro coeficientes. En el análisis principal, se estudió la precisión y sesgo de cada estrategia usando métodos diferentes. El sesgo, el cual se consideró significativo o no significativo de acuerdo con los resultados de un t-test, fue definido como la desviación de la media de las cien estimaciones del coeficiente objetivo (ver Figura 4.2). No obstante, la precisión, medida por el error relativo, se definió como el tamaño de las diferencias entre el coeficiente objetivo y las cien estimaciones (ver Figura 4.2).

Como puede observarse en la Tabla 4.1, la estrategia 11 representa la propuesta exacta de Bliemer y Rose (2010a). Al contrario, la estrategia 1 representa el procedimiento seguido en numerosas investigaciones, donde se computa un diseño D-eficiente basado en priors iguales a cero, y se presenta al total de encuestados. El resto de las estrategias usan un número constante o no constante de encuestados por iteración (por ejemplo, las estrategias 2, 4, 6, 7, 10 y 11 son constantes porque usan el mismo número de encuestados por iteración. Por otro lado, las estrategias 3, 6, 7, y 9 pueden ser consideradas no constantes). En un análisis secundario, los resultados obtenidos en el análisis principal fueron puestos a prueba modificando algunas características de los DCE simulados.

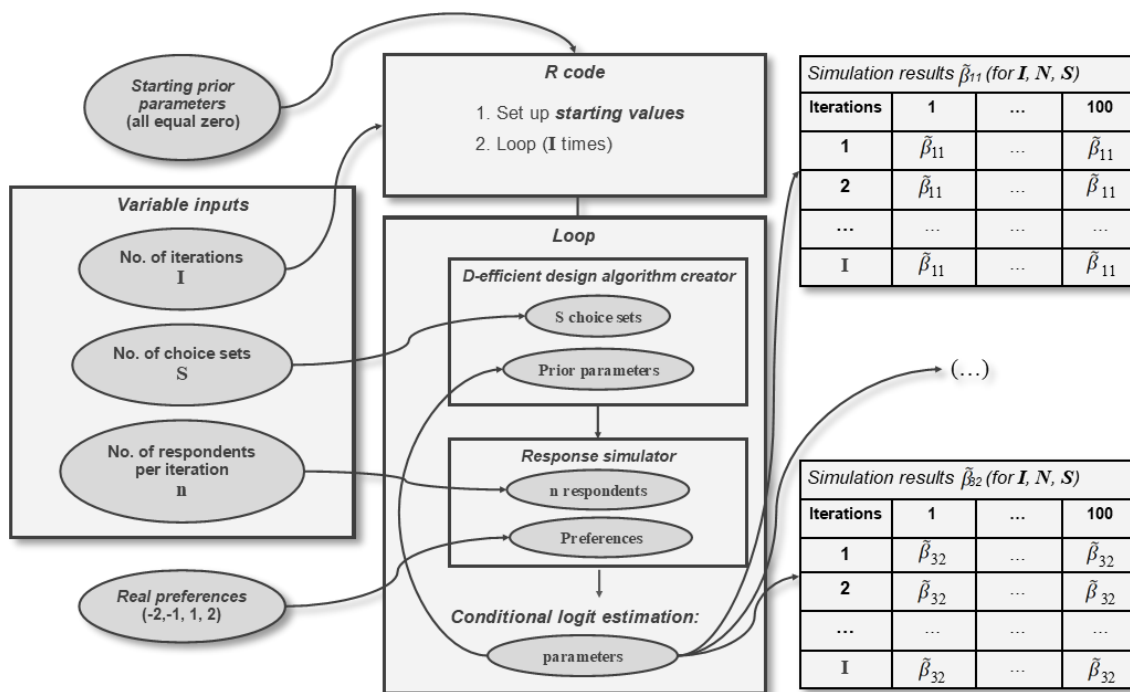


Figura 4.1 Representación del Código R empleado para las simulaciones

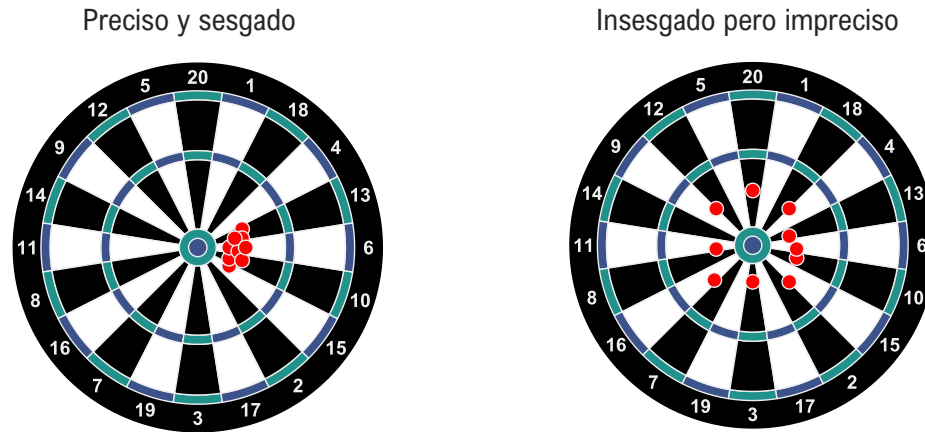


Figure 4.2 Ilustraciones de la precisión y la sesgidez

El paquete *idefix* (Traets et al., 2020) emplea la función CEA (algoritmo de intercambio coordinado – coordinate exchange algorithm). La función CEA trata a los priors como variables aleatorias, por lo que todas las estrategias están basadas en el enfoque bayesiano. Este planteamiento es equivalente a la primera definición de diseños secuenciales bayesianos individualmente adaptados (IASB) dada en Yu (2011) pero no empleada.

4.4 RESULTADOS

Se llevaron a cabo un total de $11 \times 9 \times 100 = 9\,900$ simulaciones usando el código de R proporcionado en el Apéndice 4. Para el análisis, nos centramos en las estimaciones alcanzadas en la última iteración de los DCE simulados (que se corresponderían con la estimación final de la que se obtendrían las preferencias). Así, los análisis de precisión y sesgo se basaron en los últimos coeficientes de cada repetición Monte Carlo de cada simulación usando un número diferente de choice sets, en cada uno de los cuatro parámetros simulados.

Precisión

La precisión de las estimaciones se midió usando la media del error relativo, calculado como $r_e = \left| \frac{\tilde{\beta}_{xy} - \beta_{xy}}{\beta_{xy}} \right|$. Cada estimación fue comparada con su correspondiente coeficiente objetivo: $\beta_{11} = -2$, $\beta_{21} = -1$, $\beta_{31} = 1$, or $\beta_{32} = 2$. La Figura 4.3 muestra la media del error relativo para los cuatro coeficientes dependiendo de la estrategia empleada y el número de choice sets empleados (una presentación alternativa de estos resultados puede encontrarse en la sección 4.2 del Apéndice 4).

Como se muestra en la Figura 4.3, las estrategias 10 ($l=40$, $n=5$) y 11 ($l=200$, $n=1$) logran un error relativo bajo sin importar el número de sets que usen, lo que implica que sus estimaciones son

muy cercanas a los coeficientes objetivo. No obstante, salvo por pocas excepciones, el error relativo del resto de estrategias refleja en general poca precisión. Las excepciones incluyen a las estrategias 6 ($I=10$, $n=20$) y 8 ($I=20$, $n=10$) sólo cuando 8 y 8 o 10 sets son usados respectivamente.

Los enfoques más precisos, estrategias 10 y 11, muestran un error relativo más bajo cuanto mayor es el número de sets que emplean. Sin embargo, esta tendencia no es generalizable al resto de estrategias. Como se ha mencionado en el párrafo anterior, el resto de estrategias, especialmente las estrategias 6 y 8, son más o menos precisas dependiendo del número de sets empleados aparentemente sin una ordenación lógica. En general, hemos detectado que la precisión es mayor cuando se usan 8, 10 o 16 sets. Este hallazgo sugiere la existencia “puntos óptimos” que pueden o no depender de otras características del diseño experimental como, por ejemplo, el número de atributos y niveles o el número de alternativas empleadas.

Hasta ahora, todas las estrategias definidas como precisas se corresponden con lo que antes nombramos “estrategias constantes” (número constante de encuestados por iteración). De esta forma, las estrategias no constantes no parecen ser viables para obtener estimaciones precisas.

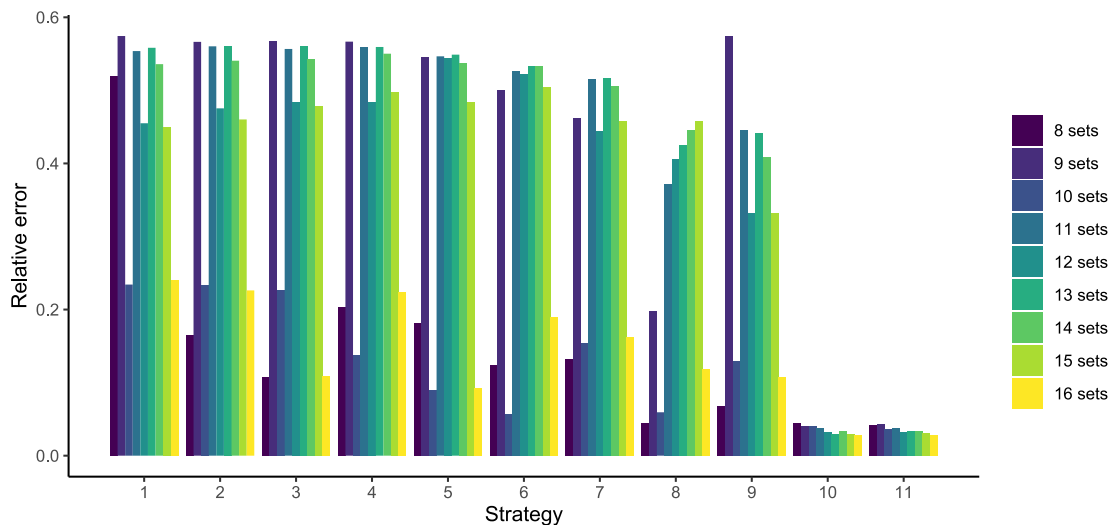


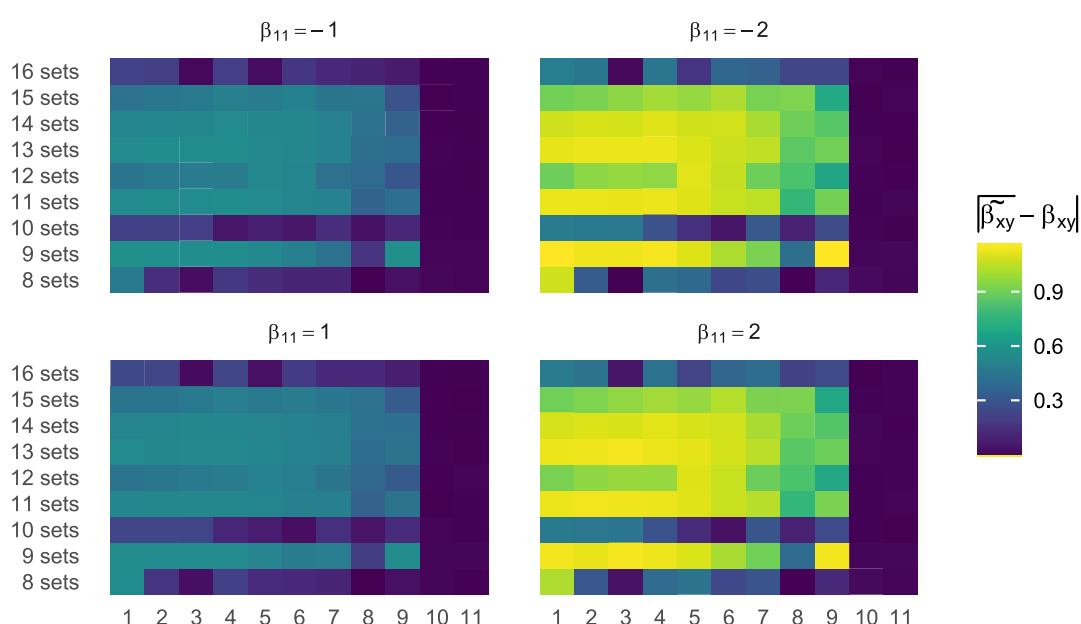
Figura 4.3. Error relativo de las estimaciones

Sesgo

En la Figura 4.4 se presenta un mapa de calor que permite una rápida visión general de la distancia absoluta entre la media de la estimación y el coeficiente objetivo. La significatividad del sesgo fue contrastada usando un t-test y presentada en la Tabla 4.3.1 de la sección 4.3 del Apéndice 4). En la Figura 4.4, cuanto más cercana esté una estimación a cero (morada), más probable es que no esté significativamente sesgada.

La mayoría de las estimaciones realizadas con las estrategias 10 ($I=40, n=5$) y 11 ($I=200, n=1$) no están significativamente sesgadas. El resto de estimaciones, con pocas excepciones, están sesgadas significativamente positiva o negativamente con respecto al coeficiente objetivo. Las excepciones incluyen las estrategias 3 ($I=3, n=[50, 50, 100]$), 6 ($I=10, n=20$), y 8 ($I=20, n=10$), las cuales no están significativamente sesgadas (o están muy cerca de no estarlo) cuando son empleadas con un número específico de choice sets.

Cuando el análisis se lleva a cabo en función del número de sets, la Figura 3 muestra algunos patrones interesantes; es fácilmente observable como la distancia relativa se reduce cuando se emplean 8, 10 y 16 choice sets.



Notas: cada panel representa las estimaciones de una de los cuatro coeficientes. Cada sector (área cuadrada) ilustra una elección de diseño: una combinación entre una estrategia de diseño (eje x) y un número de choice sets (eje y). El color de las áreas cuadradas ilustra cómo de cerca está la estimación del coeficiente objetivo en valor absoluto. $\bar{\beta}_{xy}$ representa la media de las estimaciones (cada número del apéndice 4.3), y β_{xy} el coeficiente objetivo.

Figura 4.4 Sesgo según la estrategia y el número de choice sets

Bajo las circunstancias en las que se han efectuado las simulaciones (3 atributos de $2^2 \cdot 3$ niveles, dos alternativas y 200 encuestados) solo unas pocas estrategias pueden ser consideradas precisas y no significativamente sesgadas. Este es el caso de las estrategias 10 y 11, cuyo rendimiento mejora cuantos más choice sets se usan. Puesto que la estrategia 11 es la propuesta original secuencial de Bliemer y Rose, este resultado era esperado. No obstante, hasta donde conocemos, es la primera vez que el rendimiento del resto de estrategias es puesto a prueba empíricamente. En estas, se ha hallado que, además de la estrategia original, la estrategia 10 ($n=40, I=5$) fue precisa y no significativamente sesgada en todos los casos analizados. Lo mismo no puede ser dicho del resto de estrategias, las cuales están significativamente sesgadas o sin

imprecisas en la mayoría de los casos. No obstante, ciertas combinaciones escapan de esta generalidad, teniendo un rendimiento bueno cuando cierta estrategia es combinada con cierto número de choice sets (por ejemplo, estrategia 6 con 10 choice sets). Para comprobar la consistencia de estos resultados, a continuación, se realizan dos análisis de sensibilidad.

Análisis de sensibilidad

En esta sección, la consistencia y fiabilidad de los resultados anteriores es puesta a prueba bajo diferentes circunstancias. Primero, el impacto de un número mayor de atributos y niveles es probado añadiendo un nuevo atributo de cuatro niveles al diseño experimental de los DCE simulados. A continuación, el impacto de una reducción del tamaño muestral se prueba reduciendo el número de respuestas simuladas a N=100.

a. Número de parámetros

La capacidad de las estrategias de ser precisas y no significativamente sesgadas se estudió al añadir un nuevo atributo de cuatro niveles (por lo que el nuevo diseño sería $2^2 \cdot 3 \cdot 4$). Sin embargo, debido a las limitaciones de hardware y tiempo (el tiempo de computación incrementó sustancialmente cuando se añadió el nuevo atributo), el método Monte Carlo se redujo a 20 repeticiones para las estrategias 7, 8, 9, 10, y 11 (de 8 a 11 sets) y 10 repeticiones en la estrategia 11 de (de 12 a 16 sets).

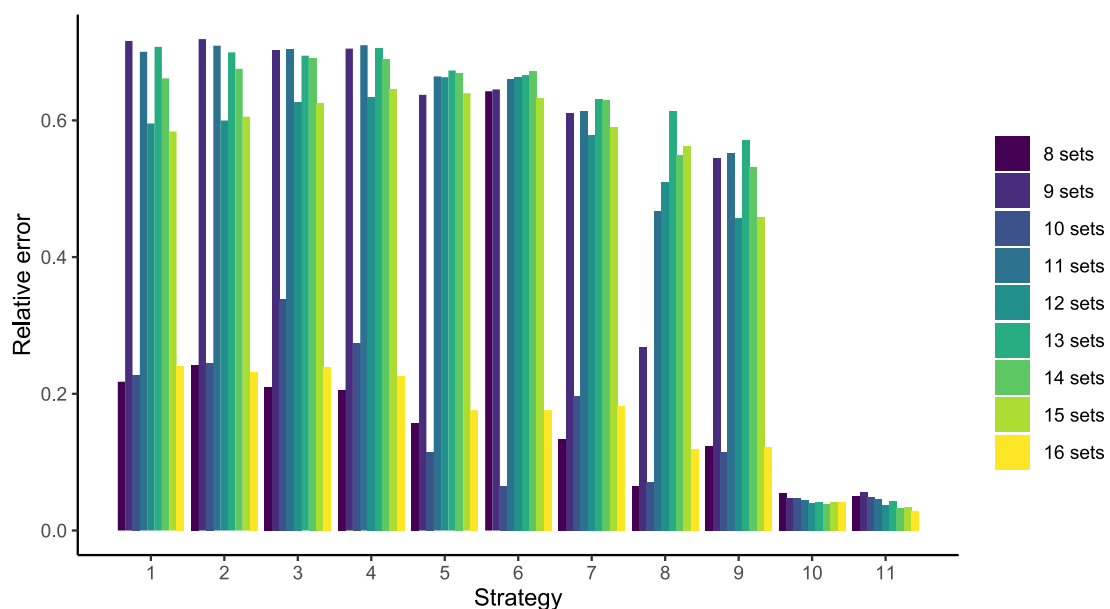


Figura 4.5 Error relativo de las estimaciones (para un diseño experimental $2^2 \cdot 3 \cdot 4$)

La magnitud de los errores relativos de cada estrategia puede observarse en la Figura 4.5 y la Tabla 4.2.2 del Apéndice 4. Como se muestra, la adición del nuevo atributo no modificó los

resultados obtenidos anteriormente (incluyendo a las estrategias 6 y 8 que siguen teniendo un bajo error relativo cuando usan 10 o 8 y 10 choice sets respectivamente).

De nuevo, la significatividad del sesgo fue comprobada usando un t-test. Los resultados, los cuales pueden ser observados en detalle en la Tabla 4.3.2 del Apéndice 4, son muy similares a los resultados del análisis principal. Las estrategias 10 y 11 muestran un sesgo no significativo, sin importar el número de choice sets empleados. Entre el resto de estrategias, la estrategia 8 no muestra un sesgo significativo para la mayoría de coeficientes cuando se emplean 8 sets, mientras que la estrategia 6 no está significativamente sesgada en algunos parámetros cuando se emplean 10 sets. A diferencia del análisis principal, la estrategia 3 mostró un sesgo significativo al incrementarse el número de parámetros.

b. Tamaño muestral

Para estudiar los efectos de la reducción del tamaño muestral se llevaron a cabo las mismas simulaciones usando 100 encuestados en lugar de 200, $N=200$. No se llevó a cabo el análisis de sensibilidad con un número mayor de encuestados pues se asumió que esto sólo aumentaría el rendimiento de las estrategias. El método Monte Carlo se limitó a 20 repeticiones como en el caso anterior. Además, sólo se simularon las estrategias 6, 8, 10 y 11 puesto que se supuso que el resto de las estrategias simplemente serían menos precisas e insesgadas. Debido a la reducción en el número de encuestados, el número de iteraciones necesarias para completar la encuesta se redujo a la mitad: estrategia 6 ($l=5$, $n=20$), estrategia 8 ($l=10$, $n=10$), estrategia 10 ($l=20$, $n=5$), estrategia ($l=100$, $n=1$).

Como se muestra en la Figura 4.6 y la Tabla 4.2.3, en las estrategias 10 y 11 el error relativo era más alto (que en las encuestas con 200 encuestados), especialmente cuantos menos choice sets eran empleados. Esta pérdida de precisión es atribuible a la reducción del tamaño muestral, pero es resuelta cuando se emplean más choice sets. Por otro lado, las estrategias 6 y 8 muestran un error relativo menor cuando se emplean 8, 10, y 16 sets. Con respecto a la significatividad del sesgo de las estimaciones, las conclusiones son análogas a aquellas que se obtuvieron en el análisis principal (ver la Tabla 4.3.4 del Apéndice 4).

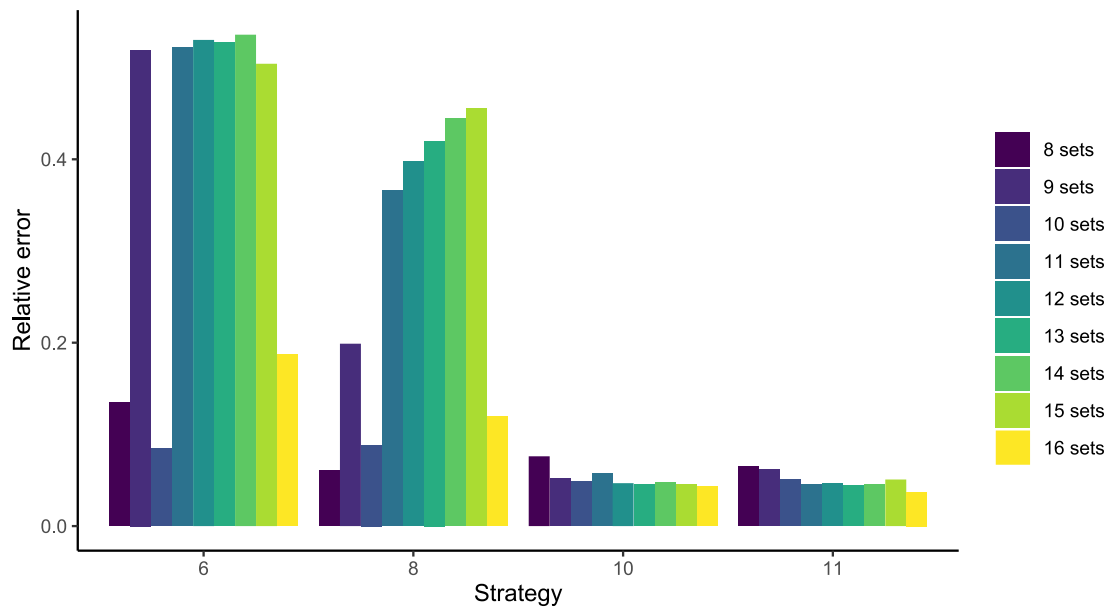


Figura 4.6 Error relativo de las estimaciones usando 100 encuestados

4.5 DISCUSIÓN Y CONCLUSIONES

En este artículo, se ha llevado a cabo una investigación para encontrar alternativas secuenciales más eficientes que la propuesta original de Bliemer y Rose (2010b). Para ello diez estrategias seriales diferentes han sido puestas a prueba y comparadas con la original a través de un estudio basado en simulaciones. Evaluando la precisión y sesgo de cada estrategia, se ha descubierto que, para que una estrategia sea precisa e insesgada, puede no ser necesario actualizar el DCE tras cada respuesta. Sin embargo, los resultados sugieren que actualizar los priors con relativa frecuencia y regularidad puede ayudar a que la estrategia tenga un rendimiento similar (o igual) a la estrategia original. Un buen ejemplo es la estrategia 10, la cual propone actualizar el diseño cada 5 respuestas, en lugar de cada respuesta como en la propuesta original, y ha demostrado tener un rendimiento similar en todos los casos comprobados.

Se halló que algunas estrategias tenían un buen rendimiento sólo cuando un número específico de choice sets se usaba. Sin embargo, no podemos afirmar que estos resultados sean independientes a las características del DCE simulado, por lo que tampoco podemos generalizar estas conclusiones a DCEs con otras características. Por otro lado, las razones por las que estas estrategias, las estrategias 6 y 8, han demostrado ser precisas e insesgadas bajo unas circunstancias específicas siguen siendo una incógnita. Nuestros resultados sugieren que existe una estrategia óptima diferente dependiendo del número de atributos, niveles, alternativas y sets empleados en el diseño. Por ejemplo, para un diseño experimental $2^2 \times 3$ con ocho sets, la estrategia 6 sería óptima porque obtiene estimaciones precisas e insesgadas a la vez que minimiza el número de iteraciones (y el coste) del experimento. Otras estrategias óptimas podrían ser

encontradas incrementando el número de simulaciones de DCEs con diferentes características. Sin embargo, nuevas simulaciones de diseños experimentales más complejos requieren un hardware más potente que el que se ha empleado en esta investigación. El código y resultados están disponibles para que otros investigadores puedan analizarlo, replicarlo y modificarlo.

Dado que puede que no existe una “estrategia óptima secuencial” universal, no podemos realizar recomendaciones prácticas universales. No obstante, si que podemos adelantar dos nuevas hipótesis. En primer lugar, estrategias análogas a la estrategia 10 (actualizar cada 5 encuestados) son probablemente tan precisas e insesgadas como la propuesta original para la mayoría de diseños experimentales (por lo que estaríamos ante una estrategia dominante). En segundo lugar, probablemente, cada configuración de DCE (definida por sus atributos, niveles, sets y alternativas) tenga su propia estrategia secuencial óptima

Capítulo 5. Diseños experimentales eficientes y accesibles



5.1 INTRODUCCIÓN

En los capítulos anteriores se han estudiado numerosos aspectos teóricos y prácticos de los experimentos de elección discreta. Hasta ahora, se han obtenido diferentes conclusiones que podemos enumerar a continuación. En primer lugar, en el ámbito teórico se han establecido los modelos de elección discreta (no lineales) como las herramientas de análisis básico de los resultados de los experimentos de elección discreta. Específicamente, hasta ahora la regresión logística condicional ha sido el modelo de elección discreta por defecto, y sus variaciones, como el logit mixto o el logit de clases latentes, han permitido superar algunas de las restricciones planteadas por el primero. El asentamiento del logit condicional supuso, por otro lado, que los diseños experimentales de los DCE fueran confeccionados para incrementar la precisión de sus estimadores. En concreto, en la literatura académica en la materia existe cierto consenso en que un DCE es eficiente (u óptimo) cuando el diseño experimental reduce (o minimiza) la matriz de varianzas covarianzas del modelo logit condicional esperado. En general, esto puede realizarse desde una perspectiva determinista o bayesiana, dependiendo de los supuestos que hagamos sobre los parámetros previos, y utilizando diferentes algoritmos de optimización que deben ser programados y automatizados para que su uso sea viable.

Con respecto a las aplicaciones prácticas, se repasaron algunos conceptos de presentación de los choice sets a los encuestados, pilotaje de la encuesta, análisis de los resultados y disposición a pagar. También, posteriormente, se presentaron los resultados e implicaciones del empleo de experimentos de elección discreta seriales y algunas alternativas de formulación.

Como se ha podido comprobar hasta ahora, el diseño e implementación de los experimentos de elección discreta es complejo y requiere el empleo de software especializado. Además, las técnicas de diseño eficiente (u óptimo) han añadido la necesidad de realizar uno o varios experimentos pilotos, que han incrementado la farragosidad del procedimiento. La realización de un solo piloto requiere que los investigadores generen un DCE, lo pasen a un número determinado de encuestados, transcriban los datos, los codifiquen, estimen los modelos necesarios, generen un nuevo DCE basado en los anteriores coeficientes y repitan el procedimiento. Cuando el método requiere la actualización constante del experimento, llevar el DCE de forma manual se vuelve imposible.

Para aliviar la carga que el empleo de un DCE puede suponer a un equipo de investigadores, a continuación, se presenta el paquete de R “DCEtool”. “DCEtool” combina todas las herramientas empleadas en un DCE bajo la misma interfaz visual, permitiendo el diseño y testeo del DCE en minutos. En las siguientes secciones se introducirán los detalles de DCEtool. Estos detalles se corresponden con la última versión de la aplicación, la cual ha ido incorporando mejoras a lo largo

del tiempo. Estas mejoras surgen a partir de los numerosos correos recibidos de investigadores de todas partes del mundo que se han interesado por DCEtool. También, por supuesto, del Taller EEconAES y los comentarios de Miguel Ángel Negrín Hernández.

5.2 DETALLES TÉCNICOS

5.2.1 Algunos detalles

Para realizar un DCE existen dos requisitos que son indispensables. El primer requisito consiste en conocer qué hay que hacer para generar un diseño experimental adecuado, cómo presentarlo a las personas encuestadas y cómo analizar sus resultados. Para cumplir este requisito, es necesario tener conocimientos teóricos sobre la materia. El segundo requisito consiste en realizar el esfuerzo necesario para llevar la teoría a la práctica. En este capítulo se presenta DCEtool, una API que ofrece a la comunidad investigadora asistencia para cumplir estos dos requisitos de forma gratuita con software libre multiplataforma. En resumidas cuentas, una persona interesada en crear un DCE, pero sin alguno de estos conocimientos, podrá realizarlo de forma rápida y correcta, sin necesidad de pasar horas estudiando la teoría o descifrando la sintaxis de algunas funciones muy complejas de R (más sobre las implicaciones de esto después).

Al igual que los requisitos expuestos anteriormente, DCEtool cumple estos a través de su código *backend* y su interfaz *frontend*. Su código *backend* consiste en una serie de funciones en R enlazadas para crear, presentar y analizar un DCE (primer requisito). Su código *frontend* consiste en una interfaz visual en JavaScript que restringe lo que el usuario puede hacer con el código *backend*, solo permitiendo buenas prácticas de diseño, encuestación, y análisis, y restringiendo los conocimientos necesarios para llevar a cabo el procedimiento (segundo requisito).

Entrando en detalles técnicos, para aquellos lectores interesados y familiarizados con R y algunas de sus funciones, DCEtool recopila los inputs del usuario a través de una interfaz programada con el paquete 'shiny' (W. Chang, 2022) y los introduce en una función propia para crear matrices de diseño basadas en priors fijos o bayesianos (los inputs que se solicitan pueden observarse en la sección 5.3). Los inputs y la matriz de diseño generada pueden ser importados y exportados de un archivo Excel generado con 'writexl' (Ooms & McNamara, 2021) y leído con 'readxl' (Wickham et al., 2022). A continuación, el usuario puede introducir los nombres de los atributos y niveles, y también de las alternativas si desea crear una encuesta local. La encuesta local se genera a través de una función propia, programada para DCEtool, pero inspirada en una función similar del paquete 'idexf' (Traets et al., 2020), que actualiza sus priors (o no) de acuerdo con la teoría vista con anterioridad y la selección del usuario. Finalmente, los resultados de la encuesta pueden ser analizados usando el modelo logit condicional del paquete 'survival' (Therneau & Lumley, 2022) o

el logit mixto del paquete 'mlogit' (Croissant, 2020). Posteriormente, se desglosarán con mayor detalle las funcionalidades de DCEtool.

DCEtool se crea no tanto para asistir teóricamente el diseño de experimentos de elección discreta, sino para agilizar el proceso a través de software libre y accesible a cualquier persona en cualquier país del mundo. En la actualidad, DCEtool permite generar diseños D-óptimos y D-eficientes (bayesianos), decodificar la matriz de diseño, presentar la encuesta de forma cíclica y registrar las respuestas del usuario, incluso aplicando el método serial y el serial alternativo propuesto en el capítulo anterior, y analizar los resultados a través de un logit condicional y logit mixto estimando su disposición a pagar y generando una representación gráfica de los coeficientes.

La primera versión de DCEtool, versión 0.0.1, se subió al repositorio oficial de R, CRAN, en abril de 2021. La última versión de DCEtool, la versión 1.0.0, la cual incorpora mejoras sustanciales en funcionalidades y estabilidad, se subió a CRAN en abril de 2022. DCEtool comenzó como un proyecto de aprendizaje de programación en R usando shiny, y en la actualidad es una herramienta robusta para generar y testear experimentos de elección discreta.

Desde su lanzamiento en el repositorio oficial y la publicación de diversos preprints con las instrucciones para su uso, numerosos usuarios se han interesado por el y han escrito emails pidiendo asistencia. Las actualizaciones de DCEtool, especialmente la 1.0.0, se ha realizado pensando en solucionar muchos de los problemas planteados en estos emails. La actualización 1.0.0 también surge tras las propuestas que se realizan tras la exposición realizada en el Taller EconAES de la Asociación de Economía de la Salud en Madrid en marzo de 2022.

Como puede observarse en la Figura 5.1, desde su lanzamiento DCEtool ha sido descargado más de 5000 ocasiones a través del repositorio oficial de R, CRAN. En el Apéndice 5 pueden encontrarse todos los enlaces de relevancia para DCEtool.

5.2.1 Requisitos

DCEtool requiere tanto R (R Core Team, 2021) como RStudio (RStudio Team, 2021) y ha sido probado en Windows (10 y 11) y Ubuntu (20.04).

5.2.2 Instalación

DCEtool está disponible en el repositorio oficial de R, CRAN, por lo que su instalación puede realizarse desde la propia interfaz de RStudio usando el código de la línea 1.

```
1. install.packages("DCEtool")
```

La interfaz gráfica se desplegará tras ejecutar las líneas de código 2 y 3.

2. `library("DCEtool")`
3. `DCEtool()`

Alternativamente, el paquete 'devtools' puede emplearse para instalar DCEtool directamente desde GitHub. Es probable que, tras desarrollar una actualización, GitHub sea el primer repositorio al que se suba esta por los restrictivos requisitos de CRAN. Para instalar desde GitHub hay que ejecutar las líneas 4,5, y 6:

4. `install.library("devtools")`
5. `library("devtools")`
6. `install_github("danielpereztr/DCEtool")`

Interfaz

- Design settings: en esta pestaña se deciden las características que tendrá el experimento de elección discreta. El usuario puede introducir los atributos, niveles, número de alternativas, número de choice sets, una alternativa nula, el método de optimización, una semilla aleatoria, y los priors del diseño.
- Design matrix: en esta sección se computará un diseño eficiente una vez se hayan introducido todos los parámetros en la sección anterior. El usuario podrá descargar la matriz de diseño en formato Excel, cambiar los nombres de los atributos y los niveles y ver las características del diseño.

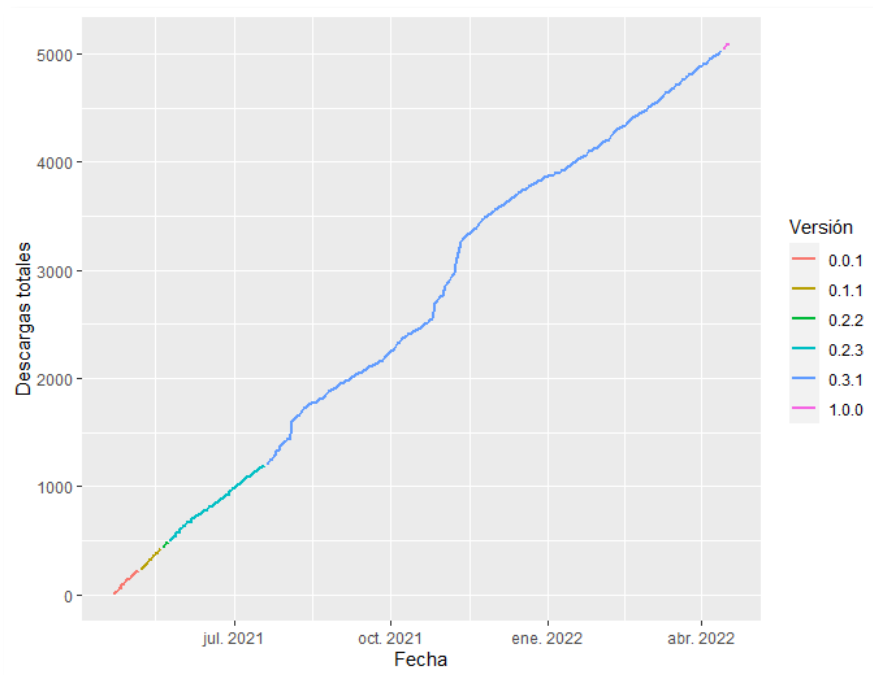


Figura 5.1. Descargas de DCEtool desde su lanzamiento según la versión

- Create a survey: en esta sección el usuario puede configurar una encuesta local usando el DCE que acaba de crear. Para ello, puede introducir un nombre para cada alternativa, crear un texto introductorio y final, y decidir si usar o no la aproximación secuencial de Bliemer y Rose (2010a) o la alternativa secuencial cada cinco encuestados en Pérez-Troncoso (2022).
- Survey: en esta sección la encuesta se presentará de forma cíclica, empezando por el texto introductorio y terminando por el texto final. El usuario puede usar esta sección para llevar a cabo un experimento de elección discreta con diferentes encuestados.
- Results: el usuario puede analizar todas las respuestas anteriormente recopiladas a través de un logit condicional y un logit mixto. El atributo precio puede ser codificado de forma lineal para estimar la disposición a pagar automáticamente y los coeficientes pueden ser representados gráficamente. El programa también permite guardar los resultados en formato Excel o cargar los resultados de otro DCE.

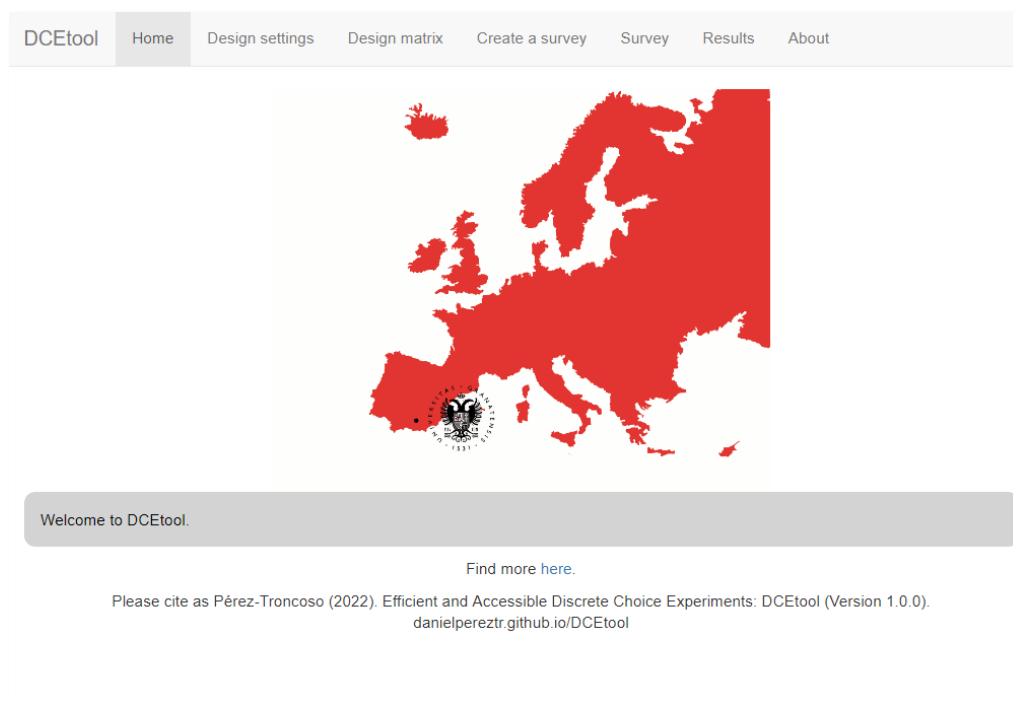


Figura 5.2. Interfaz principal de DCEtool

5.3 GENERACIÓN DE UN DCE

En este apartado se generará, se testeará y se analizará un experimento de elección discreta empleando DCEtool. Para ello, todos los pasos serán pormenorizados, para que el experimento pueda ser reproducido.

Configuración del experimento

Para comenzar, en la pestaña ‘Design settings’ debemos especificar las características del diseño experimental. En esta sección emplearemos la misma configuración que en Pérez-Troncoso et al. (2021): cuatro atributos, dos de ellos con dos niveles y otros dos con cuatro, y dos alternativas sin alternativa nula. En este caso, en lugar de emplear dieciséis choice sets como en el experimento original, usaremos doce. El método de optimización será bayesiano, los priors serán todos cero y la semilla aleatoria será 1234.

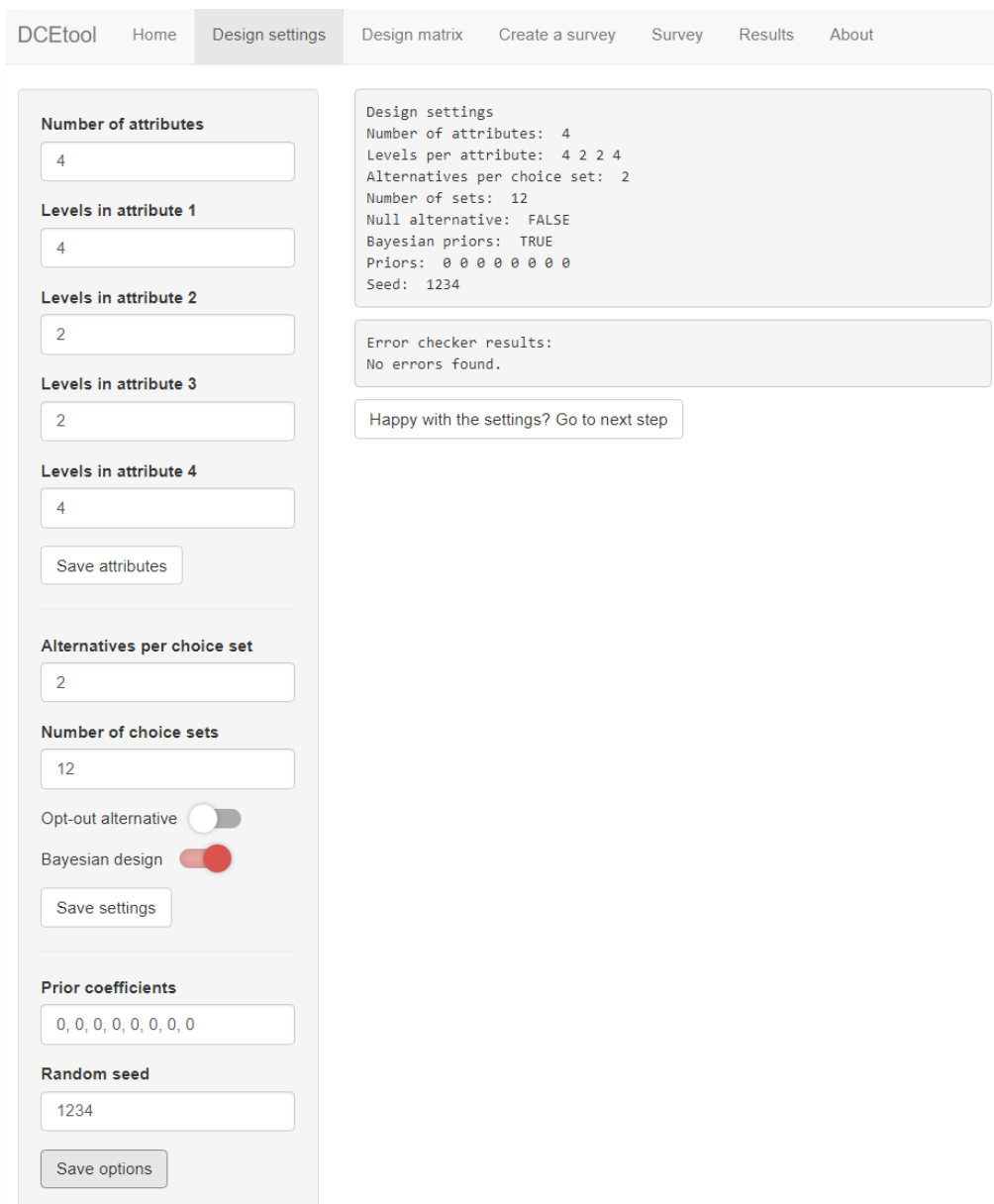


Figura 5.3 Configuración del experimento de elección discreta

En el siguiente apartado, ‘Design matrix’ se pueden elegir los nombres de los atributos y niveles del experimento. En este caso seguiremos empleando el ejemplo del DCE de nutrición personalizada. Como puede observarse en la Figura 5.4, cuando los nombres de las variables de la matriz de diseño incluyan los nombres de algunos de los niveles es recomendable guardar el

diseño experimental. DCEtool utiliza archivos Excel tanto para importar como para exportar datos. Una vez guardado este diseño experimental, el usuario podría cerrar y volver a abrir el programa y cargar este diseño directamente desde esta pestaña.

Generate design
Save design
Browse... No file selected
Show design details
Name the attributes

Name of attribute 1
Personalización

Name of attribute 2
Formato

Name of attribute 3
Rutinas deportivas

Name of attribute 4
Precio

Save names
Change level names

Show 10 entries

	task	alt	Nivel 2	Nivel 3	Nivel 4	Digital	Si	20 euros	30 euros	40 euros
1	1	1	0	1	0	0	1	1	0	0
2	1	2	0	0	0	1	0	0	0	0
3	2	1	0	0	0	0	1	0	1	0
4	2	2	0	1	0	1	0	0	0	0
5	3	1	0	0	1	0	0	0	1	0
6	3	2	0	0	0	1	1	0	0	0
7	4	1	0	1	0	0	0	0	1	0
8	4	2	0	0	1	0	1	1	0	0
9	5	1	0	1	0	1	1	0	0	1
10	5	2	1	0	0	0	1	0	0	0

Showing 1 to 10 of 24 entries

Number of attributes: 4
Levels per attribute: 4 2 2 4
Alternatives per choice set: 2
Number of sets: 12
Null alternative: FALSE
Bayesian priors: TRUE
Priors: 0 0 0 0 0 0 0
Attribute names: Personalización Formato Rutinas deportivas Precio
Level names: Nivel 1 Nivel 2 Nivel 3 Nivel 4 No Sí Tradicional Digital 10 euros 20 euros 30 euros 40 euros
Details: Coordinate exchange algorithm used to generate an efficient Bayesian design

Figura 5.4 Nombre de los atributos y niveles en la matriz de diseño

En el siguiente apartado, 'Create a survey', se pueden especificar las visuales de la encuesta: la introducción, el texto final, y cuál es el aspecto de un choice set. Tanto el texto introductorio como el final deben ser redactados en lenguaje Markdown. Esto permite la inclusión de texto de diferentes tamaños, tablas, enlaces, e incluso imágenes (aunque el espacio disponible es limitado) (Figura 5.5).

A continuación, en el apartado 'survey', el primer paso será elegir qué modo de encuestación va a emplearse. Al elegir el método de Bliemer & Rose, los priors serán reestimados cada vez que un encuestado responda a todos los choice sets (en este caso 12). Al elegir el método serial de cada cinco encuestados, los priors serán reestimados cada cinco encuestados. Al elegir 'No', la misma encuesta (la representada en la matriz de diseño que antes hemos guardado) se repetirá de forma cíclica (Figura 5.6).

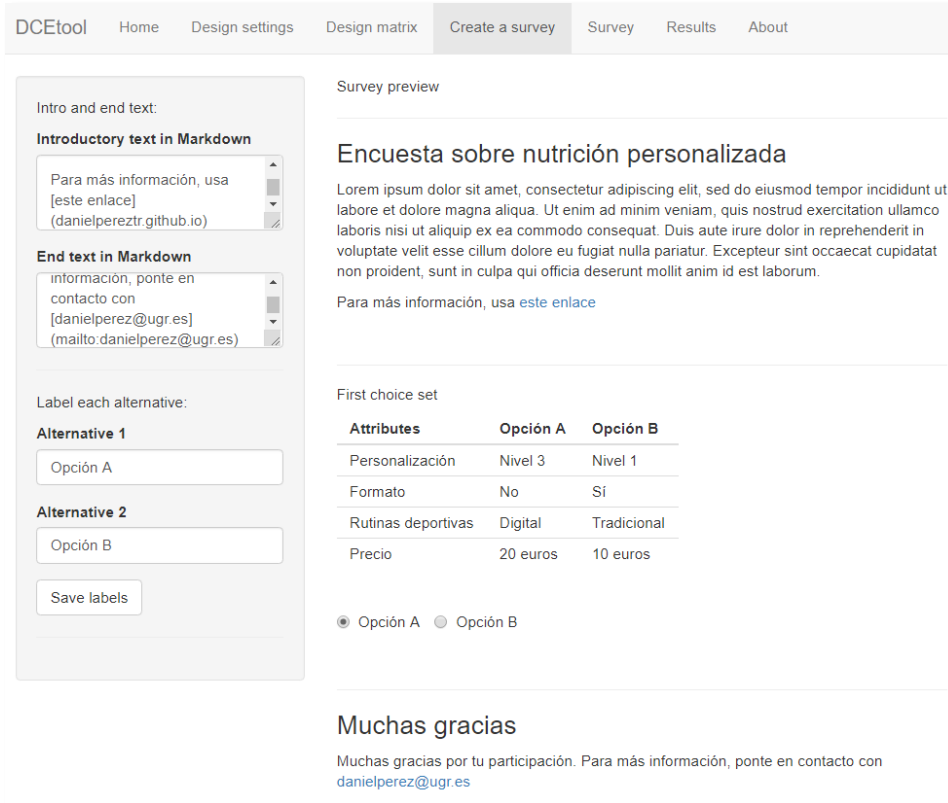


Figura 5.5 Características de la encuesta

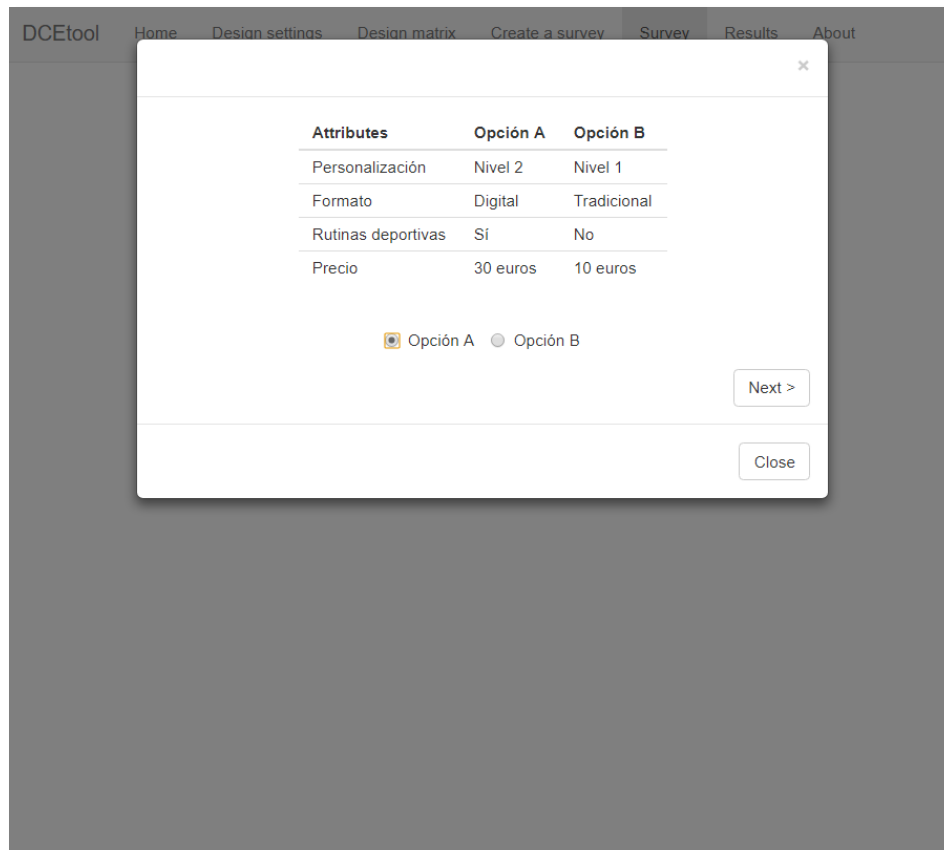


Figura 5.6 Encuesta

Una vez se hayan registrado las respuestas que el usuario desee, mientras se muestra el texto final en la encuesta, el encuestado debe pulsar sobre 'Analyze the results' y guardar una copia de los resultados.

Price variable coding: Code price as continuous variable

Select price levels:

- pid
- task
- alt
- Nivel 2
- Nivel 3
- Nivel 4
- Digital
- Si
- 20 euros
- 30 euros
- 40 euros
- choice
- gid

	pid	task	alt	Nivel 2	Nivel 3	Nivel 4	Digital	Si	20 euros
1	1	1	1	0	0	0	0	0	0
2	1	1	2	0	1	0	1	0	1
9	1	2	1	0	0	0	1	0	0
10	1	2	2	0	0	1	0	0	0
11	1	3	1	1	0	0	1	1	0
12	1	3	2	0	0	0	0	0	0
13	1	4	1	0	0	1	1	0	0
14	1	4	2	0	0	0	0	1	0
15	1	5	1	0	0	0	1	1	1
16	1	5	2	1	0	0	0	0	0

Metodología para el cálculo de la DAP como estimador del valor monetario de una mejora en salud

Figura 5.7. Codificando el atributo precio como variable continua

La tabla que se muestra contiene tanto la matriz de diseño como las respuestas proporcionadas, todo codificado para ser analizado. No obstante, como se vio en los temas anteriores, para estimar la disposición a pagar es necesario establecer el atributo precio como una variable continua. Para ello, en el desplegable 'Price variable coding' seleccionaremos 'Code price as continuous variable'

y seleccionaremos todos las variables que se refieran a niveles del atributo precio. A continuación, los valores de los precios deberán ser proporcionados como números enteros (Figura 5.7).

Tras pulsar sobre 'Add variable to the data frame', una nueva variable, 'cont_price', debe aparecer en la tabla. A continuación, podemos proceder a realizar las estimaciones pertinentes usando la pestaña 'Estimation' del menú lateral izquierdo de la pestaña 'Results'.

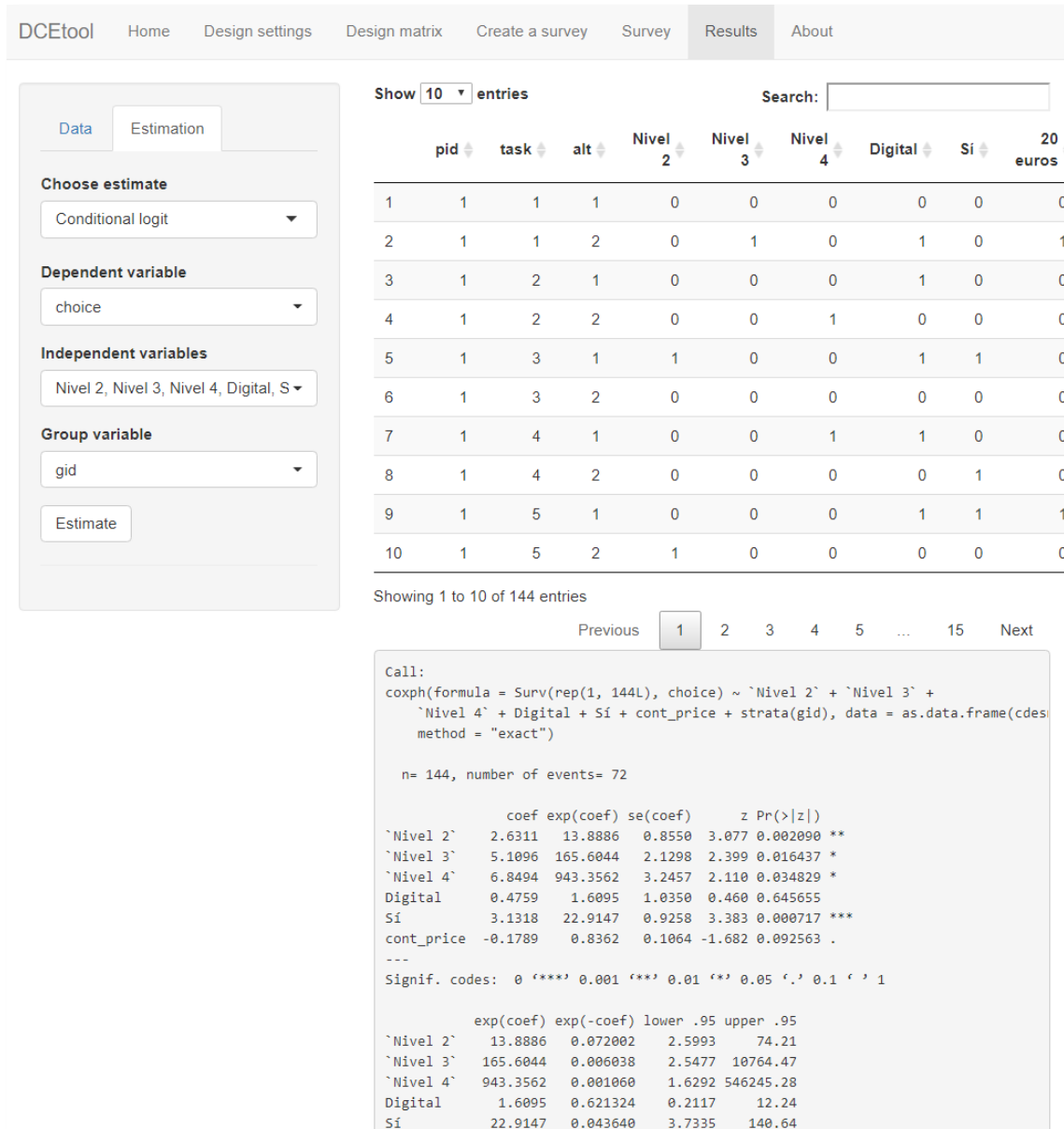


Figura 5.8. Logit condicional

En primer lugar, emplearemos una regresión logística condicional. Para ello, en el menú desplegable seleccionaremos 'Conditional logit', la variable dependiente será 'choice', las variables independientes serán todos los niveles (menos los niveles de precio, que deberán ser sustituidos por la variable continua recién creada), y la variable de agrupación será la variable 'gid' (tanto 'choice' como 'gid' son nombres predeterminados por DCEtool para estas variables). Tras pulsar

sobre 'Estimate', el programa debe ofrecer una salida con los resultados de la regresión (Figura 5.8). Los resultados provienen de seis respuestas registradas por el autor durante la realización de este capítulo.

De la misma manera, puede estimarse el logit mixto. El logit mixto requiere también que se especifique una variable identificadora de cada individuo ('pid' por defecto) y una variable identificadora de cada alternativa ('alt' por defecto).

Por último, DCEtool nos permite estimar tanto la disposición a pagar como representar gráficamente los resultados. Para ello, en primer lugar debemos estimar un modelo condicional logit y después acceder a cada uno de los apartados: 'Willingness to pay' (Figura 5.9) y Figures (Figura 5.10).

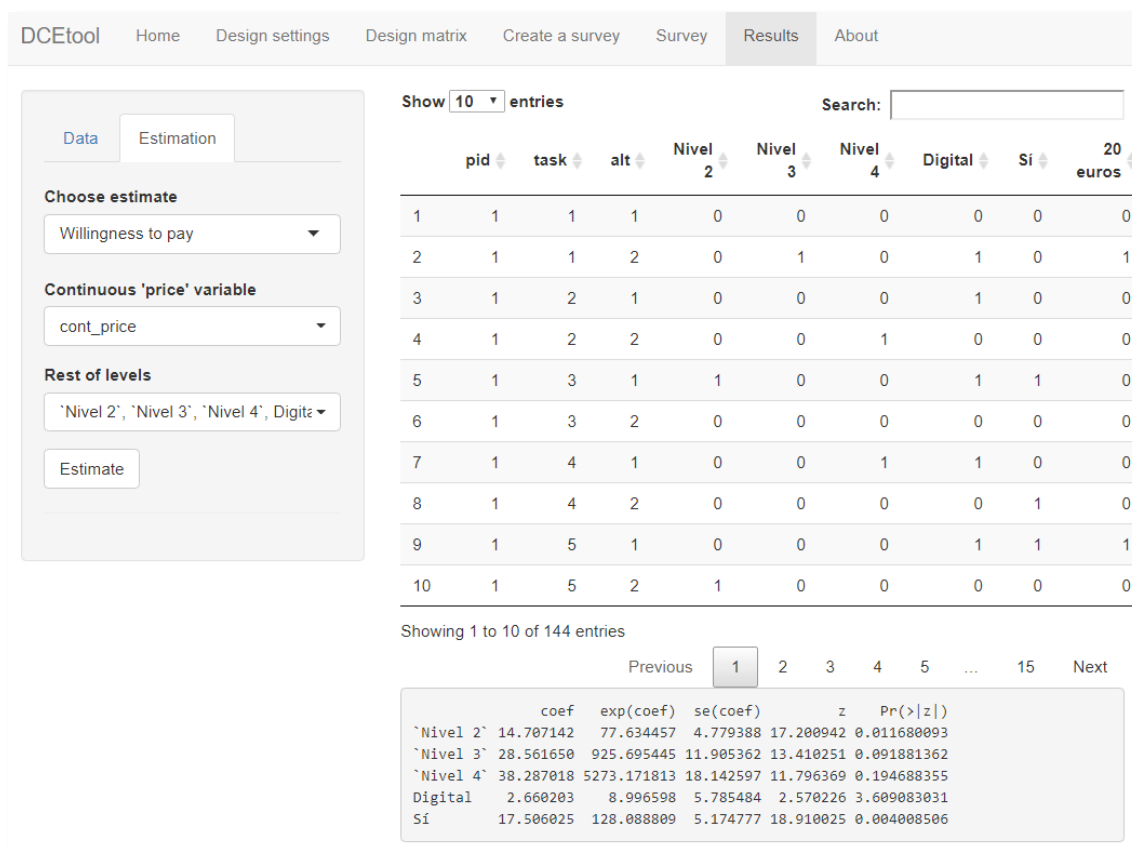


Figure 5.9. Cálculo de la disposición a pagar



Figura 5.10 Representación gráfica de los resultados

5.4 VALIDEZ INTERNA

Una de las ventajas de DCEtool es que permite que la persona que esté diseñando el DCE compruebe su funcionamiento directamente en la aplicación. Para comprobar cómo se comporta DCEtool bajo estas circunstancias, se ha reproducido el DCE en Pérez-Troncoso et al. (2021) y se ha respondido trece veces en cada una de las modalidades de encuesta disponibles en DCEtool: no serial, serial según Bliemer & Rose, y serial cada cinco encuestados. Aunque el cometido principal de este ejercicio es poner a prueba a DCEtool, también será interesante realizar una comparación de los resultados pues todas las encuestas serán respondidas bajo las mismas circunstancias (por la misma persona, con los mismos sesgos y en la misma interfaz).

	Simple		Serial Bliemer & Rose		Serial (n = +5)	
	Coef.	SE (p-valor)	Coef.	SE (p-valor)	Coef.	SE (p-valor)
Pers: N1	(Coeficiente restringido a cero; nivel base)					
Pers: N2	1.3344	0.4735 (0.005)	2.3654	0.5695 (0.000)	4.1463	1.0909 (0.000)
Pers: N3	3.8683	0.7066 (0.000)	3.6942	0.6928 (0.000)	6.7799	1.5180 (0.000)
Pers: N4	3.9406	0.8228 (0.000)	5.4487	1.0546 (0.000)	9.3736	2.2040 (0.000)
F: Trad.	(Coeficiente restringido a cero; nivel base)					
F: Digital	0.3402	0.8228 (0.371)	-0.1192	0.2657 (0.654)	-0.8919	0.5972 (0.135)
R. De: No	(Coeficiente restringido a cero; nivel base)					
R. De: Sí	0.2549	0.4118 (0.536)	1.3373	0.4149 (0.001)	1.2149	0.6588 (0.065)
Precio	-0.0742	0.0223 (0.001)	-0.1083	0.0226 (0.000)	-0.2209	0.0545 (0.000)
	$p - R^2 = 0.6076$ Log-ver.: -42.42		$p - R^2 = 0.5074$ Log-ver.: -52.92		$p - R^2 = 0.5779$ Log-ver.: -45.64	

Tabla 5.1. Representación gráfica de los resultados

En la Tabla 5.1 pueden observarse los resultados de las regresiones logísticas condicionales de cada una de las modalidades de DCE. Dos hechos saltan a la vista con respecto a los coeficientes de los modelos: la escala y los signos. Puede observarse que existen diferencias de escala entre los coeficientes de los diferentes modelos, aunque estas no deben percibirse como una prueba de que una estimación sea mejor que otra. Solo las diferencias de utilidad son relevantes en los modelos de elección discreta (Train, 2009). Por otro lado, el signo de los coeficientes de los modelos se corresponde para todos aquellos coeficientes que son significativos.

Con respecto a la significatividad de los coeficientes, observamos que, por norma general, los coeficientes de los modelos seriales son más significativos y tienen un error estándar menor. Las medidas de ajuste, por otro lado, muestran la imagen contraria, pues es el modelo no serial el que consigue un pseudo R^2 mayor. Muy probablemente esto no sea resultado de un mejor ajuste sino de la existencia de sobreajuste (en mayor medida que en los otros modelos, recordando que McFadden (1974) determinó que un pseudo R^2 entre 0.2-0.4 era indicador de un buen ajuste).

Para comparar con mejor perspectiva las diferencias relativas entre utilidades (coeficientes) de los diferentes modelos, se estimaron las disposiciones a pagar¹⁷ por cada uno de los niveles en cada uno de los modelos. Para incluir en la estimación los intervalos de confianza de las DAP se ha empleado el paquete 'wtp' de Stata (Hole, 2007) el cual usa el método Delta para la estimación de los intervalos.

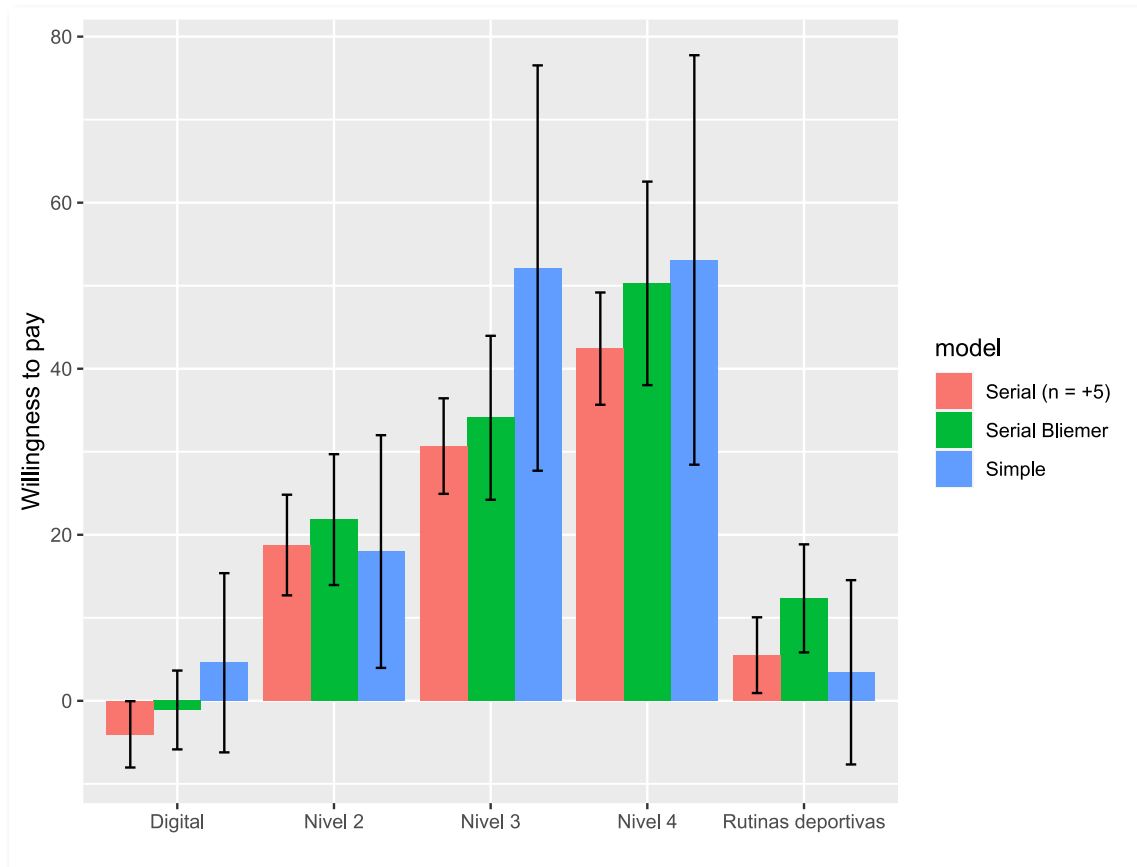


Figura 5.11 Representación gráfica de los resultados en DAP de cada método

Como puede observarse en la Figura 5.11, todas las regresiones obtienen resultados de DAP muy similares con algunas peculiaridades. El método de encuestación simple obtiene unas disposiciones a pagar muy altas para los niveles de personalización 3 y 4. Además, para todas las estimaciones obtenidas a través de la regresión de la encuesta simple los IC 95% son mayores, lo que denota mayor incertidumbre entorno a la estimación. Por ejemplo, usando la encuesta simple no podría rechazarse la hipótesis nula de igualdad entre el nivel 2 y 4 de personalización, mientras que hay diferencias significativas cuando se emplean las otras técnicas. Con respecto al

¹⁷ La estimación de la disposición a pagar implica la normalización de la utilidad a euros y permite comparar, globalmente, los coeficientes entre modelos.

resto, tanto la opción serial pura (Bliemer & Rose) como la alternativa obtienen estimaciones bastante precisas y similares.

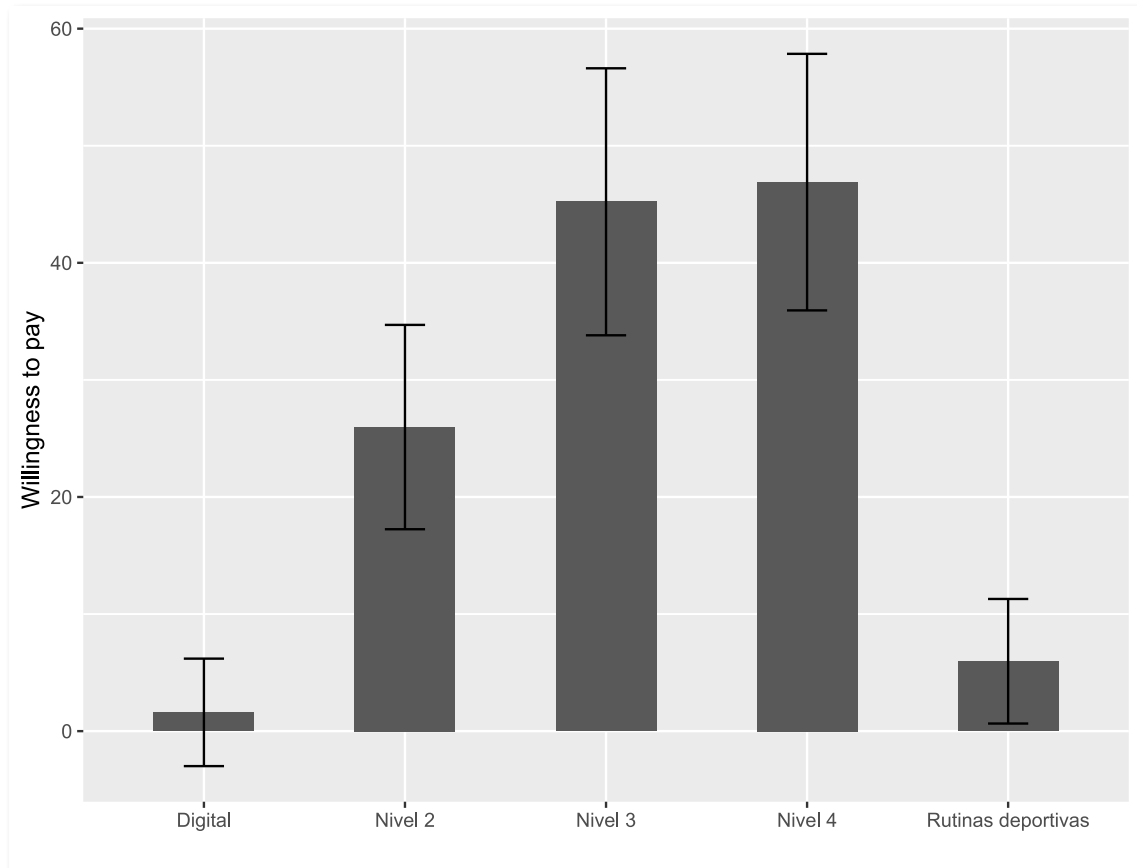


Figura 5.12 Representación gráfica de los resultados en DAP totales

En la Figura 5.12 se presentan los resultados en disposición a pagar de la agregación de todas las encuestas (N = 39). La regresión logística condicional de la que provienen estos resultados obtuvo un $R^2 = 0.44$, lo que confirmaría que en los modelos de la Tabla 5.1 se estaría produciendo sobreajuste debido al tamaño de la base de datos. En esta situación, el método que obtiene un R^2 más cercano al rango deseado de 0.2-0.4 es, como cabe esperar, el secuencial puro.

5.5 CARGA COMPUTACIONAL

DCEtool emplea una función propia para generar los diseños experimentales D-eficientes o D-óptimos según la elección del usuario. Esta función, llamada internamente 'dce_toolbox' (que puede ser llamada como una función normal tras cargar el paquete DCEtool¹⁸) contiene tanto el algoritmo de Fedorov modificado como el algoritmo CEA que puede encontrarse en Traets et al.

¹⁸ Tras cargar el paquete DCEtool en R usando `library(DCEtool)`, la función 'dce_toolbox' puede ser llamada y un DCE puede ser generado utilizando los parámetros adecuados. Para más información, 'help(dce_toolbox)' en R.

(2020). Ambos algoritmos han sido reescritos¹⁹ para DCEtool. El primero de ellos, el algoritmo de Fedorov modificado, se emplea sólo para los diseños no bayesianos. El segundo algoritmo se emplea en los diseños bayesianos ya que es más rápido, pero también menos exhaustivo.

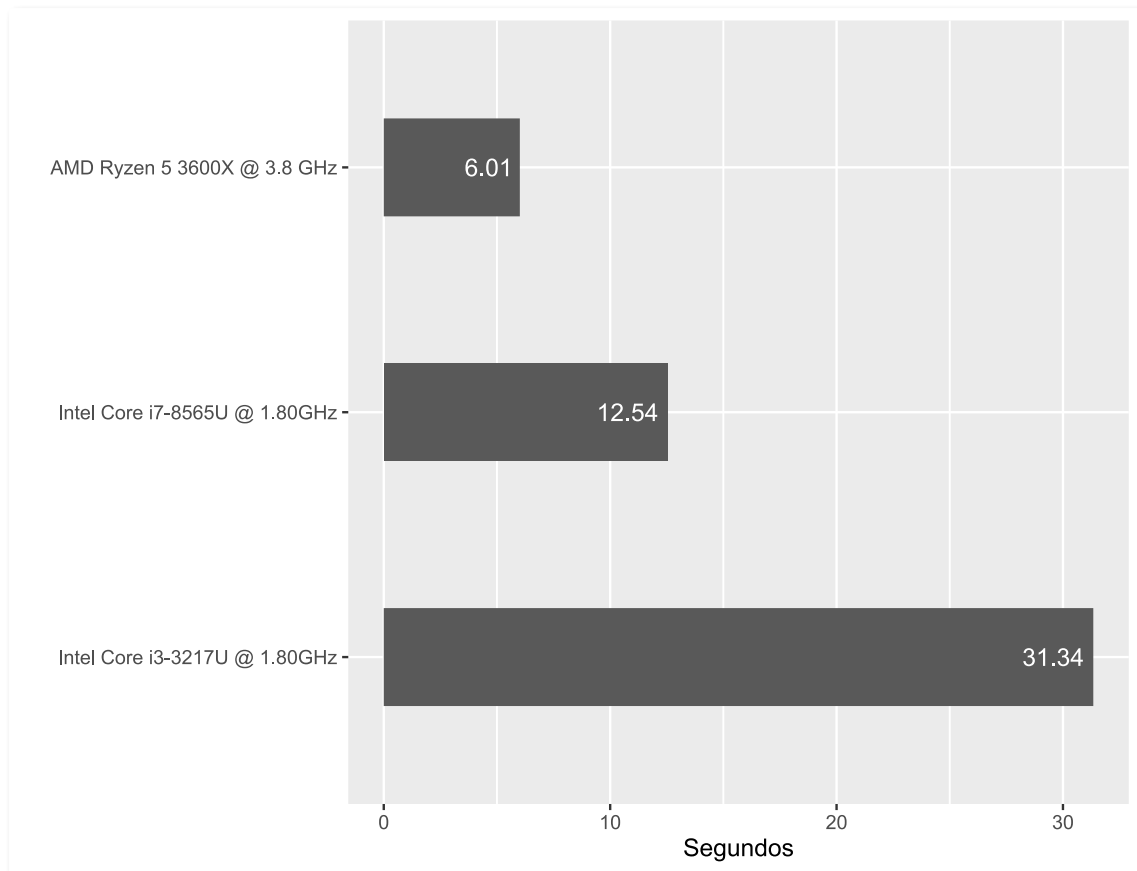


Figura 5.13 Representación gráfica de los resultados en DAP totales

En la Figura 5.3 se presentan los tiempos de computación de DCEtool para generar un diseño bayesiano $3^3 \cdot 2$ con ocho choice sets y dos alternativas por choice set. Como puede observarse, pese a que el diseño es relativamente pequeño, su cálculo puede conllevar más de diez segundos en sistemas modernos y más de 30 en sistemas relativamente antiguos. A diseños más grandes, es esperable que se incremente sustancialmente el tiempo de computación. Por otro lado, para diseños similares no bayesianos podemos esperar un tiempo de computación menor con el algoritmo de Fedorov, aunque este tiempo de computación se puede ver incrementado cuando se emplean muchos atributos y niveles.

¹⁹ El algoritmo de Fedorov fue reescrito en R gracias al apoyo que el profesor Arne Risa Hole me brindó a través de email. La función de cálculo del D-error en el algoritmo CEA fue extraída del paquete idfix de Traets et al. (2020) pues está programada en C++, un lenguaje de programación mucho más complejo pero mucho más rápido, algo que era necesario en esta ocasión.

5.6 CONCLUSIONES

En este capítulo se han abordado los orígenes, detalles y empleo de DCEtool en la generación de experimentos de elección discreta. Si bien DCEtool está pensado principalmente para aliviar la carga de trabajo de los diseñadores de experimentos de elección discreta, también tiene una labor pedagógica clara. En DCEtool, una persona recién llegada al mundo de los experimentos de elección discreta puede generar diseños experimentales con buenas propiedades sin tener conocimiento teórico. Si bien esto puede suponer el uso de estas herramientas sin los conocimientos teóricos básicos necesarios, la alternativa presente hoy en día podría ser aún peor. Lo cierto es que el autoaprendizaje de la materia es arduo y lleva a malentendidos. Existe un gran grado de confusión alrededor de los nombres, supuestos y métodos de optimización empleados en la literatura (Louviere et al., 2010). Tampoco ayuda que los artículos metodológicos más citados y famosos contenga nociones antiguas ya superadas (Bliemer & Rose, 2010a). Por este motivo, la existencia de una alternativa fácil de emplear y con buenas propiedades de diseño por defecto puede ayudar a muchas personas.

En próximas versiones de DCEtool, se incluirán las propuestas de algunos de los asistentes al curso metodológico de DCEs que se impartió en marzo en la Facultad de Económicas de la Universidad de Granada empleando R y DCEtool. También, gracias a las propuestas de mejora del director de esta tesis, David Epstein, se incluirá una pestaña que contendrá una guía de aprendizaje, con conceptos organizados y referencias útiles que guíen al usuario hacia las referencias correctas y la comprensión de la teoría, sin tener que pagar el caro peaje del aprendizaje a través de artículos antiguos y erróneos.

Capítulo 6. Conclusiones



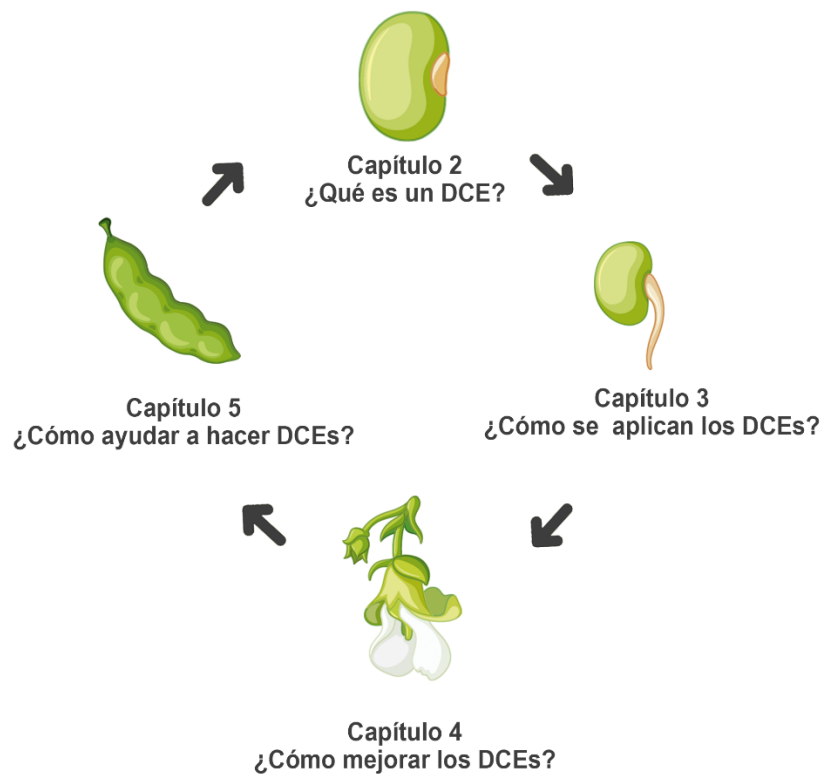
Este trabajo ha presentado los experimentos de elección discreta como una herramienta metodológica adecuada para la estimación de la disposición a pagar con el cometido de realizar un análisis coste-beneficios. Como se ha mencionado en la introducción, esta metodología es adecuada para realizar análisis en marcos alineados con el *bienestarismo* ('welfarism') clásico, donde la propia persona interesada es la que evalúa la utilidad en base a sus preferencias y realiza el desembolso. En Evaluación Económica de productos sanitarios, el empleo de DCEs es adecuado para casos como los que se exponen en el Capítulo 3, aunque su ámbito de actuación no queda restringido a este dominio, siempre que el empleo de esta metodología se encuentre debidamente justificado.

Adoptando una visión global del trabajo, este ha comprendido tres labores que se han considerado de relevancia para el campo de conocimiento. En primer lugar, en el Capítulo 2 se realizó una labor de compilación y desarrollo de los aspectos teóricos básicos de la materia, con un nivel de detalle suficiente como para permitir a cualquier interesado comprender la mayoría de aspectos teóricos de la materia. Esta labor de recopilación y translación surgió de la percepción de confusión y multiplicidad de fuentes existente en la materia, luego confirmada por otros autores, y permitió un proceso de aprendizaje y estructuración del conocimiento. Pese a que este trabajo termine aquí, el cometido de facilitar la comprensión y empleo de la metodología no terminará con esta tesis. Simultáneamente al desarrollo de este texto, se realizaron dos cursos metodológicos en la Facultad de Ciencias Económicas y Empresariales de la Universidad de Granada, y otro en el Taller de EEconAES de la Asociación de Economía de la Salud, además de la publicación de otros textos y tutoriales online (también en relación con el contenido del Capítulo 5). Este trabajo de divulgación pretende ser continuado en la medida de lo posible de aquí en adelante.

Por otro lado, el Capítulo 3 presenta una aplicación práctica de la metodología en el ámbito de la Evaluación Económica de tecnologías relacionadas con la salud. Esta publicación tiene especial relevancia pues fue la primera experimentación real y primera publicación científica de esta tesis, la cual ocasionó nuevos conocimientos prácticos fundamentales para el empleo de la metodología. Este trabajo no hubiera sido posible sin el apoyo y los conocimientos del director de esta tesis, David Epstein, y el profesor José Alberto Castañeda.

El Capítulo 4 supuso un gran reto, y constituye una de las partes del trabajo que más tiempo han supuesto, pero que a la vez han ofrecido mayor recompensa. Esta investigación partió del hallazgo fortuito de una investigación donde los profesores John Rose y Michiel Bliemer, de la Universidad de Sídney, proponen realizar DCEs de forma secuencial y sugieren que puede haber una forma de incrementar la eficiencia del proceso secuencial originalmente planteado. Durante el verano de 2020, comenzó el trabajo que más tarde se materializaría en la publicación condensada en el

Capítulo 4. Al principio, este trabajo consistió en programar en R un bucle que simulara la propuesta original de Bliemer y Rose, y después el código se modificó para aceptar variaciones de esta propuesta, tratando de responder a la sugerencia que los autores dejaron escrita en su trabajo. Debido a que el bucle usaba un planteamiento Monte Carlo, cada simulación debía ser repetida 100 veces, incrementando así el tiempo de computación de forma desmedida en algunas estrategias. En noviembre de 2020, un primer borrador de este trabajo fue enviado a Journal of Choice Modelling, revista de la que es editor el autor del libro donde se publicó el trabajo de Bliemer y Rose, Stephane Hess, quien acogió el ensayo, aunque aún quedaba mucho trabajo por hacer. Tras dos años de revisiones, donde los revisores anónimos -quizás alguno de los autores de la propuesta original- fueron de gran ayuda, el trabajo fue aceptado para publicación tras la inclusión de nuevas simulaciones y bastantes horas de mejora del inglés.



El Capítulo 5 surge tras el esfuerzo de programación en R para el Capítulo 4, y el descubrimiento del paquete de R 'shiny' gracias a la propuesta de David Epstein de desarrollar la herramienta ICERtool (<https://icertool.shinyapps.io/ICERtool/>) en base a su código para representar la frontera de eficiencia en un plano de coste-efectividad (doi.org/ 10.17632/7g77npfpg.2). Ambos hechos, unidos a la dificultad que se observó en uno de los cursos de DCEs de los asistentes para tener una comprensión global de la materia y llevar a cabo los procedimientos necesarios para generar un DCE, se decidió programar DCEtool. Desde su creación, DCEtool ha ido mejorando poco a poco, aunque con las limitaciones de solo tener una persona a su cargo que no se dedica a la

programación de forma profesional. Pese a ello, en la medida de lo posible, DCEtool irá incorporando mejoras y pretende convertirse en la alternativa gratuita más sencilla y adecuada actualmente disponible, teniendo en cuenta que no existen muchas (o ninguna) otras opciones con sus características.

Pese a que los DCEs son más populares cada día, aún existe mucho desconocimiento y la Economía de la Salud debería hacer mayor uso de ellos en el futuro. Como hemos visto, estos son ideales para comprender las preferencias de los pacientes, pero también para calcular las ponderaciones y trade-offs de los decisores, tarea para la cual en muchas ocasiones se emplean otras metodologías con menos base teórica que los DCEs. Igualmente, podrían emplearse para realizar análisis coste-beneficios de intervenciones públicas en salud como los aumentos de impuestos a las bebidas azucaradas o los programas de promoción de la actividad física.

Esta tesis ha abarcado aquellos aspectos con los que se ha trabajado durante su periodo de redacción, aunque quedan otros de relevancia que no serán tratados. No obstante, se emplaza al lector a seguir de cerca los trabajos que Milad Haghani, Michiel Bliemer, John Rose, Emily Lancsar y compañía, quienes actualmente tienen dos *preprints* en arXiv sobre el sesgo hipotético en DCEs, que pronto podrán leerse posiblemente en Journal of Choice Modelling. Por otro lado, también recomendar la publicación de junio de 2022 de Daniel McFadden, igualmente en Journal of Choice Modelling, sobre la inestabilidad del modelo logit mixto para estimar demandas.

Para finalizar, plantear una futura línea de investigación y un deseo. Por un lado, la futura línea de investigación simulará la realización de un DCE online cuyos choice sets se actualizarán, usando como priors las respuestas ya recopiladas, cada 10, 30, 60 o 90 minutos, y hallará la periodicidad de actualización óptima dada una distribución de las respuestas en el tiempo. Por otro lado, el deseo es que este trabajo, o parte de él, sea alguna vez de utilidad para cualquier persona que un día se interese por esta metodología.

AGRADECIMIENTOS

Gracias a David Epstein por confiar en mi durante todos estos años y darme la oportunidad de hacer esta tesis. Sin su confianza y apoyo habría sido imposible aprender todo lo que he aprendido estos años, participar en proyectos tan interesantes y diversos, y tener tiempo para estudiar e investigar lo que hoy se encuentra en este documento.

Gracias a mis padres, quienes tanto cariño y apoyo me han dado desde que me fui de casa para perseguir esta meta. Sin unos padres como ellos, nunca hubiera llegado hasta aquí. Gracias a mi abuelo, quien siempre ha sido una inspiración y está muy orgulloso de que vaya a convertirme en doctor.

Gracias a Javi, Laura, Jesús, Adri, Fran, Edu, Alex, Dani, Belén, Norberto, Gonzalo, Fernando, Natalia, Javi C y, por supuesto, Rita, por aguantarme siempre pese a que a veces puedo ser la persona más pesada del planeta.

Gracias al grupo de investigación Economía Pública y Globalización (EPIC), al Departamento de Economía Aplicada y tod@s l@s compañer@s que me han acogido con tanto cariño desde que llegué. Gracias a mis colegas de la facultad Juanda y Modou, quienes me han guiado innumerables veces desde que aterricé aquí.

Gracias a los miembros del tribunal que tan amablemente han aceptado participar en la defensa de esta tesis.

Gracias al grupo EEconAES de la Asociación de Economía de la Salud por su ayuda y necesaria labor ayudando e integrando a nuev@s investigador@s.

Gracias al proyecto PAPAartis, Stance4Health y al Imperial College of Science Technology and Medicine por la financiación de este proyecto de investigación.

Esta ha sido una época apasionante que será determinante para el resto de mi vida. Estoy muy agradecido.

FINANCIACIÓN

Durante el desarrollo de esta tesis, se ha recibido financiación de las siguientes fuentes:

- Contrato de investigación con cargo al proyecto PAPA-artis (2019-2022)
- Contrato de investigación con cargo al proyecto NESIC (Imperial College) (2022)
- Contrato de investigación con cargo a grupo (SEJ393) (UGR) (2022)
- Financiación de la investigación en el Capítulo 3 por el proyecto Stance4Health (2021)
- Taller EEconAES, curso *Experimentos de Elección Discreta* para la ev. econ. (2021)

REFERENCIAS

- Addelman, S. (1962). Orthogonal Main-Effect Plans for Asymmetrical Factorial Experiments. *Technometrics*, 4(1), 21. <https://doi.org/10.2307/1266170>
- Ballco, P., & de Magistris, T. (2019). Spanish consumer purchase behaviour and stated preferences for yoghurts with nutritional and health claims. *Nutrients*, 11(11). <https://doi.org/10.3390/nu11112742>
- Bansback, N., Brazier, J., Tsuchiya, A., & Anis, A. (2012). Using a discrete choice experiment to estimate health state utility values. *Journal of Health Economics*, 31(1), 306–318. <https://doi.org/10.1016/J.JHEALECO.2011.11.004>
- Bliemer, M. C. J., & Rose, J. M. (2010a). Serial Choice Conjoint Analysis for Estimating Discrete Choice Models. In *Choice Modelling: The State-of-the-art and The State-of-practice* (pp. 137–161). Emerald Group Publishing Limited. <https://doi.org/10.1108/9781849507738-006>
- Bliemer, M. C. J., & Rose, J. M. (2010b). Serial Choice Conjoint Analysis for Estimating Discrete Choice Models. In *Choice Modelling: The State-of-the-art and The State-of-practice* (pp. 137–161). Emerald Group Publishing Limited. <https://doi.org/10.1108/9781849507738-006>
- Blomquist, G. C. (2015). Value of Life, Economics of. *International Encyclopedia of the Social & Behavioral Sciences: Second Edition*, 14–20. <https://doi.org/10.1016/B978-0-08-097086-8.71016-6>
- Boardman, A. E., Greenberg, D. H., Vining, A. R., & Weimer, D. L. (2012). *Cost-benefit analysis: concepts and practice* (Cambridge University Press, Ed.). <http://testbankmango.eu/Test-bank-for-Cost-Benefit-Analysis-4-9780137002696>
- Bridges, J. F. P., Hauber, A. B., Marshall, D., Lloyd, A., Prosser, L. A., Regier, D. A., Johnson, F. R., & Mauskopf, J. (2011). Conjoint Analysis Applications in Health—a Checklist: A Report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. *Value in Health*, 14(4), 403–413. <https://doi.org/10.1016/j.jval.2010.11.013>
- Brouwer, W. B. F., Culyer, A. J., van Exel, N. J. A., & Rutten, F. F. H. (2008). Welfarism vs. extra-welfarism. *Journal of Health Economics*, 27(2), 325–338. <https://doi.org/10.1016/J.JHEALECO.2007.07.003>
- Bunch, D., Louviere, J., & Anderson, D. (1996). A Comparison of Experimental Design Strategies for Multinomial Logit Models : The Case of Generic Attributes. In *Computer Science*.

<https://www.semanticscholar.org/paper/A-Comparison-of-Experimental-Design-Strategies-for-Bunch-Louviere/7c75e3c852ea3cae24465c2b6757960dfbbe5ae8>

- Burcharth, H. F., Hawkins, S. J., Zanuttigh, B., & Lamberti, A. (2007). Design tools related to socio-economics. *Environmental Design Guidelines for Low Crested Coastal Structures*, 347–374. <https://doi.org/10.1016/B978-008044951-7/50035-X>
- Cantillo, J., Martín, J. C., & Román, C. (2020). Discrete choice experiments in the analysis of consumers' preferences for finfish products: A systematic literature review. *Food Quality and Preference*, 84, 103952. <https://doi.org/10.1016/J.FOODQUAL.2020.103952>
- Carlsson, F., & Martinsson, P. (2003). Design techniques for stated preference methods in health economics. *Health Economics*, 12(4), 281–294. <https://doi.org/10.1002/hec.729>
- Carson, R. T., & Czajkowski, M. (2014). The discrete choice experiment approach to environmental contingent valuation. *Handbook of Choice Modelling*, 202–235. <https://doi.org/10.4337/9781781003152.00015>
- Carson, R. T., & Louviere, J. J. (2011). A Common Nomenclature for Stated Preference Elicitation Approaches. *Environmental and Resource Economics*, 49(4), 539–559. <https://doi.org/10.1007/s10640-010-9450-x>
- Chang, J., Savage, S. J., & Waldman, D. M. (2017). Estimating Willingness to Pay for Online Health Services with Discrete-Choice Experiments. *Applied Health Economics and Health Policy*, 15(4), 491–500. <https://doi.org/10.1007/S40258-017-0316-Z/TABLES/8>
- Chang, W. (2022). *CRAN - Package shiny*. <https://cran.r-project.org/web/packages/shiny/index.html>
- Coast, J., Al-Janabi, H., Sutton, E. J., Horrocks, S. A., Vosper, A. J., Swancutt, D. R., & Flynn, T. N. (2012). Using qualitative methods for attribute development for discrete choice experiments: issues and recommendations. *Health Economics*, 21(6), 730–741. <https://doi.org/10.1002/HEC.1739>
- Confraria, J., Ribeiro, T., & Vasconcelos, H. (2017). Analysis of consumer preferences for mobile telecom plans using a discrete choice experiment. *Telecommunications Policy*, 41(3), 157–169. <https://doi.org/10.1016/J.TELPOL.2016.12.009>
- Croissant, Y. (2020). Mlogit: Random utility models in r. *Journal of Statistical Software*, 95, 1–41. <https://doi.org/10.18637/JSS.V095.I11>

- de Bekker-Grob, E. W., Donkers, B., Jonker, M. F., & Stolk, E. A. (2015). Sample Size Requirements for Discrete-Choice Experiments in Healthcare: a Practical Guide. *Patient, 8*(5), 373–384. <https://doi.org/10.1007/s40271-015-0118-z>
- Determann, D., Lambooij, M. S., Gyrd-Hansen, D., de Bekker-Grob, E. W., Steyerberg, E. W., Heldoorn, M., Pedersen, L. B., & de Wit, G. A. (2017). Personal health records in the Netherlands: Potential user preferences quantified by a discrete choice experiment. *Journal of the American Medical Informatics Association, 24*(3), 529–536. <https://doi.org/10.1093/jamia/ocw158>
- Determann, D., Lambooij, M. S., Steyerberg, E. W., de Bekker-Grob, E. W., & de Wit, G. A. (2017). Impact of Survey Administration Mode on the Results of a Health-Related Discrete Choice Experiment: Online and Paper Comparison. *Value in Health, 20*(7), 953–960. <https://doi.org/10.1016/J.JVAL.2017.02.007>
- Devlin, N. J., Shah, K. K., Feng, Y., Mulhern, B., & van Hout, B. (2018). Valuing health-related quality of life: An EQ-5D-5L value set for England. *Health Economics, 27*(1), 7–22. <https://doi.org/10.1002/HEC.3564>
- Drummond, M. F., Sculpher, M. J., Claxton, K., Stoddart, G. L., & Torrance, G. W. (2015). *Methods for the Economic Evaluation of Health Care Programmes* (4th ed.). Oxford University Press.
- Epstein, D. (2021). Nutrition Economics: Celebrating Cross-Disciplinary Collaboration. *Value in Health, 24*(3), 313–316. <https://doi.org/10.1016/J.JVAL.2020.12.004>
- Feng-Gu, E., Everett, J., Brown, R. C. H., Maslen, H., Oakley, J., & Savulescu, J. (2021). Prospective Intention-Based Lifestyle Contracts: mHealth Technology and Responsibility in Healthcare. *Health Care Analysis, 1*–24. <https://doi.org/10.1007/s10728-020-00424-8>
- Fields, B. E., Bell, J. F., Bigbee, J. L., Thurston, H., & Spetz, J. (2018). Registered nurses' preferences for rural and urban jobs: A discrete choice experiment. *International Journal of Nursing Studies, 86*, 11–19. <https://doi.org/10.1016/J.IJNURSTU.2018.05.012>
- Fischer, A. R. H., Berezowska, A., van der Lans, I. A., Ronteltap, A., Rankin, A., Kuznesof, S., Póinhos, R., Stewart-Knox, B., & Frewer, L. J. (2016). Willingness to pay for personalised nutrition across Europe. *European Journal of Public Health, 26*(4), 640–644. <https://doi.org/10.1093/eurpub/ckw045>

- Foltz, J. L., Harris, D. M., & Blanck, H. M. (2012). Support among U.S. adults for local and state policies to increase fruit and vegetable access. *American Journal of Preventive Medicine*, 43(3 SUPPL.2). <https://doi.org/10.1016/j.amepre.2012.05.017>
- Greene, W. H., & Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37(8), 681–698. [https://doi.org/10.1016/S0191-2615\(02\)00046-2](https://doi.org/10.1016/S0191-2615(02)00046-2)
- Gundlach, A., Ehrlinspiel, M., Kirsch, S., Koschker, A., & Sagebiel, J. (2018). Investigating people's preferences for car-free city centers: A discrete choice experiment. *Transportation Research Part D: Transport and Environment*, 63, 677–688. <https://doi.org/10.1016/J.TRD.2018.07.004>
- Güney, O. I., & Giraldo, L. (2020). Consumers' attitudes and willingness to pay for organic eggs: A discrete choice experiment study in Turkey. *British Food Journal*, 122(2), 678–692. <https://doi.org/10.1108/BFJ-04-2019-0297/FULL/PDF>
- Haghani, M., Bliemer, M. C. J., & Hensher, D. A. (2021). The landscape of econometric discrete choice modelling research. *Journal of Choice Modelling*, 40, 100303. <https://doi.org/10.1016/J.JOCM.2021.100303>
- Haghani, M., Bliemer, M. C. J., Rose, J. M., Oppewal, H., & Lancsar, E. (2021a). Hypothetical bias in stated choice experiments: Part I. Integrative synthesis of empirical evidence and conceptualisation of external validity. *Journal of Choice Modelling*, 41. <https://doi.org/10.1016/j.jocm.2021.100309>
- Haghani, M., Bliemer, M. C. J., Rose, J. M., Oppewal, H., & Lancsar, E. (2021b). Hypothetical bias in stated choice experiments: Part II. Macro-scale analysis of literature and effectiveness of bias mitigation methods. *Journal of Choice Modelling*, 41. <https://doi.org/10.1016/j.jocm.2021.100322>
- Hair, J. F., Ringle, C. M., Gudergan, S. P., Fischer, A., Nitzl, C., & Menictas, C. (2019). Partial least squares structural equation modeling-based discrete choice modeling: an illustration in modeling retailer choice. *Business Research*, 12(1), 115–142. <https://doi.org/10.1007/S40685-018-0072-4/TABLES/4>
- Hauber, A. B., González, J. M., Groothuis-Oudshoorn, C. G. M., Prior, T., Marshall, D. A., Cunningham, C., IJzerman, M. J., & Bridges, J. F. P. (2016). Statistical Methods for the Analysis of Discrete Choice Experiments: A Report of the ISPOR Conjoint Analysis Good

- Research Practices Task Force. *Value in Health*, 19(4), 300–315.
<https://doi.org/10.1016/j.jval.2016.04.004>
- Haugen, H. A., Tran, Z. v., Wyatt, H. R., Barry, M. J., & Hill, J. O. (2007). Using Telehealth to Increase Participation in Weight Maintenance Programs**. *Obesity*, 15(12), 3067–3077.
<https://doi.org/10.1038/oby.2007.365>
- Hausman, J. (2012a). *Contingent valuation: A critical assessment*. Elsevier Science BV.
- Hausman, J. (2012b). Contingent valuation: From dubious to hopeless. *Journal of Economic Perspectives*, 26(4), 43–56. <https://doi.org/10.1257/jep.26.4.43>
- Hicks, R. L. (2002). *Stated Preference Methods for Environmental Management: Recreational Summer Flounder Angling in the Northeastern United States*.
- Hoffman, S. D., & Duncan, G. J. (1988). Multinomial and conditional logit discrete-choice models in demography. *Demography*, 25(3), 415–427. <https://doi.org/10.2307/2061541>
- Hole, A. R. (2007). *WTP: Stata module to estimate confidence intervals for willingness to pay measures*. Boston College Department of Economics.
<https://ideas.repec.org/c/boc/bocode/s456808.html>
- Hole, A. R. (2017). *DCREATE: Stata module to create efficient designs for discrete choice experiments*.
- Hole, A. R., Hole, & Arne. (2016). *MIXLOGIT: Stata module to fit mixed logit models by using maximum simulated likelihood*. <https://EconPapers.repec.org/RePEc:boc:bocode:s456883>
- Hoyos, D. (2010). The state of the art of environmental valuation with discrete choice experiments. In *Ecological Economics* (Vol. 69, Issue 8, pp. 1595–1603). Elsevier.
<https://doi.org/10.1016/j.ecolecon.2010.04.011>
- Huber, J., & Zwerina, K. (1996). The Importance of Utility Balance in Efficient Choice Designs. *Journal of Marketing Research*, 33(3), 307–317.
<https://doi.org/10.1177/002224379603300305>
- IBM Corp. (2020). *SPSS (27.0)*. IBM Corp.
- Johnson, F. R., Lancsar, E., Marshall, D., Kilambi, V., Mühlbacher, A., Regier, D. A., Bresnahan, B. W., Kanninen, B., & Bridges, J. F. P. (2013). Constructing experimental designs for discrete-choice experiments: Report of the ISPOR conjoint analysis experimental design

good research practices task force. *Value in Health*, 16(1), 3–13.

<https://doi.org/10.1016/j.jval.2012.08.2223>

Johnston, R. J., Boyle, K. J., Vic Adamowicz, W., Bennett, J., Brouwer, R., Ann Cameron, T., Michael Hanemann, W., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., & Vossler, C. A. (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, 4(2), 319–405. <https://doi.org/10.1086/691697>

Kanninen, B. J. (2018). Optimal Design for Multinomial Choice Experiments:

<https://doi.org/10.1509/Jmkr.39.2.214.19080>, 39(2), 214–227.

<https://doi.org/10.1509/JMKR.39.2.214.19080>

Klose, T. (1999). The contingent valuation method in health care. In *Health Policy* (Vol. 47, Issue 2, pp. 97–123). Elsevier. [https://doi.org/10.1016/S0168-8510\(99\)00010-X](https://doi.org/10.1016/S0168-8510(99)00010-X)

Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. In *Computers and Security* (Vol. 64, pp. 122–134). Elsevier Ltd. <https://doi.org/10.1016/j.cose.2015.07.002>

König, A., & Grippenkov, J. (2020). Modelling travelers' appraisal of ridepooling service characteristics with a discrete choice experiment. *European Transport Research Review*, 12(1), 1–11. <https://doi.org/10.1186/S12544-019-0391-3/TABLES/5>

Kuhfeld, W. F., Tobias, R. D., & Garratt, M. (1994). Efficient Experimental Design with Marketing Research Applications. *Journal of Marketing Research*, 31(4), 545. <https://doi.org/10.2307/3151882>

Lancsar, E., & Louviere, J. (2008). Conducting Discrete Choice Experiments to Inform Healthcare Decision Making. *PharmacoEconomics*, 26(8), 661–677. <https://doi.org/10.2165/00019053-200826080-00004>

Loomis, J. B. (1990). Comparative reliability of the dichotomous choice and open-ended contingent valuation techniques. *Journal of Environmental Economics and Management*, 18(1), 78–85. [https://doi.org/10.1016/0095-0696\(90\)90053-2](https://doi.org/10.1016/0095-0696(90)90053-2)

Louviere, J. J., Flynn, T. N., & Carson, R. T. (2010). Discrete choice experiments are not conjoint analysis. *Journal of Choice Modelling*, 3(3), 57–72. [https://doi.org/10.1016/S1755-5345\(13\)70014-9](https://doi.org/10.1016/S1755-5345(13)70014-9)

- Louviere, J. J., & Lancsar, E. (2009). Choice experiments in health: the good, the bad, the ugly and toward a brighter future. *Health Economics, Policy and Law*, 4(4), 527–546.
<https://doi.org/10.1017/S1744133109990193>
- Mangham, L. J., Hanson, K., & McPake, B. (2009). How to do (or not to do)...Designing a discrete choice experiment for application in a low-income country. In *Health Policy and Planning* (Vol. 24, Issue 2, pp. 151–158). <https://doi.org/10.1093/heapol/czn047>
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics*. (pp. 105–42). Academic Press.
<http://elsa.berkeley.edu/reprints/mcfadden/zarembka.pdf>
- McFadden, D. (1977). Quantitative Methods for Analyzing Travel Behaviour of Individuals: Some Recent Developments. *Cowles Foundation Discussion Papers*.
<https://ideas.repec.org/p/cwl/cwldpp/474.html>
- McFadden, D. (1994). Contingent Valuation and Social Choice. *American Journal of Agricultural Economics*, 76(4), 689–708. <https://doi.org/10.2307/1243732>
- McFadden, D., & Train, K. (2000). MIXED MNL MODELS FOR DISCRETE RESPONSE. *Applied Econometrics*. <https://doi.org/10.1002/1099-1255>
- McGuire, S. (2011). U.S. Department of Agriculture and U.S. Department of Health and Human Services, Dietary Guidelines for Americans, 2010. 7th Edition, Washington, DC: U.S. Government Printing Office, January 2011. *Advances in Nutrition*, 2(3), 293–294.
<https://doi.org/10.3945/an.111.000430>
- McIntosh, E., Clarke, P., Frew, E. J., & Louviere, J. J. (2010). *Applied Methods of Cost-Benefit Analysis in Health Care* (1st ed.). Oxford University Press.
- Miller, A. J., & Nguyen, N.-K. (1994). Algorithm AS 295: A Fedorov Exchange Algorithm for D-Optimal Design. *Applied Statistics*, 43(4), 669. <https://doi.org/10.2307/2986264>
- Mitchell, R. Cameron., & Carson, R. T. (1989). *Using surveys to value public goods: the contingent valuation method*. Resources for the Future.
https://books.google.es/books/about/Using_Surveys_to_Value_Public_Goods.html?id=1R75c1UxVE0C&redir_esc=y
- Monchambert, G. (2020). Why do (or don't) people carpool for long distance trips? A discrete choice experiment in France. *Transportation Research Part A: Policy and Practice*, 132, 911–931. <https://doi.org/10.1016/J.TRA.2019.12.033>

- Mozaffarian, D. (2016). Dietary and Policy Priorities for Cardiovascular Disease, Diabetes, and Obesity. *Circulation*, *133*(2), 187–225.
<https://doi.org/10.1161/CIRCULATIONAHA.115.018585>
- Mustapa, M. A. C., Amin, L., & Frewer, L. J. (2020). Predictors of stakeholders' intention to adopt nutrigenomics. *Genes and Nutrition*, *15*(1), 1–15. <https://doi.org/10.1186/s12263-020-00676-y>
- Ooms, J., & McNamara, J. (2021). *CRAN - Package writexl*. Cran Repository. <https://cran.r-project.org/web/packages/writexl/index.html>
- Pearce, D., Atkinson, G., & Mourato, S. (2006). Cost-benefit analysis and the environment: Recent developments. In *Cost-benefit analysis and the environment: Recent developments* (Vol. 9789264010055). Organisation for Economic Cooperation and Development (OECD).
<https://doi.org/10.1787/9789264010055-EN>
- Pérez-Troncoso, D., Epstein, D. M., & Castañeda-García, J. A. (2021). Consumers' Preferences and Willingness to Pay for Personalised Nutrition. *Applied Health Economics and Health Policy* *2021* *19*:5, *19*(5), 757–767. <https://doi.org/10.1007/S40258-021-00647-3>
- Poínhos, R., van der Lans, I. A., Rankin, A., Fischer, A. R. H., Bunting, B., Kuznesof, S., Stewart-Knox, B., & Frewer, L. J. (2014). Psychological Determinants of Consumer Acceptance of Personalised Nutrition in 9 European Countries. *PLoS ONE*, *9*(10), e110614.
<https://doi.org/10.1371/journal.pone.0110614>
- Qualtrics. (2020). *Qualtrics*. Qualtrics.
- R Core Team. (2021). *R: A language and environment for statistical computing* (4.1.0). Foundation for Statistical Computing. <https://www.r-project.org/>
- Rakotonarivo, O. S., Schaafsma, M., & Hockley, N. (2016). A systematic review of the reliability and validity of discrete choice experiments in valuing non-market environmental goods. *Journal of Environmental Management*, *183*, 98–109.
<https://doi.org/10.1016/J.JENVMAN.2016.08.032>
- Ramos-Goñi, J. M., Craig, B. M., Oppe, M., Ramallo-Fariña, Y., Pinto-Prades, J. L., Luo, N., & Rivero-Arias, O. (2018). Handling Data Quality Issues to Estimate the Spanish EQ-5D-5L Value Set Using a Hybrid Interval Regression Approach. *Value in Health*, *21*(5), 596–604.
<https://doi.org/10.1016/J.JVAL.2017.10.023/ATTACHMENT/8E441C35-E3AE-4939-AC61-47E6B3D8BD37/MMC3.DOCX>

- Rankin, A., Bunting, B. P., Poinhos, R., van der Lans, I. A., Fischer, A. R. H., Kuznesof, S., Almeida, M. D. V., Markovina, J., Frewer, L. J., & Stewart-Knox, B. J. (2018). Food choice motives, attitude towards and intention to adopt personalised nutrition. *Public Health Nutrition*, *21*(14), 2606–2616. <https://doi.org/10.1017/S1368980018001234>
- Rankin, A., Kuznesof, S., Frewer, L. J., Orr, K., Davison, J., de Almeida, M. D. V., & Stewart-Knox, B. (2017). Public perceptions of personalised nutrition through the lens of Social Cognitive Theory. *Journal of Health Psychology*, *22*(10), 1233–1242. <https://doi.org/10.1177/1359105315624750>
- Rose, J. M., & Bliemer, M. (2005). Constructing Efficient Choice Experiments. *INSTITUTE of TRANSPORT and LOGISTICS STUDIES*. <http://www.itls.usyd.edu.au>
- Rose, J. M., & Bliemer, M. C. J. (2004). *THE DESIGN OF STATED CHOICE EXPERIMENTS: THE STATE OF PRACTICE AND FUTURE CHALLENGES*. <https://trid.trb.org/view/704838>
- Rose, J. M., & Bliemer, M. C. J. (2014). Stated choice experimental design theory: The who, the what and the why. In *Handbook of Choice Modelling* (pp. 152–177). <https://doi.org/10.4337/9781781003152.00013>
- Rose, J. M., & Bliemer, M. C. J. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews*, *29*(5), 587–617. <https://doi.org/10.1080/01441640902827623>
- Rose, J. M., Scarpa, R., & Bliemer, M. C. J. (2008). *Incorporating model uncertainty into the generation of efficient stated choice experiments: A model averaging approach*.
- RStudio Team. (2021). *RStudio: Integrated Development Environment for R* (1.4.1106). <http://www.rstudio.com/>
- Samuelson, P. A. (1938). A Note on the Pure Theory of Consumer's Behaviour. *Economica*, *5*(17), 61. <https://doi.org/10.2307/2548836>
- Sándor, Z., & Wedel, M. (2001). Designing Conjoint Choice Experiments Using Managers' Prior Beliefs. <https://doi.org/10.1509/Jmkr.38.4.430.18904>, *38*(4), 430–444. <https://doi.org/10.1509/JMKR.38.4.430.18904>
- Shah, K. K., Tsuchiya, A., & Wailoo, A. J. (2015). Valuing health at the end of life: A stated preference discrete choice experiment. *Social Science & Medicine*, *124*, 48–56. <https://doi.org/10.1016/J.SOCSCIMED.2014.11.022>
- Smith, K. J., Kuo, S., Zgibor, J. C., McTigue, K. M., Hess, R., Bhargava, T., & Bryce, C. L. (2016). Cost effectiveness of an internet-delivered lifestyle intervention in primary care patients

with high cardiovascular risk. *Preventive Medicine*, 87, 103–109.

<https://doi.org/10.1016/j.ypmed.2016.02.036>

- Soekhai, V., de Bekker-Grob, E. W., Ellis, A. R., & Vass, C. M. (2019). Discrete Choice Experiments in Health Economics: Past, Present and Future. In *PharmacoEconomics* (Vol. 37, Issue 2, pp. 201–226). <https://doi.org/10.1007/s40273-018-0734-2>
- StataCorp LLC. (2017). *Stata Statistical Software* (Release 15).
- Stewart-Knox, B., Rankin, A., Kuznesof, S., Póinhos, R., Vaz De Almeida, M. D., Fischer, A., & Frewer, L. J. (2015). Promoting healthy dietary behaviour through personalised nutrition: Technology push or technology pull? *Proceedings of the Nutrition Society*, 74(2), 171–176. <https://doi.org/10.1017/S0029665114001529>
- Street, D. J., Burgess, L., & Louviere, J. J. (2005). Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing*, 22(4), 459–470. <https://doi.org/10.1016/J.IJRESMAR.2005.09.003>
- Therneau, T., & Lumley, T. (2022). *CRAN - Package survival*. Cran Repository. <https://cran.r-project.org/web/packages/survival/index.html>
- Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, 34(4), 273–286. <https://doi.org/10.1037/h0070288>
- Torrance, G. W. (1986). Measurement of Health State Utilities for Economic Appraisal. *Journal of Health Economics*, 5.
- Traets, F., Gil Sánchez, D., & Vandebroek, M. (2020). idefix: Efficient Designs for Discrete Choice Experiments. *Journal of Statistical Software*, 96(3). <https://cran.r-project.org/package=idefix>
- Train, K. E. (2008). EM Algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling*, 1(1), 40–69. [https://doi.org/10.1016/S1755-5345\(13\)70022-8](https://doi.org/10.1016/S1755-5345(13)70022-8)
- Train, K. E. (2009). Discrete choice methods with simulation, second edition. In *Discrete Choice Methods with Simulation, Second Edition* (Vol. 9780521766555). Cambridge University Press. <https://doi.org/10.1017/CBO9780511805271>
- Vallée Marcotte, B., Cormier, H., Garneau, V., Robitaille, J., Desroches, S., & Vohl, M.-C. (2018). Nutrigenetic Testing for Personalized Nutrition: An Evaluation of Public Perceptions, Attitudes, and Concerns in a Population of French Canadians. *Lifestyle Genomics*, 11(3–6), 155–162. <https://doi.org/10.1159/000499626>

- Veldwijk, J., Lambooi, M. S., van Gils, P. F., Struijs, J. N., Smit, H. A., & de Wit, G. A. (2013). Type 2 diabetes patients' preferences and willingness to pay for lifestyle programs: A discrete choice experiment. *BMC Public Health*, *13*(1), 1099. <https://doi.org/10.1186/1471-2458-13-1099>
- Wang, Y., Wang, Z., Wang, Z., Li, X., Pang, X., & Wang, S. (2021). Application of Discrete Choice Experiment in Health Care: A Bibliometric Analysis. *Frontiers in Public Health*, *9*, 591. <https://doi.org/10.3389/FPUBH.2021.673698/BIBTEX>
- Wansink, B., & Sobal, J. (2016). Mindless Eating: The 200 Daily Food Decisions We Overlook. <http://Dx.Doi.Org/10.1177/0013916506295573>, *39*(1), 106–123. <https://doi.org/10.1177/0013916506295573>
- WHO. (2012). *How to conduct a discrete choice experiment for health workforce recruitment and retention in remote and rural areas: a user guide with case studies*.
- Wickham, H., Bryan, J., & RStudio. (2022). *CRAN - Package readxl*. Cran Repository. <https://cran.r-project.org/web/packages/readxl/index.html>
- Willis, K. G. (2014). The Use of Stated Preference Methods to Value Cultural Heritage. *Handbook of the Economics of Art and Culture*, *2*, 145–181. <https://doi.org/10.1016/B978-0-444-53776-8.00007-6>
- Yu, J., Goos, P., & Vandebroek, M. (2011). Individually adapted sequential Bayesian conjoint-choice designs in the presence of consumer heterogeneity. *International Journal of Research in Marketing*, *28*(4), 378–388. <https://doi.org/10.1016/J.IJRESMAR.2011.06.002>
- Zeevi, D., Korem, T., Zmora, N., Israeli, D., Rothschild, D., Weinberger, A., Ben-Yacov, O., Lador, D., Avnit-Sagi, T., Lotan-Pompan, M., Suez, J., Mahdi, J. A., Matot, E., Malka, G., Kosower, N., Rein, M., Zilberman-Schapira, G., Dohnalová, L., Pevsner-Fischer, M., ... Segal, E. (2015). Personalized Nutrition by Prediction of Glycemic Responses. *Cell*, *163*(5), 1079–1094. <https://doi.org/10.1016/j.cell.2015.11.001>
- Zwerina, K., Huber, J., & Kuhfeld, W. F. (1996). A general method for constructing efficient choice designs. *Durham, NC: Fuqua School of Business, Duke University*.

APÉNDICE CAPÍTULO 2. EXPERIMENTOS DE ELECCIÓN DISCRETA

2.1 Código de Stata

```
1.  ssc install dcreate /// Instalación del paquete
2.  //// En primer lugar crearemos un diseño 23 con 8 sets.
3.  matrix levmat = 2, 2, 2 // Vector llamado levmat con un número // por cada
   atributo indicando el número de niveles
4.  genfact, levels(levmat) // Creamos todas las alternativas posibles
5.  matrix priors = 0.5, 0.76, 1.5 // Damos un valor a los priors
6.  dcreate i.x1 i.x2 i.x3, nalt(2) nset(8) bmat(priors) //Diseño
7.  //// Ahora creamos el mismo diseño usando priors bayesianos
8.  clear // Borramos el diseño anterior
9.  genfact, levels(levmat) // Generamos de nuevo las alternativas
10. matrix I = I(3) // Generamos una matriz identidad para las
   // varianzas de los priors
11. dcreate i.x1 i.x2 i.x3, nalt(2) nset(8) bmat(priors) vmat(I)
12. // Diseño experimental con alternativa constante (nula)
13. clear all // Lo borramos todo
14. matrix levmat = 2, 2, 2
15. genfact, levels(levmat)
16. matrix optout = 1,1,1 // Alternativa nula con niveles base
17. matrix priors = 0.5, 0.76, 1.5
18. dcreate i.x1 i.x2 i.x3, nalt(2) nset(8) bmat(priors) fixedalt(optout) asc(3)
```

2.2 Código de R

```
1. install.packages("idefix") # Instalación del paquete
2. library(idefix) # Activación del paquete
3. niveles <- c(2,2,2) # Vector 'niveles' con 3 atrib. de 2 niv.
4. codif <- c("D","D","D") # Indicamos la codif. de cada atrib.
5. priors <- c(0.5, 0.76, 1.5) # Valor para los priors
6. I <- diag(length(priors)) # Matriz diagonal tamaño priors
7. sim <- MASS::mvrnorm(n=100, mu=priors, Sigma=I) # Distribución
8. dis <- CEA(lvls = niveles, coding = codif, n.alts = 2, n.sets = 16, par.draws = sim, no.choice = FALSE, best = TRUE).
9. dis$design # El diseño experimental
10. dis$error # D-error
11. # Diseño experimental con alternativa nula
12. priors <- c(0, 0.5, 0.76, 1.5) # Valor para los priors
13. sim <- list(sim[,1:1], sim[, 2:(length(priors))])
14. constante <- c(0,0,1)
15. dis <- CEA(lvls=niveles, coding=codif, n.alts=3, n.sets=16, alt.cte=constante, par.draws=sim, no.choice=TRUE, best=TRUE)
16. dis$design # El diseño experimental
17. dis$error # D-error
```

2.3 Código de R

El código íntegro para generar los diseños puede ser descargado de:

- https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/dce_toolbox.R

Este mismo código puede encontrarse en el paquete 'DCEtool', abordado en el Capítulo 5 de esta tesis. Este paquete contiene la función 'dce_toolbox', la cual es la misma programación que puede encontrarse en el enlace anterior. Este paquete si contiene documentación que describe como usar la función:

```
1. install.packages("DCEtool")
2. library(DCEtool)
3. help(dce_toolbox)
```

2.4 Simulaciones y análisis de los resultados

El código íntegro para generar las simulaciones y analizar los resultados puede ser descargado de:

- https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/C2S2.6.R
- El LCL fue estimado usando Stata:
https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/C2S2.6.b.do

APÉNDICE TEMA 3. METODOLOGÍA DE LA INVESTIGACIÓN

3.1 Referencias y hallazgos de la revisión literaria

Referencia	Población	Intervención	Ctr.	Tipo	Hallazgos
(Ballco & De Magistris, 2019)	Adultos en España	Yogur con reclamos nutricionales	-	DCE	Tres subgrupos poblacionales: uno centrado en salud y nutrición, otro sólo en salud, y el tercero indiferente a ambos.
(Veldwijk et al., 2013)	Adultos en Países Bajos con Diabetes tipo II	Intervención de hábitos	-	DCE	El ratio de participación de los pacientes crónicos está entre el 50 y 60 p.c.
(Rankin et al., 2018)	Muestra representativa europea	Selección de alimentos dependiendo de nueve factores	-	Encuesta	El coste directo al usuario determina su disposición.
(Molema et al., 2019)	Adultos en Países Bajos con Diabetes tipo II o enfermedad cardiov.	Incentivos por logros	-	DCE	Los pacientes aceptan los incentivos sólo si cumplieron los objetivos designados.
(Determann et al., 2017)	Muestra de varias federaciones de Países Bajos	Uso de seguimientos personales de salud	-	DCE	Tres grupos dependiendo de su actitud (a favor, neutrales, en contra).
(Foltz et al., 2012)	Encuesta de salud 2008 EEUU	Medidas locales de promoción de verduras	-	Encuesta	Los consumidores apoyan las medidas
(Haugen et al., 2007)	Adultos EEUU que completaron un plan de 24 semanas de pérdida de peso	Programa de telesalud	Trad.	Quasi-experim.	Los usuarios puntúan la experiencia como positiva y útil. Los programas de telesalud son efectivos.
(Smith et al., 2016)	Pacientes con sobrepeso / obesidad y riesgo cardio	Programa online de prevención diabetes	BMT	Markov (cost-effect)	Los programas online podrían ser una solución coste-efec.
(Han et al., 2019)	Personas con IMC > 25 kg/m ² .	App móvil con algoritmo de prescripción de dieta (12 s.)	-	Quasi-experim.	Bajó el IMC y otras medidas mejoraron
(Vallée Marcotte et al., 2018)	Muestra representativa población Fr-canadiense	Actitudes hacia nutrigenética	-	Encuesta	Interesados en el uso general pero preocupados por la privacidad
(Feng-Gu et al., 2021)	Público general	Actitudes hacia la responsab. Individual	-	Encuesta	La gente cree generalmente que cada

					individuo debe cuidar su salud.
(Poinhos et al., 2014)	Muestra de nueve países europeos	Determinantes de la adopción de NP	-	Encuesta	Las actitudes hacia la NP están determinadas por los resultados esperados
(Stewart-Knox et al., 2015)	Muestra de nueve países europeos	Recolección de evidencias de Food4Me	-	Encuesta (cuali y cuanti)	The perceived benefit increased the intention to undergo a personalised nutrition intervention.
(Rankin et al., 2017)	Dos grupos de debate en dos regiones de UK	Actitudes y percepciones hacia la nutrición personalizada	-	Debate	El seguimiento y entrenamiento por un profesional aumenta las probabilidades de que la intervención sea efectiva. Quien paga, desea someterse a la intervención.
(Fischer et al., 2016)	Muestra de nueve países europeos	DAP hacia NP	-	VC DAP	Un tercio está dispuesto a pagar entre 30 y 40 euros al mes por NP
(Mustapa et al., 2020)	Adultos de Malaysia	Disposición a adoptar nutrigenética	-	Encuesta	Un equilibrio entre riesgos y beneficios es fundamental para decidir
(Mattei & Alfonso, 2020)	Nutricionistas de Puerto Rico	Estrategias de promoción de cambio conductural	-	Encuesta	Una horientación personalizada es de gran importancia
(Chang et al., 2017)	Muestra representative de hogar de EEUU	DAP por servicios de telesalud	-	DCE	La media son 5\$ al mes, pero esta se incrementa con mayores costes de oportunidad.

Tabla A3.1 Resumen de la revisión de la literatura

3.2 Obtención de atributos y niveles

Nombre de la aplicación	Versión	Precio	Descripción
Nootric (https://www.nootric.es/)	V.1.	0 €/mes	Menús semanales y recetas limitadas
	V.2.	9.99 €/mes	Menús y recetas ilimitadas + seguimiento de nutricionista
	V.3.	23 €/mes	Menús y recetas ilimitadas + seguimiento de nutricionista + dieta
Fitia (https://www.fitiaapp.com/)	-	7 €/mes	Personalización basada en edad, peso, altura, hábitos deportivos + funcionalidad lista de la compra + seguimiento virtual
Yazio (https://www.yazio.com/)	V.1	0 €/mes	Personalización basada en edad, peso, altura y objetivos de la dieta
	V.2.	3 €/mes	Personalización basada en edad, peso, altura y objetivos de la dieta + plan nutricional y ayunos + compatible con smartwatches.
Lifesum (https://lifesum.com/)	V.1.	0 €/mes	Personalización basada en edad, peso, altura y diario de comidas
	V.2.	8 €/mes	Personalización basada en edad, peso, altura y diario de comidas + menús y recetas
FitMenCook (https://fitmencook.com/)	-	0 €/mes	Recetas
MyFitnessPal (https://www.myfitnesspal.com/)	V.1.	0 €/mes	Contador calórico y objetivo diario de macronutrientes
	V.2.	9.99 €/mes	Contador calórico y objetivo diario de macronutrientes + panel de nutrientes + análisis de comida + objetivo de deporte diario
Looselt!	V.1.	0 €/mes	Plan de pérdida de peso personalizado + apoyo comunitario + sincronización
	V.2.	3.33 €/mes	Plan de pérdida de peso personalizado + apoyo comunitario + sincronización+ seguimiento avanzado, comidad y ejercicio, patrones y celebraciones
MyNetDiary (https://www.mynetdiary.com/)	-	8.28 €/mes	Diet planning, food logging, exercise tracking, weight tracking Plan de dieta, seguimiento de comida, seguimiento de ejercicio y de peso.
DietPoint (http://www.dietpointed.com/)	-	15 €/mes	Dieta personalizada basada en edad, peso, altura y objetivos.

Tabla A3.2 Aplicaciones empleadas para la definición de atributos y niveles

3.3 Primera selección de atributos y niveles

Atributos	Niveles	
Nivel de personalización (cantidad de información recopilada para diseñar la dieta)	Nivel 1	Medidas corporales
	Nivel 2	+ sugerencias para alergias / intolerancias
	Nivel 3	+ seguimiento calórico diario
	Nivel 4	+ análisis de la microbiota intestinal
Seguimiento (seguimiento del cumplimiento de la dieta)	Sí	Seguimiento mensual por nutricionista
	No	Sin seguimiento
Plan de entrenamiento (personalizado basado en datos del usuario)	Sí	Plan de entrenamiento diario a medida
	No	Sin plan de entrenamiento
Video tutoriales de recetas	Sí	Videos paso a paso de las recetas
	No	Sin videos
Precio (cantidad mensual)	10 €	10 € / mes
	20 €	20 € / mes
	30 €	30 € / mes
	40 €	40 € / mes

Tabla A3.3 Selección de atributos y niveles para el pretest

3.4 Resultados de la regresión logística condicional en el DCE piloto

Código para estimar el logit condicional del experimento piloto en R

```

1 load(url('https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/dat.rda'))
2 load(url('https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/des.rda'))
3 install.packages("DCEmgmt")
4 library(DCEmgmt)
5 dat <- DCEmgmt(data = datos, create.id = TRUE, options = c("A", "B"), alts = 2,
6 sets = 16, blocks = c("B_"), design = diseno)
7 install.library("survival")
8 library(survival)
9 clogit(choice ~ pers2 + pers3 + pers4 + monit2 + rut2 + recip2 + price2 +
10 price3 + price4 + strata(gid), data = dat)
dat$precio <- ifelse(dat$price1 == 1, 10, ifelse(dat$price2 == 1, 20,
ifelse(dat$price3 == 1, 30, 40)))
clogit(choice ~ pers2 + pers3 + pers4 + monit2 + rut2 + precio + strata(gid),
data = dat)

```

Atributos	Niveles	Coef.	SE (P>z)	P-valor
Person.	Nivel 2	0.452	0.173	0.009
	Nivel 3	0.591	0.207	0.004
	Nivel 4	0.904	0.166	> 0.000
Monitorización	Sí	1.197	0.110	> 0.000
Rutinas	Sí	0.273	0.104	0.009
Recetas	Sí	0.151	0.104	0.146
Price	(lineal)	-0.056	0.072	> 0.000

N = 40, Observaciones = 1280 [40·16·2], Pseudo R² = 0.3017

Tabla A3.4 Resultados del logit condicional con el precio como variable lineal

Atributos	Niveles	Coef.	SE (P>z)	P-valor
Person.	Nivel 2	0.489	0.171	0.009
	Nivel 3	0.581	0.210	0.004
	Nivel 4	0.848	0.165	> 0.000
Monitorización	Sí	1.191	0.109	> 0.000
Rutinas	Sí	0.283	0.105	0.009
Recetas	Sí	0.155	0.105	0.146
Price	20€	-0.614	0.177	0.001
	30€	-0.937	0.219	> 0.000
	40€	-1.790	0.212	> 0.000

N = 40, Observaciones = 1280 [40·16·2], Pseudo R² = 0.3059

Tabla A3.5 Resultados del logit condicional con el precio como variables binarias

3.5 Segunda especificación de atributos y niveles

Attribute	Levels	
Nivel de personalización (cantidad de información recopilada para diseñar la dieta)	Level 1	Hábitos saludables y educación nutricional
	Level 2	Dieta personalizada basada en altura, peso y estatura.
	Level 3	Dieta personalizada basada en altura, peso, estatura, hábitos deportivos y análisis de sangre
	Level 4	Dieta personalizada basada en altura, peso, estatura, hábitos deportivos, análisis de sangre y análisis de la microbiota intestinal.
Formato de provision	Físico	Servicio ofrecido en un lugar físico.
	Digital	Servicio ofrecido vía una aplicación móvil.
Seguimiento (seguimiento del cumplimiento de la dieta)	Sí	Seguimiento mensual de un nutricionista.
	No	Sin seguimiento.
Plan de entrenamiento (personalizado basado en datos del usuario)	Sí	Plan de entrenamiento basado en información del usuario
	No	Sin plan de entrenamiento
Suplementos nutricionales (como vitaminas o minerales)	Yes	El servicio recomienda suplementos si se detecta un déficit.
	No	Sin suplementos.
Precio (mensual)	10 €	10 € / mes
	20 €	20 € / mes
	30 €	30 € / mes
	40 €	40 € / mes

Tabla A3.6 Segunda selección de atributos y niveles

3.6 Creación de la encuesta final

Código para la creación del diseño eficiente en Stata

```
1 use "https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/pre.dta"
2 clogit choice i.pers i.monit i.exrout i.recip i.price, group(gid)
3 ssc install dcreate
4 clear
5 matrix atlev = 4,2,2,4
6 genfact, levels(atlev)
7 matrix priors = (_b[1.pers], _b[2.pers], _b[3.pers], 0, _b[1.exrout],
8 _b[1.price], _b[2.price], _b[3.price])
9 rename x1 pers
10 label variable pers "Personalization level (0, 1, 2, 3)"
11 rename x2 form
12 label variable form "Intervention format (0=traditional, 1=app)"
13 rename x3 exrout
14 label variable exrout "Exercise routines recommendations(0=no, 1=yes)"
15 rename x4 price
16 label variable price "Price (10€, 20€, 30€, 40€)"
17 dcreate i.pers i.form i.exrout i.price, nalt(2) nset(16) bmat(priors)
18 blockdes block, nblock(2)
```

3.7 Estimaciones

Análisis de los resultados en Stata

```
1 use "https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/dcedataset.dta"
2 tabulate pers, gen(pers)
3 tabulate form, gen(form)
4 tabulate rut, gen(rut)
5 /// Regresión logística condicional (los datos deben estar en formato 'largo')
6 clogit choice i.pers i.form i.rut price, strata(qid)
7 /// Regresión logística de parámetros aleatorios
8 ssc install mixlogit //instalar si se requiere
9 mixlogit choice, rand(pers2 pers3 pers4 form2 rut2 price) id(pid) group(qid)
10 /// Disposición a pagar desde el logit de parámetros aleatorios
11 nlcom -_b[pers2]/_b[price] // WTP nivel pers. 2
12 nlcom -_b[pers3]/_b[price] // WTP nivel pers. 3
13 nlcom -_b[pers4]/_b[price] // WTP nivel pers. 4
14 nlcom -_b[form2]/_b[price] // WTP formato digital
15 nlcom -_b[rut2]/_b[price] // WTP formato rutinas deportivas
16 /// Regresión logística de clases latentes
17 ssc install lclogit2
18 forvalues c = 2/5 {
19     lclogit2 choice, id(pid) group(qid) nclasses(`c') rand(per2 pers3 form1
20     rut1
21     price) membership(income_r educ age dis inter)
22     matrix n = e(b)
23     matrix ic = nullmat(ic) \ `e(nclasses)', `e(ll)',`=colsof(n)', `e(aic)',
24     `e(caic)', `e(bic)'
25 }
26 matrix colnames ic = "Classes" "LLF" "Nparam" "AIC" "CAIC" "BIC"
27 matlist ic, name(columns)
28 lclogit2 choice, id(pid) group(qid) nclasses(4) rand(per2 pers3 form1 rut1
29 price) membership(income_r educ age dis inter)
```

Comparación de diferencias entre coeficientes (Pers 3 vs. Pers4)

```
1 use "https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/dcedataset.dta"
2 tabulate pers, gen(pers)
3 tabulate form, gen(form)
4 tabulate rut, gen(rut)
5 clogit choice pers2 pers3 pers4 form2 rut2 price, strata(qid)
6 ttest pers3 = pers4
```

Calculando la heterogeneidad en el logit mixto

```
1 use "https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/dcedataset.dta"
2 tabulate pers, gen(pers)
3 tabulate form, gen(form)
4 tabulate rut, gen(rut)
5 mixlogit choice, group(qid) id(pid) rand(pers2 pers3 pers4 form2 rut2 price)
5 // Heterogeneidad
6 matrix coef = e(b)
7 scalar sdpers2 = coef[1,7]
8 scalar sdpers3 = coef[1,8]
9 scalar sdpers4 = coef[1,9]
10 scalar sdform2 = coef[1,10]
11 scalar sdrut2 = coef[1,11]
12 dis 100*normal(-_b[pers2]/scalar(sdpers2))
13 dis 100*normal(-_b[pers3]/scalar(sdpers3))
14 dis 100*normal(-_b[pers4]/scalar(sdpers4))
15 dis 100*normal(-_b[form2]/scalar(sdform2))
16 dis 100*normal(-_b[rut2]/scalar(sdrut2))
```


3.8 Análisis coste-beneficios

Código para el análisis coste-beneficios

```
1 use "https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/dcedataset.dta"
2 /// Función para calcular las disposiciones a pagar
3 program estwtp, rclass
4     clogit choice i.pers i.form i.rut i.price, strata(qid)
5     scalar wtp_pers1 = -_b[1.pers]/_b[price]
6     scalar wtp_pers2 = -_b[2.pers]/_b[price]
7     scalar wtp_pers3 = -_b[3.pers]/_b[price]
8     scalar wtp_form1 = -_b[1.form]/_b[price]
9     scalar wtp_rut1 = -_b[1.rut]/_b[price]
10    return scalar wtp = wtp_pers2 + wtp_form1 + wtp_rut1
11 End
12 /// Bootstrap WTP (todas las clases)
13 bootstrap r(wtp), r(500) saving(total, replace) seed(1357): estwtp
14 /// Clase 1
15 drop if class != 1
16 program estwtp1, rclass
17 /// Pegar líneas 3-13
18 clear
19 /// Clase 2
20 use "https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/pre.dta"
21 drop if class != 2
22 /// Pegar líneas 3-13
23 clear
24 /// Clase 3
25 use "https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/pre.dta"
26 drop if class != 3
27 /// Pegar líneas 3-13
28 clear
29 /// Clase 4
30 use "https://raw.githubusercontent.com/danielpereztr/DPT_Thesis_Data/main/pre.dta"
31 drop if class != 4
32 /// Pegar líneas 3-13
```

APÉNDICE TEMA 4. MATERIAL SUPLEMENTARIO

4.1 Código de R para las simulaciones

Código para las simulaciones de estrategias “constantes”

```
1 f <- id #estimation id number
2 iterations <- I #no. of iterations
3 resp <- as.numeric(N) #respondents per iteration
4 lev <- c(2,2,3) #attributes and levels
5 cod <- c("E","E","E") #effects coding
6 sets <- S #no. of choice sets
7 library(idefix)
8 priors <- c(0, 0, 0, 0, 0) #first priors set to zero
9 dataset.f <- c() #creation of the final dataset
10 i <- 0 #iterations counter
11 montecarlo <- 0 #Monte Carlo iterations counter
12 results <- c() #storage of coefficients
13 library(survival) #storage of coefficients
14 out.fin <- list() #storage of coefficients
15 x11.fin <- list() #storage of coefficients
16 x21.fin <- list() #storage of coefficients
17 x31.fin <- list() #storage of coefficients
18 x32.fin <- list() #storage of coefficients
19
20 while (montecarlo<100){ #Monte Carlo loop
21
22   while (i<iterations) { #Bliemer and Rose loop
23     I <- diag(length(priors))
24     sim <- MASS::mvrnorm(n=100, mu=priors, Sigma=I)
25     sim <- list(sim[,1:1], sim[, 2:(length(priors))])
26     con <- c(0,0,1)
27     d <- CEA(lvls=lev, coding=cod, n.alts=3, n.sets=sets, alt.cte=con, par.draws=sim,
28     no.choice=TRUE, best=TRUE)
29     des <- d$design
30     truePREF <- c(0, -2, -1, 1, 2)
31     desm <- do.call(rbind, replicate(resp, des, simplify=FALSE))
32     choice <- RespondMNL(par = truePREF, des=desm, n.alts=3)
33     x <- nrow(desm)
34     cs <- rep((i*240):(x/3), each= 3, length.out=240)
35     dataset.t <- cbind(cs, desm, choice)
36     dataset.t <- as.data.frame(dataset.t)
```

```
37 dataset.f <- rbind(dataset.f,dataset.t)
38 res <- clogit(choice ~ alt3.cte + Var11 + Var21 + Var31 + Var32 + strata(cs),
39 data = dataset.f, method="efron")
40 truco <- (summary(res)$coefficients[,1])
41 truco <- as.data.frame((cbind(truco, summary(res)$coefficients[,4])))
42 truco <- cbind(truco,ifelse(abs(truco[,2]) >= 1.96,1,0))
43 truco <- cbind(truco, truco[,1]*truco[,3])
44 priors <- truco[,4]
45 #fin
46 i <- i+1
47 results <- rbind(results, priors)
48 n.it <- "iteration"
49 print(paste(n.it, i))
50 } #end of Bliemer and Rose loop
51
52 i <- 0
53 x11.fin <- as.data.frame(cbind(x11.fin, results[,2])) #storage of coefficients
54 x21.fin <- as.data.frame (cbind(x21.fin, results[,3])) #storage of coefficients
55 x31.fin <- as.data.frame(cbind(x31.fin, results[,4])) #storage of coefficients
56 x32.fin <- as.data.frame(cbind(x32.fin, results[,5])) #storage of coefficients
57 dataset.f <- c() #clear dataset
58 results <- c() #clear results array
59 monti <- "Iteration MONTECARLO" # Monte Carlo iteration counter
60 print(paste(monti, montecarlo)) # Monte Carlo iteration counter
61 montecarlo<-montecarlo+1 # Monte Carlo iteration counter
62 }
63
64 library(writexl) #print 4 excel files
65 write_xlsx(x11.fin, sprintf("sim%fx11.xlsx",f))
66 write_xlsx(x21.fin, sprintf("sim%fx21.xlsx",f))
67 write_xlsx(x31.fin, sprintf("sim%fx31.xlsx",f))
68 write_xlsx(x32.fin, sprintf("sim%fx32.xlsx",f))
```

Código para las simulaciones de estrategias “no constantes”

```
1 f <- id #estimation id number
2 iterations <- I #no. of iterations
3 resp <- as.numeric(N) #respondents per iteration
4 lev <- c(2,2,3) #attributes and levels
5 cod <- c("E","E","E") #effects coding
6 sets <- S #no. of choice sets
7 library(idefix)
8 priors <- c(0, 0, 0, 0, 0) #first priors set to zero
9 dataset.f <- c() #creation of the final dataset
10 i <- 0 #iterations counter
11 montecarlo <- 0 #Monte Carlo iterations counter
12 results <- c() #storage of coefficients
13 library(survival) #storage of coefficients
14 out.fin <- list() #storage of coefficients
15 x11.fin <- list() #storage of coefficients
16 x21.fin <- list() #storage of coefficients
17 x31.fin <- list() #storage of coefficients
18 x32.fin <- list() #storage of coefficients
19
20 while (montecarlo<100){ #Monte Carlo loop
21
22   while (i<iterations) { #Bliemer and Rose loop
23     I <- diag(length(priors))
24     sim <- MASS::mvrnorm(n=100, mu=priors, Sigma=I)
25     sim <- list(sim[,1:1], sim[, 2:(length(priors))])
26     con <- c(0,0,1)
27     d <- CEA(lvls=lev, coding=cod, n.alts=3, n.sets=sets, alt.cte=con, par.draws=sim,
28     no.choice=TRUE, best=TRUE)
29     des <- d$design
30     truePREF <- c(0, -2, -1, 1, 2)
31     desm <- do.call(rbind, replicate(resp, des, simplify=FALSE))
32     choice <- RespondMNL(par = truePREF, des=desm, n.alts=3)
33     x <- nrow(desm)
34     cs <- rep((i*240):(x/3), each= 3, length.out=240)
35     dataset.t <- cbind(cs, desm, choice)
36     dataset.t <- as.data.frame(dataset.t)
37     dataset.f <- rbind(dataset.f,dataset.t)
38     res <- clogit(choice ~ alt3.cte + Var11 + Var21 + Var31 + Var32 + strata(cs),
39     data = dataset.f, method="efron")
```

```
40   truco <- (summary(res)$coefficients[,1])
41   truco <- as.data.frame((cbind(truco, summary(res)$coefficients[,4])))
42   truco <- cbind(truco,ifelse(abs(truco[,2]) >= 1.96,1,0))
43   truco <- cbind(truco, truco[,1]*truco[,3])
44   priors <- truco[,4]
45   #fin
46   i <- i+1
47   results <- rbind(results, priors)
48   n.it <- "iteration"
49   print(paste(n.it, i))
50 } #end of Bliemer and Rose loop
51
52 #last iteration out of the loop
53 I <- diag(length(priors))
54 sim <- MASS::mvrnorm(n=100, mu=priors, Sigma=I)
55 sim <- list(sim[,1:1], sim[, 2:(length(priors))])
56 con <- c(0,0,1)
57 d <- CEA(lvls=lev, coding=cod, n.alts=3, n.sets=sets, alt.cte=con, par.draws=sim,
58 no.choice=TRUE, best=TRUE)
59 des <- d$design
59 truePREF <- c(0, -2, -1, 1, 2)
60 desm <- do.call(rbind, replicate(respseg, des, simplify=FALSE))
61 choice <- RespondMNL(par = truePREF, des=desm, n.alts=3)
62 cs <- rep(max(cs):(nrow(desm)/3), each=3)
63 dataset.t <- cbind(cs, desm, choice)
64 dataset.t <- as.data.frame(dataset.t)
65 dataset.f <- rbind(dataset.f,dataset.t)
66 res <- clogit(choice ~ alt3.cte + Var11 + Var21 + Var31 + Var32 + strata(cs), data =
67 dataset.f)
68 truco <- (summary(res)$coefficients[,1])
69 truco <- as.data.frame((cbind(truco, summary(res)$coefficients[,4])))
70 truco <- cbind(truco,ifelse(abs(truco[,2]) >= 1.96,1,0))
71 truco <- cbind(truco, truco[,1]*truco[,3])
72 priors <- truco[,4]
73 i <- i+1
74 results <- rbind(results, priors)
75 n.it <- "iteration"
76 print("Final iteration")
77 #last iteration out of the loop
78
79 i <- 0
80 x11.fin <- as.data.frame(cbind(x11.fin, results[,2])) #storage of coefficients
```

```
80 x21.fin <- as.data.frame (cbind(x21.fin, results[,3])) #storage of coefficients
81 x31.fin <- as.data.frame(cbind(x31.fin, results[,4])) #storage of coefficients
82 x32.fin <- as.data.frame(cbind(x32.fin, results[,5])) #storage of coefficients
83 dataset.f <- c() #clear dataset
84 results <- c() #clear results array
85 monti <- "Iteration MONTECARLO" # Monte Carlo iteration counter
86 print(paste(monti, montecarlo)) # Monte Carlo iteration counter
87 montecarlo<-montecarlo+1 # Monte Carlo iteration counter
88 }
89
90 library(writexl) #print 4 excel files
91 write_xlsx(x11.fin, sprintf("sim%fx11.xlsx",f))
92 write_xlsx(x21.fin, sprintf("sim%fx21.xlsx",f))
93 write_xlsx(x31.fin, sprintf("sim%fx31.xlsx",f))
94 write_xlsx(x32.fin, sprintf("sim%fx32.xlsx",f))
```

4.2 Errores relativos

Strategy	8 sets	9 sets	10 sets	11 sets	12 sets	13 sets	14 sets	15 sets	16 sets	Mean
1	0.5191	0.5744	0.2342	0.5538	0.455	0.5583	0.5359	0.4501	0.2406	0.4579
2	0.1648	0.5664	0.2331	0.5603	0.4754	0.5607	0.5406	0.4602	0.2262	0.4209
3	0.1077	0.5674	0.2263	0.5567	0.4845	0.5608	0.5423	0.4784	0.1094	0.4037
4	0.2026	0.5667	0.1378	0.5598	0.4832	0.5594	0.5502	0.4972	0.2244	0.4201
5	0.1814	0.5451	0.0892	0.5466	0.5438	0.5489	0.5378	0.4843	0.0926	0.3966
6	0.124	0.5006	0.0564	0.5256	0.5226	0.5335	0.5333	0.5046	0.1894	0.3878
7	0.1315	0.4622	0.1544	0.5155	0.4444	0.5163	0.5054	0.4578	0.1617	0.3721
8	0.0442	0.198	0.0594	0.3719	0.4058	0.425	0.4451	0.4578	0.1181	0.2806
9	0.0681	0.5744	0.1295	0.4456	0.3315	0.4418	0.4092	0.3319	0.1075	0.3155
10	0.0438	0.0407	0.0397	0.0373	0.0323	0.0297	0.0334	0.0299	0.0281	0.035
11	0.0414	0.0429	0.0362	0.037	0.0328	0.0329	0.0331	0.0303	0.0281	0.035
Mean	0.148	0.4217	0.1269	0.4282	0.3828	0.4334	0.4242	0.3802	0.1387	0.3205

Each number represents the mean relative error of every coefficient estimated using the strategy in rows and the number of choice sets in columns.

Tabla 4.2.1. Errores relativos de las estimaciones

Strategy	8 sets	9 sets	10 sets	11 sets	12 sets	13 sets	14 sets	15 sets	16 sets	Mean
1	0.2177	0.7162	0.2275	0.7004	0.5955	0.7075	0.6615	0.5838	0.2413	0.5168
2	0.2425	0.7189	0.245	0.7093	0.5998	0.6995	0.6753	0.6054	0.2322	0.5253
3	0.2093	0.7035	0.338	0.7048	0.6273	0.6947	0.6917	0.6256	0.2397	0.5372
4	0.2056	0.705	0.2743	0.71	0.6342	0.706	0.6903	0.6455	0.2254	0.5329
5	0.1566	0.6374	0.1154	0.6643	0.6632	0.673	0.6687	0.6398	0.1759	0.4883
6	0.6426	0.6453	0.0653	0.6609	0.6636	0.6663	0.6717	0.6332	0.177	0.5362
7	0.1333	0.611	0.1972	0.6137	0.5793	0.6312	0.6293	0.5909	0.1828	0.4632
8	0.0652	0.2683	0.0714	0.4675	0.5101	0.6139	0.5492	0.5625	0.1191	0.3586
9	0.1235	0.545	0.1152	0.5517	0.4576	0.5709	0.5325	0.459	0.1218	0.3864
10	0.0555	0.047	0.047	0.0448	0.0409	0.0412	0.0387	0.0415	0.0415	0.0442
11	0.0507	0.0561	0.0488	0.0466	0.0377	0.043	0.0324	0.0347	0.0287	0.042077778
Mean	0.1911	0.5140	0.1586	0.5340	0.4917	0.5497	0.5310	0.4929	0.1623	0.4028

Each number represents the mean of the relative error of every coefficient incurred by the strategy in rows and the number of choice sets in columns.

Tabla 4.2.2. Errores relativos de las estimaciones (Analis. Sens, + atrib. y niv.)

Strategy	8 sets	9 sets	10 sets	11 sets	12 sets	13 sets	14 sets	15 sets	16 sets	Mean
6	0.1345	0.5191	0.085	0.5221	0.5302	0.5272	0.5359	0.5042	0.1871	0.3974
8	0.061	0.1989	0.0884	0.3661	0.3978	0.42	0.4448	0.4551	0.1196	0.2887
10	0.0761	0.0515	0.0486	0.0572	0.0468	0.0452	0.0475	0.0451	0.0432	0.0473
11	0.065	0.0621	0.051	0.045	0.0464	0.0439	0.0454	0.0508	0.0371	0.0463
Mean	0.08415	0.2079	0.06825	0.2476	0.2553	0.259075	0.2684	0.2638	0.09675	0.194925

Each number represents the mean of the relative error of every coefficient estimated by combining the strategy in rows and the number of choice sets in columns.

Tabla 4.2.3. Errores relativos de las estimaciones (Analis. Sens, - Muestra)

4.3 Sesgo

Strat.	8 sets	9 sets	10 sets	11 sets	12 sets	13 sets	14 sets	15 sets	16 sets
$\beta_{11} = -2$									
1	-3.0780***	-0.8360***	-2.4833***	-0.8705***	-1.0937***	-0.8773***	-0.9203***	-1.0821***	-2.4914***
2	-2.3271***	-0.8609***	-2.4718***	-0.8664***	-1.0339***	-0.8680***	-0.9056***	-1.0673***	-2.4556***
3	-1.9971	-0.8609***	-2.4658***	-0.8763***	-1.0170***	-0.8695***	-0.9100***	-1.0300***	-1.9758
4	-2.4180***	-0.8516***	-1.7089***	-0.8659***	-1.0272***	-0.8704***	-0.8916***	-0.9925***	-2.4521***
5	-1.6091***	-0.8927***	-2.1582***	-0.8937***	-0.8872***	-0.8966***	-0.9183***	-1.0183***	-1.8218***
6	-2.2463***	-0.9772***	-2.0647***	-0.9330***	-0.9352***	-0.9262***	-0.9170***	-0.9757***	-2.3802***
7	-1.7292***	-1.0669***	-2.3142***	-0.9473***	-1.0935***	-0.9449***	-0.9833***	-1.0735***	-1.6461***
8	-2.0020	-1.5791***	-2.1056***	-1.2275***	-1.1660***	-1.1312***	-1.0905***	-1.0609***	-2.2412***
9	-1.8762***	-0.8360***	-2.2618***	-1.0806***	-1.3158***	-1.0859***	-1.1417***	-1.2946***	-1.7596***
10	-2.0258***	-2.0131	-2.0067	-1.9959	-2.0073	-2.0106*	-2.0052	-1.9990	-2.0115**
11	-2.0117	-2.0174*	-1.9973	-2.0129*	-2.0024	-2.0008	-1.9984	-1.9913	-2.0033
$\beta_{21} = -1$									
1	-1.4698***	-0.4189***	-1.2179***	-0.4494***	-0.5576***	-0.4467***	-0.4685***	-0.5643***	-1.2262***
2	-1.1512***	-0.4218***	-1.2234***	-0.4387***	-0.5234***	-0.4359***	-0.4673***	-0.5417***	-1.2164***
3	-1.0342**	-0.4320***	-1.2096***	-0.4431***	-0.5229***	-0.4369***	-0.4611***	-0.5291***	-1.0211
4	-1.1858***	-0.4278***	-0.9304***	-0.4343***	-0.5222***	-0.4393***	-0.4468***	-0.5037***	-1.2088***
5	-0.8590***	-0.4499***	-1.1014***	-0.4492***	-0.4511***	-0.4509***	-0.4632***	-0.5151***	-0.9665***
6	-1.1137***	-0.4964***	-1.0668***	-0.4699***	-0.4706***	-0.4611***	-0.4609***	-0.4946***	-1.1813***
7	-0.8839***	-0.5599***	-1.1464***	-0.4868***	-0.5689***	-0.4865***	-0.4905***	-0.5541***	-0.8732***
8	-1.0001	-0.8263***	-1.0507***	-0.6382***	-0.5981***	-0.5791***	-0.5595***	-0.5463***	-1.1107***
9	-0.9533***	-0.4189***	-1.1206***	-0.5848***	-0.6893***	-0.5872***	-0.6338***	-0.7053***	-0.9178***
10	-1.0181***	-1.0103**	-1.0075	-0.9952	-1.0037	-1.0069*	-1.0021	-0.9993	-1.0052
11	-1.0103*	-1.0108*	-0.9956	-1.0055	-1.0035	-1.0036	-0.9967	-0.9963	-1.0039
$\beta_{31} = 1$									
1	1.5558***	0.4380***	1.2408***	0.4649***	0.5541***	0.4459***	0.4743***	0.5511***	1.2513***
2	1.1788***	0.4444***	1.2423***	0.4581***	0.5429***	0.4560***	0.4688***	0.5528***	1.2399***
3	1.0438***	0.4435***	1.2366***	0.4616***	0.5247***	0.4590***	0.4634***	0.5287***	1.0326**
4	1.2175***	0.4479***	0.8780***	0.4603***	0.5219***	0.4536***	0.4641***	0.5138***	1.2431***
5	0.8621***	0.4741***	1.0920***	0.4721***	0.4830***	0.4636***	0.4737***	0.5303***	0.9549***
6	1.1287***	0.5208***	1.0418***	0.4979***	0.5048***	0.4861***	0.4894***	0.5163***	1.2001***
7	0.8839***	0.5138***	1.1580***	0.4968***	0.5584***	0.4991***	0.5042***	0.5456***	0.8680***
8	1.0011	0.7941***	1.0573***	0.6452***	0.6122***	0.5877***	0.5683***	0.5605***	1.1288***
9	0.9554***	0.4380***	1.1348***	0.5587***	0.6740***	0.5557***	0.5850***	0.6675***	0.9024***
10	1.0149**	1.0120**	1.0090	1.0002	1.0018	1.0068	0.9985	1.0019	1.0042
11	1.0014	1.0141**	0.9948	1.0037	1.0093*	0.9995	0.9973	0.9992	0.9973
$\beta_{32} = 2$									
1	3.0240***	0.8547***	2.4732***	0.8705***	1.0684***	0.8706***	0.9068***	1.0864***	2.4782***
2	2.3134***	0.8754***	2.4617***	0.8573***	1.0309***	0.8623***	0.8972***	1.0619***	2.4411***
3	1.9440**	0.8489***	2.4524***	0.8605***	1.0120***	0.8523***	0.9026***	1.0277***	1.9353**
4	2.3961***	0.8637***	1.7053***	0.8665***	1.0192***	0.8690***	0.8851***	0.9948***	2.4394***
5	1.5594***	0.8988***	2.1422***	0.8905***	0.8941***	0.8831***	0.9057***	1.0164***	1.7667***
6	2.2523***	0.9838***	2.0498***	0.9269***	0.9333***	0.9114***	0.9161***	0.9657***	2.3724***
7	1.7007***	1.0881***	2.3116***	0.9618***	1.0967***	0.9536***	0.9839***	1.0650***	1.5889***
8	2.0016	1.5964***	2.1059***	1.2308***	1.1668***	1.1349***	1.0936***	1.0630***	2.2242***
9	1.8696***	0.8547***	2.2609***	1.0675***	1.3061***	1.0936***	1.1478***	1.3048***	1.7411***
10	2.0283***	2.0150*	2.0057	1.9948	2.0048	2.0086	2.0143*	1.9963	2.0010
11	2.0091	2.0163*	1.9999	2.0101	2.0076	2.0060	1.9979	1.9939	2.0062

Each number represents the mean of each coefficient's estimate (for the combination of the strategy in rows, number of choice sets in columns, and the coefficient in the top bar). The asterisks represent the result of a *t*-test comparing the estimates' mean with their target coefficient (***: $\hat{\beta} \neq \beta$ with $\alpha = 0.01$, **: $\hat{\beta} \neq \beta$ with $\alpha = 0.05$, *: $\hat{\beta} \neq \beta$ with $\alpha = 0.1$).

Table 4.3.1. Media y t-test de las estimaciones

		$\beta = -2$								
		8	9	10	11	12	13	14	15	16
1	-	2.4323***	-0.6727***	-2.4571***	-0.6327***	-0.8696***	-0.6078***	-0.7036***	-0.8591***	-2.4920***
2	-	2.4463***	-0.5877***	-2.4584***	-0.5774***	-0.8123***	-0.6067***	-0.6505***	-0.7919***	-2.4757***
3	-	2.4250***	-0.5727***	-1.3849***	-0.5990***	-0.7462***	-0.5930***	-0.6190***	-0.7710***	-2.4693***
4	-	2.4259***	-0.5557***	-1.4946***	-0.5580***	-0.7495***	-0.5900***	-0.6112***	-0.7180***	-2.4633***
5	-	2.2990***	-0.7293***	-1.8671***	-0.6677***	-0.6678***	-0.6511***	-0.6641***	-0.7303***	-2.3908***
6	-	0.7166***	-0.7059***	-2.0145	-0.6749***	-0.6697***	-0.6586***	-0.6528***	-0.7322***	-2.3992***
7	-	2.2394***	-0.8287***	-2.3279***	-0.7961***	-0.8714***	-0.7268***	-0.7357***	-0.8835***	-2.3802***
8	-	-2.0234	-1.4744***	-2.1234***	-1.0742***	-0.9791***	-0.9296***	-0.8960***	-0.8862***	-2.2429***
9	-	2.2514***	-0.9457***	-2.2128***	-0.9536***	-1.1668***	-0.8896***	-0.9748***	-1.1363***	-2.2676***
10	-	-2.0220	-2.0264	-2.0462**	-1.9971	-2.0206	-2.0002	-1.9913	-2.0135	-2.0324**
11	-	-2.0356**	-2.0411*	-2.0216	-2.0389**	-2.020	-1.992	-1.996	-2.001	-2.034**
		$\beta = -1$								
		8	9	10	11	12	13	14	15	16
1	-	1.1972***	-0.3920***	-1.2380***	-0.3323***	-0.4806***	-0.3550***	-0.3763***	-0.4636***	-1.2313***
2	-	1.2202***	-0.3268***	-1.2170***	-0.3019***	-0.4417***	-0.3315***	-0.3377***	-0.4444***	-1.2326***
3	-	1.1987***	-0.3163***	-0.7271***	-0.3264***	-0.4104***	-0.3213***	-0.3298***	-0.4358***	-1.2304***
4	-	1.2202***	-0.3268***	-1.2170***	-0.3019***	-0.4417***	-0.3315***	-0.3377***	-0.4444***	-1.2326***
5	-	1.1438***	-0.4087***	-0.9538***	-0.3771***	-0.3696***	-0.3633***	-0.3588***	-0.3978***	-1.1932***
6	-	0.3854***	-0.3970***	-1.0289***	-0.3780***	-0.3744***	-0.3591***	-0.3595***	-0.4072***	-1.1921***
7	-	1.1336***	-0.4408***	-1.1671***	-0.4779***	-0.4859***	-0.4257***	-0.4087***	-0.5084***	-1.1790***
8	-	-1.0210	-0.7779***	-1.0696***	-0.5862***	-0.5326***	-0.5266***	-0.5001***	-0.5052***	-1.1252***
9	-	1.0894***	-0.5417***	-1.0981***	-0.5197***	-0.5850***	-0.4996***	-0.5244***	-0.6299***	-1.1375***
10	-	-1.0184	-1.0319	-1.0053	-1.0283	-1.0158	-1.0135	-1.0020	-1.0021	-1.0077**
11	-	-1.0267**	-1.0276*	-1.0152	-1.0117	-0.9962	-1.0199*	-0.9844*	-1.0208	-1.0069
		$\beta = 1$								
		8	9	10	11	12	13	14	15	16
1	-	1.2266***	0.2731***	1.2164***	0.2704***	0.4040***	0.2685***	0.3011***	0.4053***	1.2371***
2	-	1.2210***	0.2535***	1.2295***	0.2576***	0.3576***	0.2380***	0.2896***	0.3588***	1.2307***
3	-	1.1962***	0.2406***	0.6292***	0.2394***	0.3465***	0.2555***	0.2498***	0.3315***	1.2364***
4	-	1.2071***	0.2443***	0.6955***	0.2535***	0.3265***	0.2528***	0.2675***	0.3071***	1.2234***
5	-	1.1485***	0.3232***	0.8984***	0.2959***	0.2982***	0.2855***	0.2980***	0.3298***	1.1803***
6	-	0.3150***	0.3189***	0.9811**	0.2975***	0.3018***	0.2919***	0.2846***	0.3230***	1.1947***
7	-	1.1384***	0.3554***	1.1673***	0.3448***	0.3791***	0.2980***	0.3237***	0.3238***	1.1671***
8	-	1.0094	0.7159***	1.0537***	0.5044***	0.4575***	0.4309***	0.4179***	0.3915***	1.1235***
9	-	1.1491***	0.4112***	1.1085***	0.3891***	0.4851***	0.3978***	0.4392***	0.5205***	1.1300***
10	-	1.0307**	1.0186	1.0269	1.0093	1.0135	0.9989	0.9999	0.9997	1.0154
11	-	1.0442**	1.0286*	0.9968	1.0319***	0.9821	1.0243	0.9909	1.0218*	1.0272**
		$\beta = 2$								
		8	9	10	11	12	13	14	15	16
1	-	2.4176***	0.7720***	2.4633***	0.6781***	0.9193***	0.7085***	0.7669***	0.9201***	2.4952***
2	-	2.4336***	0.6883***	2.4543***	0.6524***	0.8521***	0.6686***	0.7060***	0.8328***	2.4728***
3	-	2.4107***	0.6499***	1.3571***	0.6421***	0.7863***	0.6559***	0.6874***	0.7961***	2.4629***
4	-	2.3958***	0.6580***	1.4812***	0.6434***	0.7885***	0.6642***	0.6754***	0.7731***	2.4569***
5	-	2.3014***	0.7784***	1.8504***	0.7386***	0.7143***	0.7125***	0.7123***	0.7800***	2.3826***
6	-	0.7603***	0.7658***	2.0076	0.7178***	0.7183***	0.7121***	0.7091***	0.7786***	2.3987***

7	2.2108***	0.8107***	2.3254***	0.8235***	0.8596***	0.7753***	0.7985***	0.9115***	2.3626***
8	2.0412	1.4896***	2.1539***	1.0957***	0.9943***	0.9667***	0.9279***	0.9178***	2.2542***
9	2.2491***	0.9092***	2.2089***	0.9501***	1.1448***	0.8636***	0.9609***	1.1070***	2.2462***
10	2.0393	2.0307*	2.0225	1.9955	2.0340*	1.9902	2.0025	2.0013	2.0308*
11	2.0460**	2.0461	2.0366*	2.0287	2.0320	1.9976	1.9924	2.0105	2.0296
$\beta = 0.5$									
	8	9	10	11	12	13	14	15	16
1	0.6103***	0.1100***	0.6101***	0.1371***	0.1804***	0.0762***	0.1407***	0.1717***	0.6212***
2	0.6059***	0.1565***	0.6102***	0.1000***	0.1668***	0.1154***	0.1353***	0.1566***	0.6165***
3	0.6012***	0.1128***	0.2930***	0.1028***	0.1437***	0.1355***	0.1215***	0.1352***	0.6175***
4	0.5949***	0.1217***	0.3431***	0.1083***	0.1352***	0.1093***	0.1325***	0.1358***	0.6074***
5	0.5489***	0.1313***	0.4478***	0.1213***	0.1390***	0.1216***	0.1360***	0.1315***	0.5895***
6	0.1537***	0.1276*	0.4842***	0.1349***	0.1298***	0.1448***	0.1311***	0.1553***	0.5938***
7	0.5688***	0.1527***	0.5748***	0.1186***	0.1957***	0.1914***	0.1634***	0.1328***	0.6051***
8	0.5090	0.3483***	0.5375***	0.2346***	0.2257***	0.2018***	0.1965***	0.1808***	0.5712***
9	0.5603***	0.2152***	0.5674***	0.1918***	0.2265***	0.2007***	0.2133***	0.2177***	0.5426***
10	0.5253	0.5004	0.5046	0.4953	0.5043	0.4947	0.5190*	0.4883	0.5214**
11	0.5066	0.5323**	0.5073	0.5042	0.5150	0.5100	0.4869	0.5092	0.5029
$\beta = 1.3$									
	8	9	10	11	12	13	14	15	16
1	1.5675***	0.4554***	1.5835***	0.4614***	0.5484***	0.4083***	0.4514***	0.5563***	1.6055***
2	1.5703***	0.3849***	1.5813***	0.4101***	0.5371***	0.4112***	0.4391***	0.5215***	1.6037***
3	1.5673***	0.4240***	0.8474***	0.4189***	0.4909***	0.4147***	0.4233***	0.4988***	1.6060***
4	1.5754***	0.3940***	0.9127***	0.4050***	0.4798***	0.3892***	0.4067***	0.4649***	1.5960***
5	1.4970***	0.4932***	1.1734***	0.4540***	0.4426***	0.4502***	0.4336***	0.4835***	1.5447***
6	0.4631***	0.4826*	1.2833***	0.4510***	0.4477***	0.4395***	0.4442***	0.4779***	1.5560***
7	1.4465***	0.4591***	1.5025***	0.5346***	0.4922***	0.4260***	0.4883***	0.5430***	1.5367***
8	1.3132	0.9153***	1.3851***	0.6782***	0.6300***	0.2018***	0.5977***	0.5600***	1.4596***
9	1.4540***	0.5496***	1.4502***	0.5357***	0.7111***	0.5096***	0.5685***	0.6881***	1.4670***
10	1.2985	1.3143	1.3189	1.3016	1.3198	1.3017	1.2979	1.3037	1.3336***
11	1.3215	1.3414**	1.3206	1.3088	1.3117	1.3352*	1.2845	1.3076	1.3183
$\beta = 2.5$									
	8	9	10	11	12	13	14	15	16
1	3.0467***	0.9651***	3.0633***	0.9568***	1.1115***	0.9060***	0.9279***	1.1365***	3.1135***
2	1.5703***	0.3849***	1.5813***	0.8713***	1.0979***	0.8735***	0.9174***	1.0902***	3.0896***
3	3.0233***	0.8895***	1.7248***	0.8815***	1.0510***	0.8637***	0.8913***	1.0402***	3.0661***
4	3.0122***	0.8811***	1.8749***	0.8544***	1.0413***	0.8610***	0.8936***	0.9963***	3.0690***
5	2.8614***	1.0275***	2.3314***	0.9544***	0.9514***	0.9211***	0.9218***	1.0093***	2.9841***
6	0.9989***	1.0116***	2.5142	0.9632***	0.9520***	0.9299***	0.9234***	1.0092***	2.9873***
7	2.8268***	1.1215***	1.5025***	1.0576***	1.1518***	0.9926***	0.9824***	1.1271***	2.9546***
8	2.5064	1.8638***	2.6588***	1.4026***	1.2919***	0.5955***	1.1816***	1.1775***	2.8113***
9	2.7913***	1.1283***	2.7665***	1.2037***	1.4272***	1.0906***	1.1922***	1.3811***	2.8165***
10	2.5484*	2.5355	2.5448**	2.5097	2.5277	2.5050	2.5041	2.5207	2.5278**
11	2.5614**	2.5321	2.5582**	2.5445*	2.5214	2.5141	2.4841	2.5262	2.5464**

Each number represents the mean of each coefficient's estimate (for the combination of the strategy in rows, number of choice sets in columns, and the coefficient in the top bar). The asterisks represent the result of a t-test comparing the estimates' mean with their target coefficient (***: $\beta \neq \beta$ with $\alpha=0.01$, **: $\beta \neq \beta$ with $\alpha=0.05$, *: $\beta \neq \beta$ with $\alpha=0.1$).

Table 4.3.2. T-test de los resultados del análisis de sensibilidad (1)

$\beta = -2$									
	8	9	10	11	12	13	14	15	16
6	-2.2568***	-0.9687***	-1.9283	-0.9331***	-0.9239***	-0.9249***	-0.9222***	-0.9718***	-2.3676***
8	-1.9891	-1.5712***	-2.1498***	-1.2382***	-1.1809***	-1.1464***	-1.0985***	-1.0635***	-2.2528***
10	-2.0718**	-2.0453**	-2.0107	-2.0200	-2.0101	-2.0266	-2.0167	-2.0287	-2.0080
11	-2.0246	-2.0194	-2.0108	-2.0475**	-2.0130	-2.0292	-1.9837	-2.0365	-2.0085
$\beta = -1$									
	8	9	10	11	12	13	14	15	16
6	-1.1430***	-0.4719***	-1.0018	-0.4789***	-0.4695***	-0.4712***	-0.4630***	-0.4944***	-1.1765***
8	-0.9892	-0.8302***	-1.0826***	-0.6364***	-0.6163***	-0.5788***	-0.5535***	-0.5488***	-1.1101***
10	-1.0184	-1.0319**	-1.0053	-1.0283	-1.0158	-1.0135	-1.0094	-1.0217**	-1.0077
11	-1.0081	-1.0107	-0.9930	-0.9997	-1.0016	-1.0169	-1.0103	-1.0026	-1.0016
$\beta = 1$									
	8	9	10	11	12	13	14	15	16
6	1.1122***	0.4881***	0.9776	0.5082***	0.4786***	0.4819***	0.4884***	0.5203***	1.2021***
8	1.0294	0.7968***	1.0860***	0.6622***	0.6160***	0.5932***	0.5702***	0.5663***	1.1276***
10	1.0286	1.0238	0.9705**	1.0146	1.0117	1.0022	1.0356***	1.0192*	0.9948
11	1.0177	1.0072	0.9940	1.0239	0.9868	0.9984	0.9941	1.0229	0.9739*
$\beta = 2$									
	8	9	10	11	12	13	14	15	16
6	2.2711***	0.9643***	1.9129*	0.9153***	0.9408***	0.9343***	0.9062***	0.9744***	2.3673***
8	1.9828	1.5674***	2.1287***	2.0234***	2.0040***	2.0141***	2.0105***	2.0227***	2.0110***
10	2.0485	2.0506	2.0101	2.0172	1.9975	2.0484**	2.0095	2.0106	2.0034
11	2.0311	2.0305	2.0089	2.0234	2.0040	2.0141	2.0105	2.0227	2.0110

Each number represents the mean of each coefficient's estimate (for the combination of the strategy in rows, number of choice sets in columns, and the coefficient in the top bar). The asterisks represent the result of a t-test comparing the estimates' mean with their target coefficient (***: $\beta \neq \beta$ with $\alpha=0.01$, **: $\beta \neq \beta$ with $\alpha=0.05$, *: $\beta \neq \beta$ with $\alpha=0.1$).

Table 4.3.3. T-test de los resultados del análisis de sensibilidad (2)



APÉNDICE TEMA 5. MATERIAL SUPLEMENTARIO

Repositorio GitHub: <https://github.com/danielpereztr/DCEtool>

Repositorio CRAN: <https://cran.r-project.org/package=DCEtool>

Guía (inglés): <https://danielpereztr.github.io/posts/DCEtool/>