



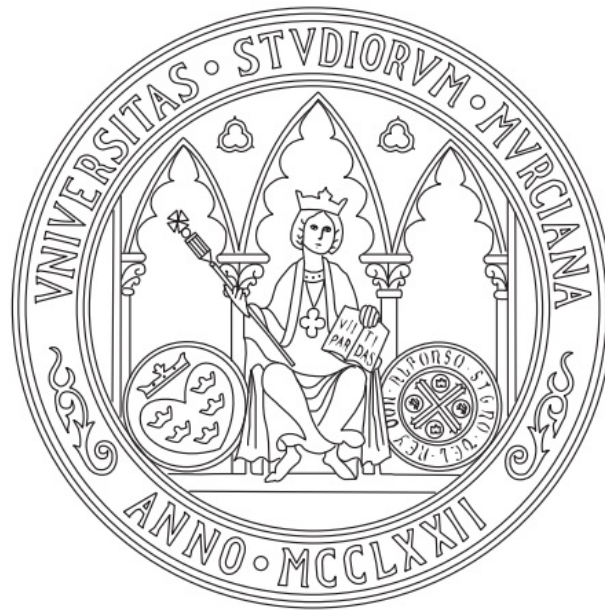
UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

Las preferencias de los turistas rurales y los factores decisivos en el proceso de elección de alojamiento

D^a. María Teresa Díaz Delfa

2021



UNIVERSIDAD DE MURCIA

ESCUELA INTERNACIONAL DE DOCTORADO

Las preferencias de los turistas rurales y los factores decisivos en el proceso de elección de alojamiento

D^a. María Teresa Díaz Delfa

Directora:

Prof. Dr. M^a Isabel Pilar Albaladejo Pina

2021

Agradecimientos

A la primera persona que quiero agradecer que por fin pueda presentar la tesis es, sin duda alguna, a Isabel Pilar Albaladejo Pina, mi directora de tesis y, sobre todo, amiga. Sin ella este final no habría sido posible. Isabel te agradezco infinitamente el apoyo brindado durante todos estos años, tu ayuda y que no me dejaras “tirar la toalla”. Muchas gracias.

En segundo lugar, a mi compañera y amiga Fuensanta Arnaldos. Ella mejor que nadie sabe la angustia y los quebraderos de cabeza que me ha supuesto el pensar que este momento nunca llegaría y, al final, a pesar de mí, ha sido posible y he logrado el objetivo tan ansiado de terminar la tesis. Ella ha sido mi apoyo y mi paño de lágrimas. Fuen, todo llega en la vida. Muchas gracias.

Por último, y no por eso menos importante, a mi familia, amigos y compañeros. A mi marido Gabriel por aguantar mi mal humor y los días de trabajo interminables que nos impedían hacer otras cosas. Gracias por cuidar de nuestros hijos. A mis hijos, Laura y Raúl, por ser como sois y enseñarme tantas cosas. A ti papá, por estar siempre ahí, en los momentos buenos, pero principalmente en los malos. A mi madre, siempre cuidándonos a todos sin pedir nada a cambio. En fin, a todas la personas que han estado conmigo durante este periplo y me han ayudado de alguna forma y que sería imposible nombrar. Muchas gracias a todos.

*A mis amores, mi marido y mis hijos,
Laura y Raúl*

Índice general

Lista de Tablas	VI
Lista de Figuras	VII
Introducción	1
1. Modelización de la elección de alojamiento rural	7
1.1. Introducción	7
1.2. Alojamientos rurales	9
1.3. Modelos de elección discreta	12
1.4. Datos de preferencias reveladas y declaradas	19
1.5. Encuesta en la Comarca Noroeste de la Región de Murcia . . .	22
1.6. Conclusiones	30
2. Demanda de turismo rural por tipo de alojamiento	33
2.1. Introducción	33
2.2. Construcción del conjunto de elección	35
2.2.1. Análisis cluster	36
2.2.2. Análisis de clases latentes	37
2.3. Datos	40
2.4. Clasificación de los alojamientos rurales	42
2.4.1. Análisis cluster	42
2.4.2. Análisis de clases latentes	49
2.4.3. Comparación de los resultados obtenidos con ambos métodos de clasificación	52
2.5. Perfil del turista por tipo de alojamiento	53
2.6. Conclusiones	61
3. Preferencias heterógenas de los turistas rurales ante la elección de alojamiento	63
3.1. Introducción	63

3.2. Modelos de elección con preferencias heterogéneas	65
3.2.1. Heterogeneidad condicional	66
3.2.2. Heterogeneidad aleatoria	67
3.3. Disposiciones a pagar	70
3.4. Investigación empírica	71
3.4.1. Datos y variables	71
3.4.2. Modelos estimados	71
3.4.3. Resultados de las estimaciones	74
3.5. Conclusiones	79
4. Los efectos de las motivaciones de los turistas rurales sobre la elección de alojamiento	81
4.1. Introducción	81
4.2. Las motivaciones de los turistas y la elección del alojamiento rural	84
4.3. Modelos híbridos de elección discreta	85
4.4. Variables latentes y ecuaciones estructurales	88
4.5. Datos y variables	91
4.6. Definición y estimación del modelo MIMIC	94
4.6.1. Definición de las variables latentes	94
4.6.2. Validación del instrumento de medida	95
4.6.3. Estimación del modelo MIMIC	99
4.7. Definición y estimación de los modelos de elección	102
4.8. Conclusiones	108
Conclusiones	111
Bibliografía	121
Anexos	137
Anexo I. Cuestionario	139
Anexo II. Programación en SAS para la obtención de las alternativas y los conjuntos de elección del experimento de preferencias declaradas	151
Anexo III. Resultados de la estimación del modelo LAC	155
Anexo IV. Comparativa de los modelos logit LACL y LAC	159

Lista de tablas

1.1. Modelos familia Logit	18
1.2. Perfil de los encuestados	25
1.3. Descripción de los indicadores	26
1.4. Atributos y niveles del diseño experimental	28
2.1. Descripción de los alojamientos	43
2.2. Análisis de correspondencia múltiple	44
2.3. Porcentaje de casas por cluster según el nivel de cada variable	48
2.4. BIC según número de clases latentes	50
2.5. Porcentaje (estimado) de casas por clase latente según el nivel de cada variable	51
2.6. Número de alojamientos por cluster según clase latente	52
2.7. Variables y sus modalidades	56
2.8. Modelo estimado con las cuatro clases obtenidas con el análisis de clases latentes (modelo LACL)	59
2.9. Modelo LACL con las Clases 1, 2 y 3 obtenidas con el análisis de clases latentes (modelo LACL)	60
3.1. Variables del modelo	72
3.2. Resultados de las estimaciones de los modelos logit	75
3.3. Disponibilidades a pagar	78
4.1. Características de los individuos incluidas en el modelo MIMIC	92
4.2. Variables de los modelos HDC y del logit multinomial	93
4.3. Análisis factorial exploratorio (AFE)	95
4.4. Fiabilidad y validez convergente del instrumento de medida	98
4.5. Validez discriminante del instrumento de medida	98
4.6. Estimación del modelo MIMIC	100
4.7. Resultados de las estimaciones de los modelos Logit	106
8. Modelo estimado con cuatro grupos determinados mediante el análisis cluster (modelo LAC)	156

9.	Modelo LCA con tres grupos	157
10.	Características socio-económicas de los turistas significativas y signo, según modelo	160
11.	Condiciones de viaje del turista significativas y signo, según modelo	161

Lista de Figuras

1.1. Mapa de la Región de Murcia [†]	23
1.2. Resultados exploratorios sobre las motivaciones para realizar turismo rural	26
1.3. Ejemplo de escenario de elección	31
2.1. Distribución de los alojamientos rurales estudiados	41
2.2. Representación gráfica de las variables nominales en el espacio	45
2.3. Contribuciones de las categorías de las variables a las dimensiones	46
4.1. Valoración de los indicadores	93
4.2. Estructura (gráfico de ruta) del modelo MIMIC	101

Introducción

El proceso de decisión llevado a cabo por un individuo ante la posibilidad de realizar un viaje está constituido por una serie de pequeñas decisiones o subdecisiones como adónde ir, cuándo ir, qué hacer durante las vacaciones y cómo llegar (Björk y Janson, 2008). Entre ellas, la elección de alojamiento es de las más importantes para los turistas, tanto desde la perspectiva económica (Sharpley, 2000 ; Chen et al., 2017) como por su valor en la experiencia turística. Según Liu y Ma (2017), el tipo y la calidad del alojamiento pueden tener un impacto en la experiencia de viaje y en el comportamiento posterior del turista. McIntosh y Siggs (2005) añaden que el alojamiento constituye un elemento esencial para determinar la satisfacción y relevancia de la experiencia turística que tiene el turista. Además, Kim y Park (2017) aseguran que conocer y entender la elección del alojamiento es una ayuda importante para los gestores y propietarios de los alojamientos en relación con una gestión eficaz y adecuada de las decisiones de inversión. Esta ayuda puede ser aún mayor para el caso de los alojamientos rurales, que a menudo carecen de recursos para promocionarse adecuadamente y tienen dificultades para adaptarse a los mecanismos del mercado actual (Albaladejo y Díaz, 2005).

En los destinos rurales existen pocos alojamientos tradicionales como hoteles. La mayor parte de la capacidad de alojamiento de estos destinos se cubre básicamente mediante pequeños hoteles rurales, B&B, casas de alquiler y casas de huéspedes (Ye et al., 2019; Albaladejo y Díaz, 2005). Estos pequeños negocios, tradicionalmente de carácter familiar, surgieron como una medida para el desarrollo del turismo rural, el alivio de la pobreza y la revitalización de las zonas rurales (Dernoi, 1991; Fleischer y Felsenstein, 2000 ; Valdés y Del Valle, 2003; Yagüe, 2002; Komppula, 2014). Pero, el número de alojamientos rurales como el de consumidores de este tipo de producto han crecido rápida y continuamente en España durante las últimas décadas (Instituto Nacional de Estadística, www.ine.es). Este crecimiento se puede explicar por cambios en el comportamiento de los consumidores de turismo y por la oferta de nuevos productos que permiten una mayor independencia y

flexibilidad, a la vez que satisfacen las necesidades y deseos inmediatos de los consumidores (Aguiló y Juaneda, 2000; Hummelbrunner y Milbauer, 1994; Yagüe, 2002).

La creciente demanda de los alojamientos y el auge de las plataformas de comercio electrónico y de economía colaborativa (por ejemplo, Airbnb), como apoyo para gestionar las reservas, han hecho de este tipo de alojamientos un producto totalmente consolidado en el mercado turístico que representa la oferta básica del turismo rural. A ello hay que añadir que durante la pandemia COVID-19, el turismo rural ha sido el sector turístico menos perjudicado a pesar de lo que está sufriendo este sector. Las restricciones existentes para los viajes y las reuniones han hecho que los turistas opten por viajar a entornos rurales y de naturaleza cercanos a su residencia habitual para disfrutar de sus vacaciones en este tipo de alojamientos normalmente independientes y aislados unos de otros. Todo ello indica que los alojamientos rurales tienen una serie de características que los distinguen de otros tipos de alojamiento y que son el resultado de una filosofía inherente a este tipo de turismo. En este contexto, consideramos valioso un estudio detallado sobre la elección de alojamiento rural.

El propósito de este trabajo es conocer las preferencias de los turistas en cuanto a los alojamientos rurales, el valor que le otorgan a las características y a los servicios que ofertan los alojamientos, así como la influencia que en estas elecciones tienen las características, sentimientos y motivaciones de los turistas. Para ello, consideraremos el alojamiento rural como un producto caracterizado por diferentes atributos físicos existentes o potenciales, que le darán un valor y establecerán una interconexión con el comportamiento del decisor, así como con sus sentimientos y motivaciones. Los modelos de elección discreta constituirán la teoría formal para la estimación de las preferencias de elección y la evaluación cuantitativa de los atributos de los alojamientos y de las características de los turistas. El primero que propuso esta modelización de elección discreta fue McFadden (1974) con el popular modelo logit multinomial (MNL). Este modelo, que combina la maximización de utilidad aleatoria y la evaluación hedónica de las alternativas, obtiene las probabilidades individuales de elección entre diferentes alternativas, que en nuestro caso serán los alojamientos. El modelo MNL es muy popular por su tratabilidad, pero ha sido criticado porque verifica la conocida propiedad de independencia de las alternativas irrelevantes (IIA) (Train, 2003) y por su limitación para captar las diferencias de gustos entre los individuos (Espino et al., 2008; Train, 2003). Así pues, se han desarrollado otros modelos de la “familia Logit” dirigidos a relajar estas restricciones, mientras se man-

tiene su tratabilidad, como el modelo logit heterogéneo, el modelo de Logit Mixto (ML) (Train, 2003) y el modelo híbrido de elección discreta (HDC) (Ben-Akiva, McFadden et al., 2002).

Los modelos de elección discreta han desempeñado y están desempeñando un papel cada vez más importante en la identificación de los atributos que determinan las elecciones y en la modelización del comportamiento de elección de los individuos en el turismo, la hospitalidad y el ocio (Crouch y Louviere, 2000). Se han aplicado principalmente para analizar la elección de destinos (Eymann y Ronning, 1997 ; Morley, 1994 ; Seddighi y Theocharous, 2002), pero también para estudiar las preferencias en cuanto al modo de viaje (Kelly et al., 2007) o sobre las atracciones culturales o patrimoniales de un destino turístico (Apostolakis y Jaffry, 2005a, 2005b), para determinar las tasas de los parques (Mmopelwa et al., 2007) o los servicios recibidos en las oficinas de información turística (Araña et al., 2016). También se pueden encontrar varias aplicaciones en el sector de la hostelería, particularmente en la elección del hotel (Martín et al., 2018; Chen et al., 2017 ; Román y Martín, 2016; Victorino et al., 2005) y en la elección de la habitación de hotel (Masiero et al., 2015; Masiero et al., 2016). En este trabajo se pone en valor el uso de los modelos de elección discreta en la elección de alojamiento rural tanto desde una perspectiva teórica como aplicada, poniendo de manifiesto que se trata de un producto claramente diferenciado del hotel.

En esta memoria, la investigación sobre las elecciones de alojamiento se basa en la utilización de datos de preferencias reveladas (PR) y declaradas (PD). Los datos de PR proporcionan información sobre la elección real de alojamiento, mientras que los experimentos de elección declarada se usan para replicar, como un experimento controlado, el procedimiento de elección que implícitamente siguen los turistas bajo escenarios hipotéticos. En cada escenario, los alojamientos se describen en términos de los diferentes niveles de sus atributos. Los datos para las aplicaciones prácticas incluidas en este trabajo se obtuvieron mediante de un experimento de elección, incorporado en una encuesta más amplia, elaborada ad-hoc y que fue realizada en la Comarca del Noroeste de la Región de Murcia durante varios fines de semana en otoño de 2003. Somos conscientes de que la encuesta tiene más de 15 años, pero los datos necesarios para mostrar las intenciones conductuales para ir al entorno rural a elegir un alojamiento son complejos y difíciles de obtener. Además, no sólo necesitábamos datos individuales de los turistas que nos permitieran recopilar información sobre el perfil de los encuestados y sus condiciones de viaje, sino que también pretendíamos identificar sus motivaciones para ir al entorno rural y averiguar las preferencias de los turistas en

relación con los atributos de los alojamientos que influían en la elección del alojamiento rural. Esta encuesta nos permitía cubrir todos estos aspectos.

Este estudio consta de cuatro capítulos además de esta Introducción y de las Conclusiones, en los que se pretende hacer un estudio en profundidad del proceso de elección de alojamiento rural. El primer capítulo se centra en abordar los conceptos, modelos y datos que son comunes a lo largo de todo el trabajo. Se define el concepto de alojamiento rural y se analiza su evolución a lo largo de los años, especialmente en el caso de España. A continuación, se presentan los aspectos más relevantes relativos a la modelización de la elección de alojamiento rural basada en el empleo de modelos de elección discreta. Se analizan tanto sus fundamentos microeconómicos como econométricos, prestando especial atención al modelo logit multinomial, que es el más sencillo, pero constituye la base de los que posteriormente se utilizarán en este trabajo. También se analizan las diferentes fuentes de datos empleadas por estos modelos, en especial la modelización con preferencias declaradas. Y por último, se explica de forma detallada el desarrollo del experimento de elección y cómo se llevó a cabo la encuesta que se realizó en la Comarca del Noroeste de la Región de Murcia y que constituye la fuente de datos para los diferentes estudios que componen este trabajo.

En el Capítulo 2 se pretende determinar el perfil turístico que se corresponde con cada uno de los diferentes tipos de alojamientos que existen en un determinado destino de turismo rural. Para ello, primero es necesario clasificar todos los alojamientos de un determinado destino en un número reducido de grupos, de modo que cada grupo constituya un tipo. Estos grupos se construyen utilizando dos métodos de agrupación diferentes, un análisis cluster precedido de un análisis de correspondencia múltiple o un análisis de clases latentes. Una vez clasificados los alojamientos, se lleva a cabo la estimación de un modelo logit multinomial para determinar las características de los turistas que con mayor probabilidad optaran por cada tipo de alojamiento. Este proceso se aplica a los datos obtenidos a la encuesta que se realizó en la Comarca del Noroeste de la Región de Murcia y los resultados indican que la gran variedad de alojamientos en cuanto a tamaño y tipo es una forma adecuada de atraer a personas de diferentes perfiles.

En el Capítulo 3 se muestra que, aunque el principal atractivo de las casas rurales para el turista se encuentra en el entorno natural en el que se sitúan junto con su carácter rural, existen otros factores físicos de estos alojamientos que también son importantes en su elección, como su tamaño, tipo de edificación, calidad de equipamiento, servicios y actividades que ofrece.

Para comprobar este hecho y a su vez evaluar las preferencias de los turistas por estos atributos físicos potenciales o efectivos, se utilizan los modelos de elección discreta. Se estimaron diferentes especificaciones logit con los datos de preferencia declarada obtenidos del experimento de elección que se llevo a cabo en la Comarca Noroeste de la Región de Murcia. Se comprobó que los modelos logit heterogéneo y logit mixto, que consideran variaciones en los gustos dentro de las preferencias de los turistas, son especificaciones más consistentes que el logit multinomial. Con este capítulo se ofrece un medio para evaluar la idoneidad de ciertos modelos de elección discreta en la elección de alojamiento rural que proporcionan ajustes mejorados sobre el logit y que muestran una gran heterogeneidad en las preferencias de los turistas. También se pone de manifiesto la utilidad de los experimentos de elección para estimar este tipo de modelos. Además, los resultados pueden ser de gran utilidad para los promotores y propietarios de casas rurales a la hora de planificar campañas de marketing y promoción o al tomar decisiones de inversión efectivas.

En el Capítulo 4, tomando como base la teoría del proceso de elección constructiva del consumidor, se propone que la elección de alojamiento rural depende también de las motivaciones de los turistas para ir al medio rural. No todos los turistas que acceden al medio rural tienen las mismas motivaciones o buscan las mismas satisfacciones (Frochot, 2005; Molera y Albaladejo, 2007; Kastenholz et al., 1999; entre otros). La consideración de las motivaciones supone una dificultad, dado que a diferencia de las variables cognitivas, éstas son variables latentes, esto es, variables no observables, muy difíciles de medir y con muchas dimensiones. Los modelos tradicionales de elección discreta no permiten la estimación de este tipo de variables. Un modelo que en su formalización admite variables observables (atributos de los alojamientos) y latentes para analizar la elección de alojamiento es el modelo híbrido de elección discreta (HDC). Este modelo amplía los modelos de elección discreta incorporando el efecto de las variables latentes a través de un modelo de ecuaciones estructurales denominado modelo de múltiples indicadores y múltiples causas (MIMIC) (Ben-Akiva, Walker et al., 2002; Bolduc y Daziano, 2010; Bahamonde-Birke et al., 2010; Bahamonde-Birke et al., 2017; Márquez et al., 2020, entre otros). En este capítulo se lleva a cabo una aplicación de este proceso de elección constructivo para estimar un modelo logit multinomial y dos especificaciones del modelo HDC, usando los datos del experimento de elección realizado en la Comarca Noroeste de la Región de Murcia. Los resultados muestran que las motivaciones afectan a la probabilidad de elección de un alojamiento rural. Además, también se observa que el efecto de las motivaciones es diferente en función de los atributos

del alojamiento.

Por último, en la sección de conclusiones se hace una recopilación de todos los estudios llevados a cabo en este trabajo y de sus resultados.

Capítulo 1

Modelización de la elección de alojamiento rural

1.1. Introducción

El alojamiento constituye una parte fundamental del viaje turístico. Según Sharpley (2000) el alojamiento supone alrededor de un tercio del gasto total del viaje. En España, los datos publicados por el Instituto Nacional de Estadística (INE) procedentes de la Encuesta de Gasto Turístico (EGATUR) indican que el gasto de los turistas extranjeros en alojamiento en 2019 supuso aproximadamente el 20 % del gasto total del viaje. Aunque es importante señalar que el valor del alojamiento para los turistas no es sólo el gasto que realizan, sino que también constituye un elemento esencial para determinar la satisfacción y relevancia de su experiencia turística (Mcintosh y Siggs, 2005).

En el contexto rural, los alojamientos, que suelen ser negocios pequeños, tales como hoteles rurales, B&B, casas de alquiler y casas de huéspedes (Ye et al., 2019), administrados generalmente por las familias locales como medida de supervivencia económica y para la revitalización de estas áreas, también tienen un valor más allá del económico. Según Ye et al. (2019), los huéspedes de estos pequeños alojamientos rurales persiguen tres aspectos con su experiencia turística, que se pueden clasificar como funcionales, emocionales y de autenticidad. El principal atractivo de las casas rurales para los turistas subyace en su entorno natural y sus características intrínsecas rurales, a la vez que implica un contacto con la gente del lugar y con su estilo de vida. Estas características están incrustadas en el significado de alojamiento rural por lo que se entienden como factores de autenticidad. Los aspectos funcionales están asociados con las características físicas y de calidad de los

alojamientos, como podrían ser el tipo de alojamiento, el tamaño, su forma de alquiler, la calidad de su mobiliario o la variedad de los servicios y actividades que oferta. Por último, los individuos, cuando hacen turismo, también buscan experimentar una serie de sentimientos, emociones y diversión.

De acuerdo con lo anterior, la decisión de un individuo ante la elección de alojamiento rural está relacionada con las propiedades, características y atributos de los alojamientos que satisfacen las necesidades del turista. Pero, además, esta elección debe considerarse como un proceso discreto, los alojamientos no son divisibles, el turista puede optar por un alojamiento completo o por una habitación, pero siempre en su totalidad. La elección también es una acción precisa, en el sentido de que el turista elige el alojamiento o no, pero no hay otras posibilidades. Dada la existencia de un conjunto amplio de alojamientos disponibles, se trata de una elección multinomial. Y por último, es una elección no ordenada, no existe un orden que implique elegir un alojamiento antes que otro (Morley, 1994). En consecuencia, la decisión de elección de alojamiento rural se puede entender como una decisión de elección discreta en cuanto que se ajusta a los criterios de estos modelos (Albaladejo y Díaz, 2005).

Los modelos de elección discreta constituyen, por tanto, una herramienta econométrica muy útil y utilizada para explicar o predecir una elección entre un conjunto de dos o más alternativas discretas (es decir, distintas, separables y mutuamente excluyentes). Además, para estimar estos modelos se pueden utilizar datos de preferencias reveladas y de preferencias declaradas. Los primeros, utilizados tradicionalmente, son observaciones de elecciones hechas por los individuos en entornos reales, mientras que los datos de preferencias declaradas muestran las elecciones realizadas por individuos reales en entornos de elección hipotéticos, que habrán sido cuidadosamente diseñados por los investigadores. Con objeto de poder utilizar ambos tipos de datos para las estimaciones de nuestros modelos de elección de alojamiento rural, en este trabajo se utilizan los datos de una encuesta realizada en la Comarca del Noroeste de la Región de Murcia durante varios fines de semana en otoño de 2003.

En este capítulo se definen y desarrollan los conceptos, modelos y datos que, de forma general, se usarán a lo largo de toda esta memoria. Así, tras hacer un pequeño análisis de los alojamientos de turismo rural en España, se introducen los modelos de elección discreta, exponiendo el más conocido de ellos, el modelo logit multinomial. Seguidamente, se definen que son los datos de preferencias reveladas y los de preferencias declaradas y se explican sus

diferencias. Posteriormente, se detalla cómo se obtuvieron los datos que serán utilizados en los diferentes estudios que se presentan en este trabajo. Por último, este primer capítulo finaliza con una breve descripción de los modelos que proponemos para estudiar los factores (observables y no observables) que influyen en la elección de alojamiento rural y las características socio-económicas y de condiciones de viaje del turista que son determinantes en su elección.

1.2. Alojamientos rurales

El concepto de turismo rural puede definirse de forma general como “un turismo que tiene lugar en el campo” (Lane, 1994). Sin embargo, este concepto ha sido entendido de forma muy distinta dependiendo del país o incluso de la región en la que se ha desarrollado¹. Ello se debe fundamentalmente a que el término “rural” no siempre hace referencia a los mismos atributos o características, especialmente en el contexto de los países desarrollados. Pero, también, a que el turismo rural es una actividad compleja y polifacética, marcada por una diversidad cada vez mayor (Lane, 2009). Bajo este concepto se engloban actividades como el ecoturismo, el agroturismo, el turismo de naturaleza, el de granjas, el de aventuras, el de deportes, el de gastronomía o el turismo cultural.

En Europa, el turismo rural cuenta con una larga tradición (Cánoves et al., 2004). Desde el comienzo del siglo XX, en el Tirol y en la Inglaterra rural, ya era posible el alojamiento en granjas, pero es en la década de los sesenta cuando el turismo rural comienza a despegar, con un crecimiento espectacular, especialmente en Francia, Alemania, Italia y los Países Bajos. En España, aparece en 1967 bajo el efímero programa organizado por el Estado denominado “Vacaciones en Casas de Labranza” y cuyo objetivo era dinamizar el medio rural incorporando actividades turísticas a la explotación ganadera tradicional. Este programa consistía en la concesión de créditos sin intereses para la remodelación, reforma o adaptación de las viviendas rurales para su incorporación a la oferta de alojamiento de turismo rural (Morales-Urrutia et al., 2018).

Sin embargo, su desarrollo en España fue lento en comparación con el resto de países europeos. A mediados de los ochenta, el desarrollo de diversos proyectos apoyados institucionalmente en diferentes zonas y regiones fueron

¹Véase Rosalinda et al. (2021) para una revisión.

una importante fuerza impulsora. De estos proyectos, quizás el más conocido sea el de Taramundi en Asturias. A partir de la integración en la Unión Europea (1986), España comenzó a contar con aportaciones de fondos comunitarios, destinadas a promover el desarrollo de las zonas rurales, como fueron los planes LEADER I, LEADER II y LEADER + (“Relaciones entre Actividades de Desarrollo de la Economía Rural”), el Programa PRODER (Programa Operativo para el Desarrollo y la Diversificación y Diversificación Económica del Medio Rural), el Programa AGADER (Agencia Gallega de Desarrollo Rural) para el caso gallego, FEADER (Fondo Europeo Agrícola de Desarrollo Rural), entre otros.

A partir de los años 90, el turismo rural en España experimenta un crecimiento considerable (Cánoves et al., 2005). Según Barke (2004), dos fueron las fuerzas que contribuyeron al desarrollo de este turismo, la disminución de las actividades rurales tradicionales y la necesidad de diversificar el turismo de sol y playa, que en España es un producto turístico de masas. Además, en esta década de los noventa, surgen nuevos patrones de comportamiento de los turistas en relación con el tiempo de ocio, la segmentación de las vacaciones y el desarrollo de los “finés de semana largos”. En este ambiente, el turismo rural permitía a los turistas disfrutar en un entorno cercano del contacto con la naturaleza, de la calidad del medio ambiente, de paz y tranquilidad y no masificación, así como de actividades relacionadas con la agricultura y el entorno rural.

Como ocurrió en Europa, el turismo rural en España comenzó basándose casi exclusivamente en la oferta de los alojamientos. Esta oferta se realizaba bajo una amplia gama de nombres comerciales debido a los amplios poderes legislativos de las Comunidades Autónomas. Mientras en Andalucía se ofertaban “Casas Rurales”, en Cantabria existían “Posadas”, “Casonas” y “Casas de Labranza” y en Asturias “Casonas” y “Casas de Aldea”. De este modo, se han usado denominaciones diferentes para alojamientos similares, mientras que algunas veces la misma denominación hacía referencia a productos diferentes (Valdés, 1996 y Valdés y Del Valle, 2003). No obstante, había ciertas características de los alojamientos de turismo rural que eran comunes para las legislaciones de todas las Comunidades Autónomas. Una de ellas era la localización del establecimiento de alojamiento que debía estar dentro de un entorno rural. Además, tenía que tener una oferta limitada en cuanto al número de habitaciones o camas para huéspedes, debía tener una equipación con servicios básicos y su arquitectura tenía que ser coherente con el estilo de los alrededores (Valdés y Del Valle, 2003). En cuanto al tipo de alquiler, los establecimientos se podían alquilar completos o por habitaciones. En este

último caso, la casa se podía compartir con otros turistas o, incluso, con los propietarios.

Con el objetivo de captar nuevos turistas y persuadirlos de volver, muchos establecimientos rurales, además del alojamiento, comenzaron a promover actividades relacionadas con la naturaleza y la vida rural, como la equitación, la pesca, la caza, las excursiones a pie, el rafting, los cursos educativos, tratamiento de terapia, la recolección de fruta, etc. Los establecimientos también empezaron a vender productos locales o del campo y a ofrecer productos más sofisticados como alimentos gourmet. Así, la oferta de establecimientos de alojamiento variaba, en gran medida, en términos de sus atributos y del rango de instalaciones y actividades, aunque cada país o región europea ponía el énfasis en una o más especialidades (Cánoves et al., 2004).

El principal problema que tenían estos establecimientos de alojamiento en un entorno rural era el de su comercialización. Estos establecimientos, tal y como establece la ley, eran pequeños, independientes y dirigidos y administrados por las familias locales. Estas familias, a menudo, disponían de pocos recursos necesarios y solían tener dificultades para adaptarse a los mecanismos del mercado cada día más competitivo, dominado por complejas técnicas de promoción y comunicación (Gannon, 1994). Así, en un principio, surgieron diferentes grupos de desarrollo y comercialización de turismo rural a nivel regional y en algunas ocasiones a nivel nacional. Estos grupos, normalmente sin fines de lucro, proporcionaban a los propietarios de los alojamientos herramientas de marketing, redes de contactos, oportunidades de formación y sistemas de control de calidad (Gössling y Lane, 2015).

A mediados y finales de la década del siglo XX, con el uso de páginas web por parte de estos grupos de desarrollo y comercialización, pero también, de modo individual por parte de algunos establecimientos, comenzaron una serie de cambios fundamentales en la comercialización de estos alojamientos de turismo rural. El desarrollo de plataformas globales de reservas que permitían la inscripción de los alojamientos independientemente de su tamaño, presupuesto y localización también fueron muy determinantes para este cambio. Una importante plataforma para este tipo de alojamientos fue Booking.com. La facilidad de uso de estas plataformas por parte de propietarios y consumidores de turismo rural, ha revolucionado totalmente la comercialización de estos alojamientos rurales en el siglo XXI. Básicamente, las plataformas de reservas de Internet son el sistema de comercialización que usan todos los establecimientos de turismo rural.

Actualmente se puede decir que el producto turismo rural está totalmente

consolidado en España y sigue centrándose en la oferta de los alojamientos. Una oferta, que ha sufrido un importante crecimiento durante el siglo XXI y que ha incluido otras tipologías de alojamiento. Así, además de las casas rurales, con sus diferentes denominaciones, se pueden encontrar en el medio rural pequeños hoteles, apartamentos y algunos albergues. Además, cada día son más las actividades complementarias de restauración, deportes, cursos y excursiones, entre otros, que se añaden. A pesar de ello, la dispersión de la oferta continúa y ciertos valores ecológicos, medioambientales, antropológicos y étnicos se mantienen todavía asociados a este tipo de turismo (Morales-Urrutia et al., 2018).

1.3. Modelos de elección discreta

Los modelos de elección discreta constituyen una de las herramientas econométricas más útiles y extendidas para modelizar la decisión de un individuo ante la elección de una alternativa discreta de entre un conjunto de alternativas exclusivo y mutuamente exhaustivo. La base teórica para analizar el comportamiento del consumidor en el contexto de los bienes discretos es la microeconomía de las elecciones discretas desarrollada por McFadden (1981). Siguiendo los postulados de Lancaster (1966), la utilidad a maximizar depende de la cantidad consumida de bienes de naturaleza continua (representados por un vector Y), así como de las características o atributos de las alternativas discretas (representadas por un vector Q). Así, el problema del consumidor puede formularse como sigue (Espino et al., 2008):

$$\begin{aligned} & \max_{j, Y} U(Y, Q_j) \\ & s.a. \sum_i P_i Y_i + c_j \leq I, \\ & Y_i \geq 0, \quad j \in M \end{aligned} \tag{1.1}$$

donde P_i es el precio de mercado del bien i , c_j es el coste de la alternativa j , I es la renta del individuo y M es el conjunto de alternativas disponibles. A partir de las condiciones de primer orden del problema (1.1) para cada j , se obtienen las funciones de demanda condicional a la elección de la alternativa j ,

$$Y_j(P, I - c_j, Q_j). \tag{1.2}$$

Sustituyendo estas funciones en la expresión de la utilidad se obtiene la utilidad indirecta condicional de la alternativa j ,

$$V_j = V_j(P, I - c_j, Q_j) \quad (1.3)$$

Esta función representa la máxima utilidad que puede obtener el individuo una vez elegida la alternativa j . Por tanto, al maximizar en j , el individuo elegirá aquella alternativa que le proporcione máxima utilidad. Así, la utilidad indirecta global V^* se define como

$$V^* = \max_j V_j(P, I - c_j, Q_j) \quad (1.4)$$

La aplicación de la identidad de Roy permite obtener la demanda de las alternativas discretas,

$$\frac{-\partial V_j^*}{\partial c_j} = \delta_i = \begin{cases} 1 & V_i \geq V_j \quad \forall j \neq i \\ 0 & \text{en otro caso} \end{cases} \quad (1.5)$$

que en este caso es una variable discreta. A partir de las derivadas parciales de V_j podemos obtener la utilidad marginal de la renta

$$\lambda_j = \frac{\partial V_j}{\partial I} = \frac{-\partial V_j}{\partial c_j} \quad (1.6)$$

y el valor subjetivo (o la medida de la disposición a pagar) del atributo q_k como

$$\frac{\partial V_j / \partial q_{kj}}{\partial V_j / \partial c_j} \quad (1.7)$$

Los modelos de elección discreta se basan en la teoría de Maximización de la Utilidad Aleatoria de McFadden (1974), que introduce el comportamiento de elección individual como un concepto intrínsecamente probabilístico. De acuerdo con esta teoría, cada individuo tiene una función de utilidad asociada a cada una de las alternativas, que no es conocida con certeza por los analistas. El valor real de la utilidad que el individuo i asocia a la alternativa j , U_{ij} , conocida como utilidad aleatoria, se puede expresar como una suma de dos partes. Una de ellas observable por el investigador, V_{ij} , que representa los gustos o preferencias del individuo por esa alternativa y que se considera que es función tanto de los atributos de la alternativa como de las características del individuo. La otra parte es un término aleatorio o error, ε_{ij} , que indica

los factores desconocidos sobre los individuos o los errores de medida o de información imperfecta de esta utilidad. Esto es,

$$U_{ij} = V_{ij} + \varepsilon_{ij} \quad (1.8)$$

Es evidente que la ecuación (1.8) se verifica exactamente, ya que ε_{ij} no es más que la diferencia entre la utilidad real, U_{ij} , y la parte de utilidad que el investigador conoce, V_{ij} . Pero, puesto que el investigador no conoce íntegramente U_{ij} , no va a poder predecir con total precisión la elección del decisor. Sin embargo, dado su conocimiento de una parte de esta utilidad, puede llegar a realizar ciertas suposiciones sobre la elección del consumidor.

La teoría de la Maximización de la Utilidad Aleatoria supone que el comportamiento de los individuos es racional en el sentido de que cada individuo elige aquella alternativa que maximiza su utilidad. Además, considera que existe un conjunto, C , que incluye todas las alternativas posibles para el grupo de individuos que se analiza, pero que no todas las alternativas tienen por que ser factibles para todos los individuos. Así, cada individuo i tiene su propio conjunto de elección, que no es más que un subconjunto de C , denotado por C_i . Bajo estas hipótesis, el individuo i elige la alternativa j si y solo si se verifica que

$$U_{ij} \geq U_{il} \quad \forall j, l \in C_i; l \neq j \quad (1.9)$$

que equivale a que

$$V_{ij} + \varepsilon_{ij} \geq V_{il} + \varepsilon_{il} \quad \forall j, l \in C_i; l \neq j \quad (1.10)$$

Dado que los ε_{ij} son variables aleatorias $\forall j \in C_i$, solo es posible hablar en términos probabilísticos. De esta forma, la probabilidad de que un individuo i elija la alternativa j del conjunto C_i , que denotamos por P_{ij} , viene dada por la expresión

$$P_{ij} = \Pr [V_{ij} + \varepsilon_{ij} \geq V_{il} + \varepsilon_{il} : \forall j, l \in C_i; l \neq j] \quad (1.11)$$

Esto es

$$P_{ij} = \Pr \left[V_{ij} + \varepsilon_{ij} \geq \max_{\substack{l \in C_i \\ l \neq j}} (V_{il} + \varepsilon_{il}) \right] \quad (1.12)$$

que está bien definida siempre que exista una única alternativa que maximice la utilidad del individuo i .

Según la expresión (1.11 o 1.12), el cálculo de la probabilidad de elección de la alternativa j por parte del individuo i , depende de las suposiciones específicas que se hagan sobre las distribuciones conjuntas de las utilidades no observadas, y sobre la forma funcional de las utilidades observadas.

El más popular de los modelos de elección discreta, el logit multinomial, surge de suponer que las componentes aleatorias, ε_{ij} , son independientes e idénticamente distribuidas entre sí con una distribución del valor extremo del tipo I o Gumbel con parámetro de localización η igual a cero y parámetro de escala μ , $\mu > 0$. Bajo esa suposición, se obtiene que la probabilidad de que un individuo i elija la alternativa j viene dada por (McFadden, 1974),

$$P_{ij} = \frac{e^{\mu V_{ij}}}{\sum_{l \in C_i} e^{\mu V_{il}}} \quad \forall j \in C_i \quad (1.13)$$

Por tanto, las probabilidades de elección del modelo logit multinomial son función del parámetro de escala, μ , y de la componente determinística de la función de utilidad, V_{ij} .

El factor de escala μ , al estar inversamente relacionado con la varianza del término de error, como muestra la ecuación (1.14),

$$Var[\varepsilon_{ij}] = \frac{\pi^2}{6\mu^2}, \quad (1.14)$$

nos informa sobre el nivel de incertidumbre asociado con la utilidad esperada (con los atributos inobservados) de cada alternativa. Esto es, si parámetro μ tiende a cero, entonces la incertidumbre sobre las características inobservadas de las alternativas es máxima. En esta situación el modelo de elección no proporciona ninguna información, y las alternativas del conjunto de elección C_i , por tanto, son igualmente probables. Si μ tiende a infinito, entonces la incertidumbre sobre las características inobservadas de las alternativas es inexistente, ya que toda la información acerca de las preferencias individuales está incluida en la parte sistemática de la función de utilidad, V_{ij} . El modelo de elección es, por tanto, determinista (Ben-Akiva y Lerman, 1985).

Lo habitual es que la componente observable de la utilidad o utilidad determinística, V_{ij} , venga dada como una especificación lineal de las características de las alternativas, esto es,

$$V_{ij} = \sum_k \beta_k x_{ijk} \quad (1.15)$$

con x_{ijk} representando el valor de la característica k de la alternativa j observada por el individuo i y β_k el parámetro a estimar para la característica k . Substituyendo la expresión (1.15) en (1.13) se obtiene que las probabilidades de elección del modelo logit multinomial son igual a

$$P_{ij} = \frac{e^{\mu \sum_k \beta_k x_{ijk}}}{\sum_{l \in C_i} e^{\mu \sum_k \beta_k x_{ilk}}} \quad \forall j \in C_i \quad (1.16)$$

La ecuación (1.16) pone de manifiesto la imposibilidad de estimar el parámetro μ de forma separada del vector de parámetros β con un sólo conjunto de datos. De ahí que, si se quiere estimar μ y β_k , sea necesario combinar dos o más fuentes de datos, tal y como señala Morikawa (1989), ya que esto nos permite identificar el cociente de los parámetros de escala. En concreto, podemos estimar parámetros de escala relativos combinando fuentes de datos de preferencias reveladas (PR) y de preferencias declaradas (PD), aplicando un método de estimación secuencial (Swait y Louviere, 1993) o simultáneo (Morikawa, 1989; Ben-Akiva y Morikawa, 1991; Hensher y Bradley, 1993 y Bhat 1995, entre otros).

No obstante, en la mayoría de aplicaciones donde únicamente se dispone de un conjunto de datos, ante la inseparabilidad del gusto (β) y la escala (μ), se normaliza la escala a algún valor concreto, generalmente uno, de forma que se mezcla su efecto con el vector de parámetros β en la estimación del modelo,

$$P_{ij} = \frac{e^{\sum_k \beta_k x_{ijk}}}{\sum_{l \in C_i} e^{\sum_k \beta_k x_{ilk}}} \quad \forall j \in C_i, \quad (1.17)$$

de manera que el modelo logit multinomial puede sobre-estimar (sub-estimar) los parámetros del modelo, β , cuando la incertidumbre sobre las características inobservadas de la alternativas es baja (alta) (Swait y Louviere, 1993).

Debido a su fácil tratabilidad y al sencillo proceso de estimación de sus parámetros, el modelo logit se ha utilizado en muchas aplicaciones empíricas en campos muy diversos y, en particular, en el turismo (Eymann et al., 1992; Costa y Manente, 1995; Albaladejo y Díaz, 2003; Eugenio-Martin, 2003; Apostolakis y Jaffry, 2005a, 2005b, entre otras). Pero también ha sido muy

crítico por verificar la conocida propiedad de independencia de las alternativas irrelevantes (IIA) (Train, 2003), según la cual, la decisión de un individuo de elegir una alternativa frente a otra es independiente de la presencia o ausencia de una tercera no escogida y de los atributos de cualquier otra alternativa del conjunto de elección. Esto se puede mostrar fácilmente considerando el cociente de las probabilidades de elección de dos alternativas cualquiera, j y l ,

$$\frac{P_{ij}}{P_{il}} = \frac{\frac{e^{V_{ij}}}{\sum_{k \in C_i} e^{V_{ik}}}}{\frac{e^{V_{il}}}{\sum_{k \in C_i} e^{V_{ik}}}} = e^{(V_{ij} - V_{il})} \quad (1.18)$$

La propiedad IIA no es la única limitación del logit multinomial. Este modelo tampoco permite captar las diferencias de gustos de los individuos (Espino et al., 2008; Train, 2003), dado que considera un conjunto de coeficientes fijos para todos los individuos de la población. El logit multinomial es también poco apropiado ante la presencia de heterocedasticidad entre individuos o atributos, dado que considera las varianzas de las componentes aleatorias de las utilidades iguales y, en consecuencia, se considera que el nivel de incertidumbre asociado con los atributos inobservados de cada alternativa se mantiene constante para todas las alternativas disponibles.

Con objeto de superar las limitaciones del logit multinomial, se han desarrollado otros modelos de la “familia Logit” (véase Tabla 1.1) que relajan alguna o algunas de estas restricciones tratando de mantener tratabilidad del modelo, como el modelo logit anidado (NL), el modelo de Valor Extremo Generalizado (GEV) (Walker y Ben-Akiva, 2002) y el modelo de Logit Mixto (ML) (Train, 2003). Por ejemplo, Chen et al. (2017) utilizaron un modelo logit anidado para investigar si el proceso de elección de alojamiento se explicaba mejor por una estructura jerárquica, en la que en primer lugar los turistas elegían entre hoteles y no hoteles, y, después, escogían un tipo de alojamiento específico. El modelo logit mixto también ha sido empleado para modelizar la elección de hospedaje porque permite introducir heterogeneidad en las preferencias sobre los atributos del alojamiento entre los turistas (Román y Martín, 2016; Masiero et al., 2015; Albaladejo y Díaz, 2009).

Tabla 1.1: Modelos familia Logit

• Homogeneidad en las preferencias de los individuos: El vector de parámetros β se mantiene constante a través de los individuos.	
<i>Modelo Logit Multinomial (MNL)</i>	<ul style="list-style-type: none"> – ε_{ij} tienen varianzas iguales (μ) para todas las alternativas (Homocedasticidad) – ε_{ij} son independientes – Patrones de sustitución entre las alternativas simétricos
<i>Modelo Logit Anidado (NL)</i>	<ul style="list-style-type: none"> – Heterocedasticidad parcial: Las varianzas de las componentes inobservadas de la utilidad dentro del nido son iguales, pero distintas entre los nidos – Dependencia parcial: Las alternativas dentro del mismo nido están correladas, pero son independientes entre los nidos – Permite formalizar la elección de alternativas con cierto grado de sustitución entre ellas
<i>Modelo del Valor Extremo Heterocedástico (HEV)</i>	<ul style="list-style-type: none"> – ε_{ij} tienen varianzas distintas entre las alternativas (μ_j) (Heterocedasticidad) El analista puede tener un grado de incertidumbre diferente sobre la utilidad esperada de cada alternativa j – ε_{ij} son independientes – Permite formalizar la elección de alternativas con diferente grado de sustitución entre ellas
<i>Modelo del Valor Extremo Heterocedástico Parametrizado (Cov-Het)</i>	<ul style="list-style-type: none"> – ε_{ij} tienen varianzas distintas entre las alternativas y los individuos (μ_{ij}) (Heterocedasticidad) La varianza de ε_{ij} tiene una forma funcional concreta y depende de covariables asociadas con el individuo i y la alternativa j, \mathbf{W}_{ij}. – ε_{ij} son dependientes – Permiten patrones de sustitución complejos entre las alternativas
• Heterogeneidad discreta: Existen S segmentos en la población cada uno de los cuales tiene su propio vector de parámetros del gusto β_s y su propia escala μ_s	
<i>Modelo logit de clase latente (CL)</i>	<ul style="list-style-type: none"> – Heterocedasticidad parcial: Los términos de error ($\varepsilon_{ij/s}$) dentro de cada segmento tienen la misma varianza (μ_s) (homocedásticos dentro del mismo segmento s), pero son distintas entre los segmentos (heterocedasticidad entre segmentos) – $\varepsilon_{ij/s}$ son independientes – Este modelo no impone la propiedad IIA a las probabilidades incondicionales observadas (aunque sí lo hace dentro de cada clase)
• Heterogeneidad continua: Los parámetros del gusto varían aleatoriamente entre los individuos β_i	
<i>Modelo Logit Multinomial Escalado (SLM)</i>	<ul style="list-style-type: none"> – $\beta_i = \mu_i \beta$ donde β es un vector columna μ_i sigue una distribución log-normal de media uno o, equivalentemente, $\mu_i = \exp\left(-\frac{\tau^2}{2} + \tau w_i\right)$ con $w_i \sim N(0, 1)$ y $\tau \geq 0$ – Heterocedasticidad de escala. La escala varía entre los individuos (μ_i), pero es invariante ante la elección – ε_{ij} son independientes
<i>Modelo Logit de Parámetros Aleatorios (RPL) o Logit Mixto (ML)</i>	<ul style="list-style-type: none"> – β_i es vector aleatorio de coeficientes inobservados para cada individuo i con función de densidad $f(\beta_i/\theta)$ con θ los parámetros de la distribución f. Las distribuciones f más habituales son la normal y la log-normal – ε_{ij} tienen varianzas iguales (μ) para todas las alternativas (Homocedasticidad) – ε_{ij} son independientes entre sí e independientes de β_i – Existen correlación entre las alternativas a través de la parte estocástica de la utilidad

1.4. Datos de preferencias reveladas y declaradas

La información utilizada para estimar los modelos de elección discreta puede proceder de dos fuentes de datos diferentes, según el criterio empleado para extraer las preferencias de los individuos ante una situación de elección. Esto es, los datos pueden derivarse de encuestas de Preferencias Reveladas (PR) y Declaradas (PD).

Hasta finales del siglo XX, prácticamente todas las aplicaciones de los modelos de elección discreta a decisiones en el ámbito del turismo y los viajes se realizaron con datos de preferencias reveladas. Los datos de este tipo se obtienen mediante la observación de la situación real de elección a la que el individuo se enfrenta. Básicamente consiste en recopilar, mediante encuestas, información sobre las elecciones de destino, hotel, medio de transporte, u otras, que el individuo ha hecho en un momento determinado de tiempo. Con estos datos se procede al análisis econométrico de los modelos de elección discreta. Sin embargo, no siempre los datos de elecciones reales son del todo idóneos a la hora de hacer las estimaciones. Con frecuencia, muestran poca variación o una variabilidad limitada para los valores de algunas variables determinantes en el proceso de estimación, reduciendo, por tanto, la precisión de los parámetros estimados. También, en muchas ocasiones, las variables explicativas están altamente correlacionadas, lo que hace difícil medir el impacto de cada una de ellas sobre la variable respuesta, reduciendo considerablemente el número de parámetros que se pueden estimar. Además, no existe posibilidad de incorporar nuevas alternativas que difieran sustancialmente de las ya existentes. Por otro lado, estos datos son propensos a errores de medida potencialmente importantes (Ortúzar y Willumsen, 2001 y Espino, 2003).

Con la intención de disponer de datos más “idóneos” para el proceso de estimación de los modelos de elección discreta, surgen los datos de preferencias declaradas. Estos datos recogen las declaraciones de los individuos en cuanto a su comportamiento frente a elecciones hipotéticas que han sido cuidadosamente diseñadas por los investigadores (Ortúzar, 2000). El origen de estos datos se encuentra en los estudios de investigación de mercado, aunque se han utilizado en muchos campos como en estudios de gestión de recursos y escenarios de valoración mediamambiental (Adamowicz et al., 1997); en la modelización de la elección de lugares de ocio o centros de compras (Oppewal et al., 1997), o en la modelización del transporte (Ortúzar y Willumsen, 2001 y Espino et al., 2006), por mencionar unos cuantos. En los estudios de

turismo, se comienzan a utilizar a principios del siglo XXI. Por ejemplo, con este tipo de datos se ha analizado la elección de destino (Huybers y Bennet, 2000; Huybers, 2003; Papatheodorou, 2001 y Crouch y Louviere, 2004) y la demanda de atracciones del patrimonio (Apostolakis y Jaffry, 2005a, 2005b).

Para la obtención de los datos de preferencias declaradas, el investigador debe diseñar un experimento en el que debe definir claramente, en primer lugar, la situación o contexto en el que el individuo declara sus intenciones, es decir, el “contexto de elección”. En segundo lugar, las alternativas de elección incluidas en él, definidas a través de la combinación de una serie de características o atributos, de manera que un cambio en el nivel de un atributo genera una nueva alternativa. Por último, la forma en la que los individuos declaran sus preferencias, que puede ser: por ordenación (*ranking*), por valoración (*rating*) o por elección (*choice*). En los datos de preferencias declaradas *ranking*, el encuestado tiene que ordenar las alternativas incluidas en el contexto de elección, sin valoración métrica. En los datos de PD *rating* se pide al encuestado que valore el conjunto de alternativas. Este método es más completo, ya que el encuestado no sólo da una ordenación, sino también un valor métrico, por lo que el investigador tiene una aproximación precisa de las preferencias. Finalmente, en los datos de PD *choice*, el encuestado tiene que seleccionar la mejor situación según él. Este tipo de tarea es la que los individuos realizan en la realidad, comparando un conjunto de situaciones y seleccionando una. Los datos de preferencias declaradas por elección son los más realistas y no requieren que se haga ninguna suposición adicional sobre medidas de orden y de cardinalidad (Louviere et al., 2000).

Así, la clave de un experimento de preferencias declaradas, en general, y de elección, como es el planteado en esta memoria, es la selección de los atributos que definen las alternativas. Esto permite al investigador modelizar la demanda de alternativas que no existen en una situación de elección real, e incluir nuevos atributos en alternativas que, aunque existen, no se dispone de ellos en una situación real. Sin embargo, debido a la variabilidad de los atributos, el número de alternativas de elección a las que tendría que enfrentarse un individuo podría ser bastante amplio. Así, para llevar a cabo el experimento sin un agotamiento de los individuos, el conjunto de alternativas de elección debe reducirse, recurriendo a la teoría estadística del diseño de experimentos. Esta teoría, además de obtener subconjuntos eficientes de las alternativas en cuanto a su número, también permite reducir la colinealidad que existe entre los atributos. De esta forma, los datos de preferencias declaradas no sólo son un complemento útil de los datos de preferencias reveladas, sino que también mejoran los niveles de validez predictiva en una amplia

variedad de contextos de elección real (Swait et al., 1994).

A pesar de estas ventajas, los datos de preferencias declaradas no siempre se consideran válidos para la estimación de un modelo de elección discreta. Dada la naturaleza hipotética de las alternativas de elección, en algunas situaciones no hay una equivalencia real con las alternativas reales existentes en el mercado. Tampoco hay una garantía de que la elección que un individuo realiza en una situación de elección hipotética, coincida con la que realmente haría si se diera dicha situación de elección en el contexto real. Además, como señalan Bradley y Kroes (1990) y Espino et al. (2004), en la propia respuesta pueden producirse sesgos de varios tipos, como de afirmación, de racionalización, de política o de no restricción² y también es posible la aparición del efecto fatiga durante la cumplimentación de la encuesta de preferencias declaradas (Bates, 1988 y Carson et al., 1994). Estos sesgos y el efecto fatiga pueden dar lugar a un error de medida en la variable dependiente (Ortúzar, 2000). Es importante señalar que si el modelo estimado tiene como objetivo la predicción, un error de este tipo puede tener importantes consecuencias.

Por consiguiente, el uso de un tipo de datos u otro tienen sus ventajas e inconvenientes. Así, es frecuente el uso simultáneo de ambos tipos de datos, dado que, como afirman Morikawa et al. (2002), su uso conjunto puede producir modelos de comportamiento del consumidor más fiables y útiles. Además, la combinación de dos o más fuentes de datos, como es el caso de datos de preferencias reveladas y de preferencias declaradas, también permite resolver el problema de la inseparabilidad del gusto (β) y de la escala (μ) en el modelo logit multinomial (Morikawa, 1989).

En este trabajo, para profundizar en el estudio de la elección de alojamiento rural, utilizamos datos de preferencias reveladas y declaradas. Estos datos proceden de una encuesta realizada en la Comarca Noroeste de la Región de Murcia, diseñada para conocer las preferencias de alojamiento rural de los turistas.

²Según Espino et al. (2004, p. 762-763), el *sesgo de afirmación* se produce cuando “el entrevistado contesta, consciente o inconscientemente, lo que cree que el entrevistador desea escuchar”; el *sesgo de racionalización* cuando el individuo “racionaliza sus respuestas con el objetivo de justificar su comportamiento en el momento de la entrevista”; el *sesgo de política* se produce cuando el encuestado contesta “en función de la expectativa que posee sobre las decisiones de política que se podrían tomar en base a los resultados de la encuesta y no en función de sus preferencias sobre el ejercicio que se le plantea”; y *sesgo de no restricción* se da cuando el entrevistado responde sin tener en cuenta todas las restricciones que afectan a su comportamiento, de manera que sus respuestas no son factibles en la práctica.

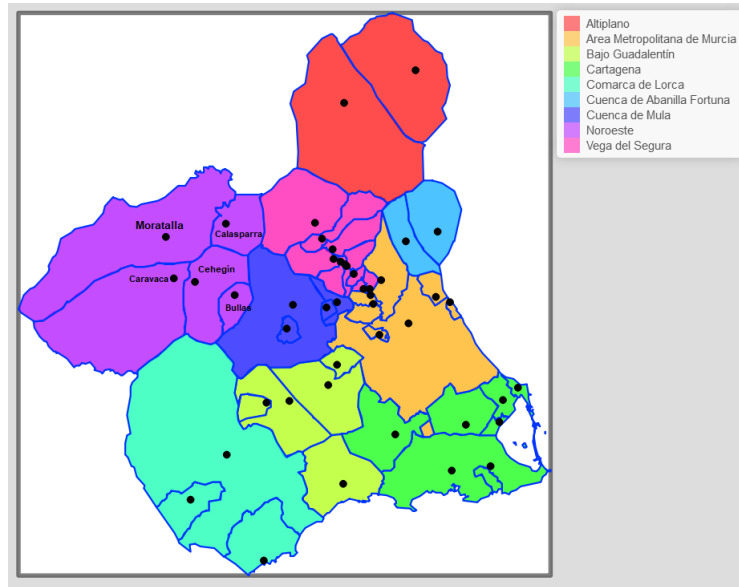
1.5. Encuesta en la Comarca Noroeste de la Región de Murcia

Los datos que se usan en las diferentes investigaciones que se proponen en este trabajo provienen de la encuesta realizada en la Comarca del Noroeste de la Región de Murcia durante varios fines de semana en el otoño de 2003. La encuesta la realizó el Grupo de Investigación Modelización Matemática para la Economía bajo el proyecto “Crecimiento Económico y Modelos de Preferencias en el Estudio del Turismo” que fue financiado por la Fundación Seneca para los años 2002 y 2003.

Aunque Murcia, como la mayoría de las regiones de la costa del Mediterráneo, es un destino generalmente asociado con el turismo de “sol y playa” el número de alojamientos turísticos en zonas rurales se incrementó de forma notable a partir del Decreto 79 (Ley 79), de 10 de septiembre de 1992, que regulaba un nuevo tipo de alojamiento en la Región, conocido como “Alojamiento Turístico Especial de Interior”. Pronto los alojamientos rurales se distribuyeron por toda la Región de Murcia, aunque es la Comarca del Noroeste la que concentra la mayor oferta. Además, esta zona fue la primera en disponer de este tipo de alojamientos. La zona Noroeste de Murcia es una zona montañosa del interior, lejos de las ciudades y pueblos de tránsito tradicionales de la región (zona de color morado en la Figura 1.1). Se caracteriza por extensas mesetas, viñedos, campos de cereales y de arroz, y picos montañosos como el Macizo de Revolcadores, el más alto de la región. Hay cinco municipios en esta área: Bullas, Calasparra, Caravaca, Cehegín y Moratalla; que ofrecen tradiciones y cultura local como complemento para al ocio, la naturaleza, el deporte y la buena comida de la zona (Albaladejo y Díaz, 2007).

A diferencia del turismo de playa, que se concentra en el período de verano, el turismo rural en Murcia tiene una demanda estacional basada en periodos de vacaciones cortos y fines de semana. Así, tras consultar con los responsables de la central de reservas de NORATUR (Asociación de proveedores de alojamiento rural y actividades de turismo rural en Murcia), que gestionaban en el período que se llevó a cabo la encuesta una gran parte de los alojamientos alquilados en la comarca del Noroeste, se decidió pasarla en otoño.

Por tanto, los participantes de la encuesta eran turistas que pasaban al menos una noche en un alojamiento rural de la Comarca. Con el fin de garantizar una muestra representativa, se entrevistaron turistas en todos los municipios del área del Noroeste. El número en cada municipio se determinó en

Figura 1.1: Mapa de la Región de Murcia[†]

[†]Elaboración propia utilizando la librería *leaflet* de R

proporción al número de camas ofertadas en este tipo de alojamientos. Las entrevistas las realizaron los investigadores del proyecto y estudiantes entrenados y supervisados por los primeros. Durante varios fines de semana de otoño, los encuestadores situados en las zonas de máxima afluencia turística de los diferentes municipios que conforman la Comarca, contactaban con los turistas y aquellos que accedían a participar en la encuesta, se les pedía que completaran un cuestionario en papel bajo su supervisión (en el Anexo I se puede consultar el cuestionario utilizado). Este cuestionario constaba de tres partes: en la primera se solicitaba datos individuales de los turistas, la segunda trataba de identificar los motivos de los turistas para hacer turismo rural, y la tercera eran preguntas destinadas a llevar a cabo el experimento de preferencias declaradas, objetivo principal de la encuesta.

Dado que la Región de Murcia disponía en el momento de realización de la encuesta de 1980 camas, de las cuales el 60 % se concentraban en la Comarca del Noroeste³ de la región (de acuerdo a los datos del Instituto Nacional de Estadística, septiembre 2003), se llevaron a cabo 307 entrevistas, con una confianza del 95 % y 4,8 % de error.

³En 2003 la Comarca del Noroeste aglutinaba el 60 % de la oferta total de camas disponibles en la Región de Murcia, situación que se ha mantenido estable a lo largo de los años hasta llegar a nuestros días.

En cuanto a los datos individuales de los turistas, el cuestionario disponía de preguntas relativas a sus características demográficas y socio-económicas, como edad, género, estado civil, nivel de educación e ingresos. También existían preguntas en relación a aspectos relacionados con el viaje, como el gasto per cápita del viaje, el medio de transporte utilizado, el tamaño del grupo con el que viajaban y las actividades que tenían contratadas o que tenían pensado contratar o realizar. La mayoría de estas variables se seleccionaron de estudios anteriores (Yagüe, 2002 ; Seddighi y Theocharous, 2002; Solsona, 1999; Kastenholz et al., 1999; Sharpley y Sharpley, 1997; Oppermann, 1996; Fuentes, 1995). En la variable “Alquiler del alojamiento” se incluyó la modalidad NORATUR ya que en ese momento era la asociación más antigua y conocida de propietarios y agencias de alojamiento y proveedores de actividades de turismo rural en la Región de Murcia. Para terminar, se les preguntaba por la casa rural en la que residían. Los perfiles de los encuestados y sus condiciones de viaje se muestran en la Tabla 1.2.

La segunda parte del cuestionario estaba dedicada a obtener información sobre las motivaciones que cada turista encuestado tenía para hacer turismo rural. Para ello, se diseñó un conjunto de indicadores, que debían ser calificados en una escala de tipo Likert del 1 al 5, donde 1 equivalía a estar “Totalmente en desacuerdo” con el motivo y el 5 “Totalmente de acuerdo”. Los indicadores sobre los que se preguntaron y que se muestran en la Tabla 1.3, fueron seleccionados de estudios anteriores sobre el turismo en general (Frochot y Morrison, 2000; Jang et al., 2002; Ryan y Glendon, 1998) y sobre el turismo rural en particular (Kastenholz et al., 1999; Sharpley y Sharpley, 1997). También se tuvo en cuenta los resultados de una encuesta piloto realizada en la primavera de 2002, siguiendo la recomendación de los gestores del Centro de Reservas de NORATUR, y que se llevó a cabo para ensayar el cuestionario y detectar errores. En la Tabla 1.3 se muestra un resumen de las principales medidas descriptivas de los indicadores y la Figura 1.2 muestra la clasificación de los indicadores de acuerdo con las valoraciones de los encuestados, mostrando mayor intensidad de color (azul más oscuro) cuando la categoría del indicador presenta mayor porcentaje de observaciones.

En la tercera parte del cuestionario se desarrollaba el experimento de preferencias declaradas por elección. Éste tenía como objetivo analizar los principales atributos o características de los alojamientos que los turistas consideraban cuando elegían un alojamiento para pasar sus vacaciones en una zona rural. El experimento fue conceptualizado como un problema en el que los individuos tenían que elegir entre pasar un fin de semana en uno de los dos alojamientos rurales hipotéticos descritos o “quedarse en su casa”.

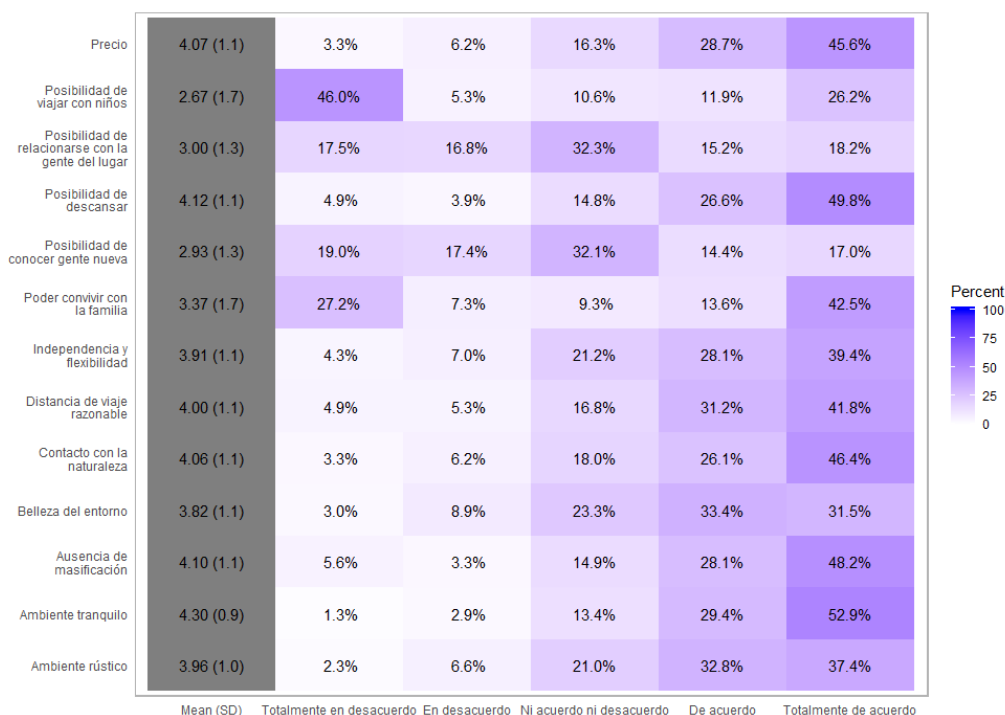
Tabla 1.2: Perfil de los encuestados

Tipo	Denominación	Nivel	Frecuencia	%
Características Socio-económicas	Género	Hombre	149	48.5
		Mujer	158	51.5
	Edad	<= 25	83	27.0
		26 - 40	186	60.6
		> 40	38	12.4
	Lugar de residencia	Alicante	67	21.8
		Murcia	222	72.3
		Resto de España	18	5.9
	Nivel de formación alcanzada	Educación primaria	59	19.2
		Educación secundaria o similar	92	30.0
		Educación superior	156	50.8
	Situación laboral	Ocupado (asalariado / autónomo)	232	75.6
		Parado	13	4.2
		Estudiante	43	14.0
		Jubilado o tareas del hogar	19	6.2
	Profesión	Empresario con trabajadores a su cuenta	2	0.9
		Empresario sin trabajadores a su cuenta	28	12.1
		Trabajador por cuenta ajena	128	55.2
Profesión liberal		29	12.5	
Funcionario		45	19.4	
Estado civil	Casado	130	42.3	
	Soltero	168	54.7	
	Divorciado o separado / viudo	9	2.9	
Ingresos familiares mensuales	Menos de 1202 €	45	14.7	
	Entre 1202 y 2404 €	131	42.7	
	Mas de 2404 €	131	42.7	
Condiciones de viaje	Alquiler del alojamiento	Directamente a los propietarios	198	64.5
		Vía NORATUR	88	28.7
		Otras vías	21	6.8
	Referencia del lugar	A través de conocidos	134	43.6
		Internet y/o guías turísticas	116	37.8
		Otros viajes, anuncios de prensa u otras vías	57	18.6
	Compañeros de viaje	Viaja únicamente con amigos	158	51.5
		Viaja únicamente con la familia	55	17.9
		Viaja con familia y amigos	94	30.6
	Viaja con niños	Sí	124	40.40
		No	183	59.60
	Gasto diario por persona y día (excluyendo alojamiento)	Hasta 30 €	68	22.1
		De 30 a 60 €	106	34.5
		De 60 a 90 €	70	22.8
		Más de 90 €	63	20.5
Número de personas con las que viaja	Menos de 6 personas	101	32.9	
	Entre 6 y 12 personas	109	35.5	
	Más de 13 personas	97	31.6	
Experiencia previa en turismo rural	Sí	248	80.9	
	No	59	19.2	

Tabla 1.3: Descripción de los indicadores

Indicador	n	media	sd	mediana	asimetría	curtosis
Ambiente rústico	305	3,964	1,027	4	-0,800	0,020
Ambiente tranquilo	306	4,297	0,902	5	-1,285	1,351
Ausencia de masificación	303	4,099	1,123	4	-1,284	0,989
Belleza del entorno	305	3,816	1,066	4	-0,653	-0,261
Contacto con la naturaleza	306	4,062	1,089	4	-1,018	0,257
Distancia de viaje razonable	304	3,997	1,115	4	-1,074	0,491
Independencia y flexibilidad	302	3,914	1,126	4	-0,847	-0,058
Poder convivir con la familia	301	3,369	1,695	4	-0,397	-1,563
Posibilidad de conocer gente nueva	305	2,931	1,327	3	0,075	-1,039
Posibilidad de descansar	305	4,125	1,111	4	-1,279	0,949
Posibilidad de relacionarse con la gente del lugar	303	2,997	1,323	3	0,023	-1,037
Posibilidad de viajar con niños	302	2,669	1,720	2	0,283	-1,667
Precio	307	4,072	1,076	4	-1,067	0,424

Figura 1.2: Resultados exploratorios sobre las motivaciones para realizar turismo rural



Los alojamientos fueron descritos en términos de diferentes niveles de una serie de atributos seleccionados (véase Tabla 1.4).

Para evitar descripciones de alojamientos poco probables o irrelevantes y no excluir atributos importantes, la selección de los factores incluidos se hizo con mucho cuidado (Hensher et al., 2005). Por ello, se recogió información sobre características, equipaciones y servicios que ofrecían los alojamientos rurales en la zona noroeste de la Región de Murcia y de otras zonas con tradición en turismo rural en España. Después, se usó un grupo focal, que comprendía expertos en turismo rural de la Región de Murcia, para decidir entre los atributos seleccionados cuales resultaban más relevantes y cuales eran los niveles apropiados para estos atributos. Así, se propusieron 14 atributos (4 atributos con 4 niveles y 10 con 2 atributos) que se muestran en la Tabla 1.4.

Todos los niveles de las primeras diez variables eran viables para las casas rurales en Murcia. El nivel “por habitación” de la variable “tipo de alquiler” y el nivel de “privado” de la variable “baño” sólo estaban disponibles en Murcia en el caso de los hoteles rurales, pero estas opciones funcionaban bien en otras casas rurales fuera de nuestra región. El nivel de “A través de internet” para la variable “Reserva del alojamiento” no estaba disponible para las casas rurales en Murcia en el momento de realización de la encuesta, pero en opinión de los investigadores del proyecto “Crecimiento Económico y Modelos de Preferencias en el Estudio del Turismo” y del grupo focal podía ser decisiva en el desarrollo y promoción de los alojamientos, tal y como posteriormente ha ocurrido. Por último, se añadió un nuevo factor como un atributo - si el establecimiento poseía la “Q” de Calidad Turística - que es la marca que representa la calidad en el sector turístico español, otorgada por el Instituto para la Calidad Turística Española (ICTE), y es un reconocimiento de los esfuerzos realizados por las empresas turísticas para incorporar los sistemas de calidad.

Con estos 14 atributos y sus niveles se generaron simultáneamente las alternativas de elección (casas rurales hipotéticas) y los conjuntos de elección que contenían a estas alternativas. Tal y como se conceptualizó el experimento, elegir pasar un fin de semana en uno de los dos alojamientos rurales propuestos o “quedarse en casa” (Adamowicz et al., 1997), el conjunto de elección al que se enfrentaba cada individuo tenía siempre 3 alternativas. Por tanto, los escenarios de elección generados en el experimento de elección fueron de tamaño constante (3 alternativas) y las alternativas descritas en ellos, genéricas y no etiquetadas (Louviere et al., 2000), por lo que cada

Tabla 1.4: Atributos y niveles del diseño experimental

Atributo	Modalidades	Denominación
Tipo de construcción	4	Original
		Nueva
		Molino
		Madera
Ubicación	4	Pueblo
		Huerta
		Campo-Montaña
		Aldea
Número de habitaciones	4	1
		2
		3
		4
Precio por habitación	4	80 €
		100 €
		120 €
		140 €
Alquiler de caballos	2	Sí
		No
Parque infantil	2	Sí
		No
Servicio de comidas	2	Sí
		No
Piscina	2	Sí
		No
Instalaciones deportivas	2	Sí
		No
Minigranja	2	Sí
		No
Cuartos de baño	2	Compartido
		Sin compartir o privado
Tipo de alquiler	2	Casa completa
		Por habitaciones
Posee la 'Q' de calidad	2	Sí
		No
Reserva del alojamiento	2	A través de internet
		Por teléfono

casa no contenía más información que la proporcionada por sus atributos. La inclusión de la tercera opción, “quedarse en casa”, servía para hacer el experimento de elección más realista. Así, si a un individuo encuestado no le parecían adecuadas ninguna de las dos casas descritas, siempre podían elegir esta tercera opción (Olsen y Swait, 1998). Además, nos permitía preservar la ortogonalidad de la matriz de datos obtenida con el diseño experimental (Adamowicz et al., 1994).

Dado que había $4^4 \times 2^{10}$ posibilidades para la primera casa y $4^4 \times 2^{10}$ posibilidades para la segunda, y una para la elección “quedarse en casa”, el diseño factorial completo (FC) involucraba $(4^4 \times 2^{10}) \times (4^4 \times 2^{10}) \times 1$ escenarios posibles. El diseño FC permite identificar los efectos principales y las interacciones⁴ de todas las variables incluidas en el experimento. En la práctica, no es posible que un sujeto pueda evaluar tantas opciones cuando establece sus preferencias, por lo que es imprescindible reducir su número sin pérdida de bondad de los datos. Con el fin de seleccionar una muestra de pares de perfiles del diseño factorial completo se realizó un diseño factorial fraccional ortogonal (FFO) (Louviere y Woodworth, 1983). El diseño FFO permite disminuir el número de opciones al sacrificar la posibilidad de medir algunas de las interacciones entre los atributos del modelo y, en algunos casos, de todas las interacciones. En concreto, con un diseño de efectos principales⁵ de resolución III se seleccionaron 64 conjuntos de elección que, posteriormente, fueron distribuidos en 16 grupos con cuatro escenarios de elección cada uno usando un factor de bloque de cuatro niveles. Esto se hizo para suavizar el efecto fatiga de los encuestados (Carson et al., 1994) y el efecto de las observaciones repetidas. Así mismo, con esta división de los conjuntos de elección utilizando un factor bloque, nos asegurábamos de que cada bloque de conjuntos de elección fuera aproximadamente equivalente desde el punto de vista estadístico (Adamowicz et al., 1994) y la estimación eficiente de los efectos principales de los atributos (Kelly et al., 2007). Además, con este diseño los atributos eran ortogonales, pero aunque la variable BLOQUE no era completamente ortogonal si fue optimizada usando el criterio D-eficiente (Huber y Zwerina 1996; Kuhfeld, Tobias y Garratt, 1994). La selección del diseño final se realizó utilizando los paquetes *Factex* y *Optex* de la versión 7 del programa SAS[®] (véase Anexo II). Finalmente, cada encuestado fue asignado

⁴Los efectos principales se definen como la respuesta al pasar al siguiente nivel de una variable cuando el resto de atributos permanecen constante (*ceteris paribus*), mientras que las interacciones consideran los efectos debidos a la variación conjunta de más de una variable (Montgomery, 1991; Espino et al., 2004).

⁵Este diseño permite estimar los efectos principales de los 14 atributos incluidos en nuestro modelo más el término constante, pero sacrifica las interacciones entre los atributos, ya que su efecto está confundido con los efectos principales o con otras interacciones.

aleatoriamente a las encuestas de manera que a cada uno de ellos se le presentó únicamente un grupo con 4 escenarios de elección para su evaluación. En la Figura 1.3 se muestra un ejemplo de conjunto de elección (Albaladejo y Díaz, 2007).

Se recogieron 309 cuestionarios, de los cuales 307 incluían las contestaciones a los escenarios de elección propuestos en el experimento de preferencias declaradas, mientras que la parte de los indicadores motivacionales únicamente fue completada por 290 individuos. Así, dependiendo de la investigación realizada y de la información requerida, el número de cuestionarios válidos podrá ser de 307 o 290, ya que los datos requeridos para las tres investigaciones propuestas en esta memoria, y que se desarrollan en los siguientes capítulos, se han obtenido de esta encuesta diseñada ad-hoc.

1.6. Conclusiones

En este capítulo hemos comenzado explicando que se entiende por turismo rural, definición que actualmente sigue siendo poco clara y sobre la que no hay un consenso universal. Esto se debe a que el turismo rural es complejo, abarca actividades multifacéticas y varía según las regiones y los países (Rosalinda et al.,2021). Seguidamente, hemos hecho una breve exposición sobre el origen y evolución del turismo rural a lo largo del tiempo en Europa y en España, prestando especial interés a los establecimientos de alojamiento en el entorno rural. A continuación, hemos introducido los modelos de elección discreta, haciendo hincapié en el más conocido de ellos, el modelo logit multinomial. Estos modelos son la piedra angular de los estudios que se presentarán en los tres capítulos siguientes de esta memoria. Luego, hemos definido los tipos de datos que se utilizarán en la estimación de los modelos propuestos, los datos de preferencias reveladas y los de preferencias declaradas, explicando sus características y diferencias. Y por último, hemos explicado, de forma detallada, el experimento de elección diseñado para obtener información sobre qué atributos de los alojamientos son relevantes en su elección y hemos especificado cómo se llevó a cabo la encuesta realizada en la Comarca del Noroeste de la Región de Murcia y que constituye la fuente de datos primaria para los diferentes estudios que componen este trabajo.

Así, en este capítulo hemos aglutinado toda la información necesaria y común a los tres estudios propuestos con el propósito de conocer las preferencias de los turistas en cuanto a los alojamientos rurales, el valor que

Figura 1.3: Ejemplo de escenario de elección

E1/1

ESCENARIO 1

	ALOJAMIENTO A	ALOJAMIENTO B
Descripción	Casa de nueva construcción con decoración rústica en la huerta rodeada de árboles 4 habitaciones y 4 cuartos de baño.	Casa de nueva construcción con decoración rústica en núcleo urbano con aproximadamente 12.500 habitantes. 1 habitación y 1 aseo con ducha.
Instalaciones y servicios	Cocina equipada, calefacción, pistas polideportivas y pequeño corral con animales.	Cocina equipada, calefacción, piscina, pistas polideportivas, pequeño corral con animales, parque infantil y picadero de caballos.
Precio del fin de semana	Únicamente es posible alquilar el alojamiento completo. Completo: 320 €	30 €
Sello del Instituto para la Calidad Turística (ICTE)	SI	NO
Reserva	Únicamente por teléfono/fax	Por Internet o por teléfono/fax

¿Cuál de estos alojamientos de turismo rural elegiría usted para pasar un fin de semana? (Rodee con un círculo su elección)

ALOJAMIENTO A
 ALOJAMIENTO B
 OTRO TIPO DE ALOJAMIENTO

le otorgan a las características y a los servicios que ofertan los alojamientos, así como la influencia que en estas elecciones tienen las características, sentimientos y motivaciones de los turistas.

Capítulo 2

Demanda de turismo rural por tipo de alojamiento

2.1. Introducción

La demanda de turismo rural ha sido caracterizada como un segmento turístico que se siente especialmente atraído por la naturaleza y la cultura rural (Lane 1994). El turista rural busca la percepción de espacio y libertad que la amplitud del medio rural le comunica, el sentimiento de paz y tranquilidad que le transmite un espacio rodeado de naturaleza o la sensación de autenticidad y tradición que el contacto con los residentes locales y la vida rural le transfiere (Molera y Albaladejo, 2007). Pero los turistas rurales de un determinado destino, al igual que los turistas de cualquier otro tipo de turismo (sol y playa, urbano, cultural, etc.), pueden ser un grupo muy heterogéneo en cuanto a sus características, preferencias de viaje o de alojamiento, en cuanto al tipo de restauración que desean o en cuanto a las actividades que les gustaría realizar.

Si entendemos el turismo como una actividad que debe contar con los servicios necesarios en términos de cantidad y calidad para satisfacer a los visitantes y proporcionarles una buena experiencia turística, y dada la posible existencia de una gran variedad de turistas, una parte esencial para el desarrollo de un destino es su demanda turística. Así, definir las características y preferencias de los turistas del destino puede ser un elemento clave para tomar decisiones relacionadas con la programación, la dotación de personal y los precios de los servicios que se pretende ofertar. También, tener información sobre la demanda del destino, es un valor para planificar campañas promocionales o de comercialización de los servicios e, incluso, estos estu-

dios puede ser una herramienta a considerar por los gobiernos a la hora de formular políticas de desarrollo turístico, como pueden ser las inversiones en infraestructura turística (Davila et al., 2002; Sharpley y Sharpley, 1997; Wen et al, 2020).

En los destinos rurales, la oferta de servicios se basa fundamentalmente en los alojamientos. Estos alojamientos suelen ser pequeños negocios con pocos recursos económicos que surgen con la intención de promover el desarrollo del destino rural, normalmente social y económicamente deprimido (Yague, 2002; Fleischer y Felsenstein, 2000; Valdés y Del Valle, 2003; Dernoï, 1991). Además, habitualmente estos negocios están regentados por personas con escasos conocimientos para hacer campañas de comercialización o promoción y suelen ser muy heterogéneos entre sí en cuanto al tamaño, al tipo de construcción, a la ubicación, a la oferta de restauración y de actividades. Ante esta situación, podría decirse que la posición de los alojamientos rurales en el extremadamente competitivo mercado turístico en el que se encuentran es bastante difícil. Por tanto, disponer de información relativa a las características y preferencias de los consumidores, en cuanto a los diferentes tipos de alojamiento rural existentes en un destino turístico en particular, puede ser de vital importancia para planificar el desarrollo económico y social de un destino rural.

El objetivo de este capítulo es desarrollar un proceso para determinar los perfiles de turistas que optan por cada una de las opciones de alojamiento existentes en un destino rural. El logit multinomial es un buen modelo para analizar estos perfiles suponiendo que su conjunto de elección está formado por todos los alojamientos distintos de un destino. Sin embargo, los alojamientos rurales que hay en un destino son muy variados en cuanto a los atributos que los definen, por lo que es necesario utilizar un método estadístico multivariante para clasificar en grupos homogéneos los alojamientos con similares características. Cada grupo será una de las alternativas del conjunto de elección. Para realizar la agrupación, dado la naturaleza cualitativa de las características, empleamos dos técnicas multivariantes, la primera, el análisis cluster combinado con un análisis de correspondencias múltiple y, la segunda, el análisis de clases latentes. Este proceso se aplica para tratar de identificar el perfil del turista quien, habiendo decidido pasar un fin de semana o periodo de descanso en la Comarca del Noroeste de la Comunidad Autónoma de Murcia, tiene mayor probabilidad de elegir un tipo particular de alojamiento que otro de entre los distintos disponibles en la zona.

El capítulo se estructura en 5 epígrafes además de esta introducción. En el segundo epígrafe se desarrollan de forma teórica las dos técnicas que son la

base para la construcción del conjunto de elección. En el tercero se detallan los datos que se usan para definir el conjunto de elección de los turistas rurales de la Región de Murcia y aquellos sobre las características de estos turistas que se usaran para la estimación del modelo logit multinomial. En el cuarto se definen las alternativas de elección entre las que eligen los turistas murcianos y en el quinto se procede a la estimación del modelo logit, que nos mostrará los perfiles de los turistas que eligen cada una de las cuatro alternativas. El capítulo termina exponiendo las conclusiones que se pueden obtener de la aplicación del procedimiento propuesto.

2.2. Construcción del conjunto de elección

Para llevar a cabo el objetivo de conocer el perfil de turista que elige un tipo concreto de alojamiento usaremos el logit multinomial. Este modelo permite determinar la probabilidad de elección de cada una de las alternativas que formen su conjunto de elección en función de las características socio-económicas y de comportamiento de viaje que definen al turista. Por tanto, el primer paso para realizar el estudio es definir el conjunto de elección.

En nuestro contexto, las alternativas que forman parte del conjunto de elección serán los distintos alojamientos situados en un destino rural. Sin embargo, la oferta de turismo rural en un destino se caracteriza, en general, por establecimientos de alojamiento flexibles, independientes y pequeños, que difieren con respecto al número de habitaciones, número de camas y de baños, el tipo de construcción o de alquiler, la ubicación, las instalaciones, etc. La combinación de todos estos atributos junto a la variedad de valores que pueden tomar, implica que para determinar los alojamientos que son similares y que constituirían cada una de las alternativas de elección de nuestro estudio, sea necesario utilizar un método estadístico multivariante. Debido a que los atributos de los alojamientos son fundamentalmente de carácter cualitativo, se propone la utilización de dos métodos que nos permiten usar este tipo de variables. El primero consiste en usar de forma combinada el análisis cluster y el análisis de correspondencias múltiple (Arimond y Elfessi, 2001) y el segundo la utilización del análisis de clases latentes (Lazarsfeld, 1950; Agresti, 2003). Ambas técnicas de agrupación requieren de la especificación a priori del número apropiado de grupos o clases, pero el análisis de clases latentes tiene la ventaja de que existen herramientas para evaluar el ajuste del modelo y para determinar un número apropiado de clases latentes, R , dado un conjunto de datos.

2.2.1. Análisis cluster

La clasificación de los objetos o alternativas en algunas categorías es posible gracias al uso de la técnica de análisis multivariante conocida como análisis cluster (Aldenderfer y Blashfield, 1991). El análisis cluster tradicional requiere el uso de datos cuantitativos. Sin embargo, las características que identifican los alojamientos (tipo de edificio, el alquiler, la posesión de algunas de las instalaciones o servicios, etc.) se basan en datos cualitativos de escala categórica y estos atributos no pueden ordenarse a lo largo de un continuo matemático. Por esto, en este trabajo hemos usado el método cluster de datos categóricos propuesto por Arimond y Elfessi (2001).

Arimond y Elfessi (2001) plantearon un método en dos etapas que permite trabajar con datos de escala categórica para identificar segmentos de mercado. En este estudio, siguiendo esta metodología en dos etapas, agrupamos los alojamientos de turismo rural en un destino por tipos, según las características que los definen. Para ello, en primer lugar, utilizamos un análisis de correspondencia múltiple para cuantificar las variables nominales que describen los alojamientos y después utilizamos el método cluster no jerárquico de las K -medias para definir las diferentes tipologías de alojamientos.

El análisis de correspondencias múltiple (ACM) (Greenacre, 1984; Lebart et al., 1984) es una técnica multivariante que describe las relaciones que existen entre algunas variables cualitativas (en escala nominal) dentro de un espacio de menor dimensión que contiene las categorías de las variables y los objetos que pertenecen a las categorías. Así, el propósito de aplicar el ACM a los datos es encontrar una cuantificación óptima para las categorías de las variables y, simultáneamente, para los objetos, de manera que los objetos que pertenecen a una misma categoría se encuentran situados próximos entre sí, mientras que aquellos de diferentes categorías se mantengan tan distantes como sea posible. Esta técnica es similar al análisis de componentes principales no lineal de datos nominales (Molera y Albaladejo, 2007). En el ACM se utiliza la distancia chi-cuadrado para medir separaciones entre puntos, ya que cumple el principio de equivalencia distribucional que postula que si dos categorías tienen perfiles idénticos pueden ser sustituidas por una sola categoría que sea la suma de sus pesos, sin que se alteren las distancias entre los demás pares de puntos.

Para la segunda fase del análisis, se utiliza el método cluster de las K -medias, que busca clasificar los alojamientos en grupos internamente homogéneos, en función de las similitudes o similaridades entre ellos, y mutuamente heterogéneos. El método de la K -medias, que requiere datos de

escala-intervalo o ratio, es un método de partición iterativa (Arimond y Elfessi, 2001). El algoritmo en este procedimiento estadístico requiere de una determinación a priori del número de grupos, K . Una vez fijado, el algoritmo comienza asignando, de forma arbitraria, una observación a cada uno de los grupos predefinidos, constituyendo el centroide inicial, para después incorporar el resto de observaciones al grupo más cercano (en términos de distancia euclídea entre la observación y el centroide del grupo). Tras esta asignación, el algoritmo calcula de nuevo el centroide para cada grupo y reasigna las observaciones al grupo más cercano. Los grupos cambian iterativamente a través de la reasignación de los objetos al centroide más cercano y al cálculo de un nuevo centroide, hasta que los grupos cambian poco entre etapas, o bien se llega al número máximo de iteraciones fijado a priori.

2.2.2. Análisis de clases latentes

El análisis de clases latentes (ACL) es una técnica estadística sobre datos categóricos multivariantes que permite clasificar a los individuos de una población en segmentos o grupos de naturaleza exhaustiva y excluyentes. Así, los modelos de clases latentes constituyen, tal y como señalan Linzer y Lewis (2011), una herramienta poderosa y útil para investigar las fuentes de confusión entre las variables categóricas observadas, para identificar y caracterizar grupos de individuos similares y aproximar la distribución de las observaciones a través de las variables de estudio.

El análisis de clases latentes se basa en la suposición de que las relaciones de dependencia entre las variables categóricas de una tabla de contingencia se deben a la existencia de una asociación entre cada una de ellas y otra variable, no observable directamente, llamada *variable latente*. Así, este análisis busca segmentar la tabla de contingencia creada a partir de las variables categóricas observadas (o “manifiestas”) por esa variable categórica no observada (“latente”), teniendo en cuenta la suposición de que las respuestas a todas las variables manifiestas son estadísticamente independientes con respecto a los valores de dicha variable latente. Este supuesto, denominado *independencia local o condicional*, implica que las variables latentes causan la relación existente entre las variables categóricas observadas, por consiguiente, no existe una relación directa entre dichas variables observadas (Castro y Tenorio, 2010).

Otro supuesto básico en los modelos de clase latente es el de *homogeneidad interna* de las variables latentes, que se utiliza para diferenciar a los individuos pertenecientes a las diferentes clases. Esto es, cada uno de los miembros

de una clase latente tiene una distribución de probabilidad igual respecto a la de la variable latente, pero diferente a la de los individuos pertenecientes a otra clase.

Consideramos un modelo en el que observamos J variables categóricas politómicas (las variables “manifiestas”) para los alojamientos $i = 1, \dots, N$, cada una de las cuales con K_j modalidades posibles. El modelo de clase latente aproxima la distribución conjunta observada de las variables manifiestas como la suma ponderada de un número finito, R , de tablas de clasificación cruzada, donde R es desconocido e indica el número de clases latentes o niveles de la variable latente inobservada que explica las relaciones existentes entre las variables categóricas observadas.

Denotemos como Y_{ijk} los valores observados de las J variables manifiestas, de forma que $Y_{ijk} = 1$ si el alojamiento i presenta la modalidad k en la variable j , e $Y_{ijk} = 0$ en caso contrario, donde $j = 1, \dots, J$ y $k = 1, \dots, K_j$ y sea $\pi_{jk/r}$ la probabilidad condicionada de que una observación en la clase $r = 1, \dots, R$ genere el resultado k en la variable j . De manera que, dentro de cada clase, para cada variable manifiesta se cumple que,

$$\sum_{k=1}^{K_j} \pi_{jk/r} = 1. \quad (2.1)$$

Por otro lado, se denotan por p_r las proporciones de mezcla que proporcionan los pesos en la suma ponderada de las tablas de componentes, con

$$\sum_{r=1}^R p_r = 1 \quad (2.2)$$

y que representan la probabilidad incondicional de que un alojamiento pertenezca a cada clase antes de tener en cuenta las respuestas Y_{ijk} proporcionadas en las variables manifiestas. A los valores de p_r también se les denominan probabilidades “a priori” de pertenencia a la clase latente.

Así, la probabilidad de que un alojamiento i de la clase r genere un determinado conjunto de J resultados sobre las variables manifiestas, suponiendo la independencia condicional de los resultados Y dada la pertenencia a la clase r , es

$$P(Y_i; \pi_r) = \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jk/r})^{Y_{ijk}} \quad (2.3)$$

La función de densidad de probabilidad a través de todas las clases es la suma ponderada

$$P(Y_i/\pi, p) = \sum_{r=1}^R p_r \prod_{j=1}^J \prod_{k=1}^{K_j} (\pi_{jk/r})^{Y_{ijk}} \quad (2.4)$$

De manera que los parámetros a estimar por el modelo de clase latente son la proporción de observaciones en cada clase latente, p_r , y las probabilidades de observar cada respuesta de cada variable manifiesta, condicional a la clase latente $\pi_{jr/k}$.

Dadas las estimaciones \hat{p}_r y $\hat{\pi}_{jk/r}$ de los parámetros p_r y $\pi_{jr/k}$, respectivamente, se determina la probabilidad *a posteriori* de pertenencia de los individuos a cada clase condicionada a los valores observados de las variables manifiestas, mediante la fórmula de Bayes:

$$\hat{P}(r_i/Y_i) = \frac{\hat{p}_r P(Y_i; \hat{\pi}_r)}{\sum_{q=1}^R \hat{p}_q P(Y_i; \hat{\pi}_q)} \quad (2.5)$$

donde $r_i \in \{1, \dots, R\}$. Por último, los alojamientos se asignan a una de las clases en función del criterio de máxima probabilidad *a posteriori* de pertenencia a una clase.

Es importante señalar que el número de parámetros independientes a estimar en el modelo de clases latentes es

$$R \sum_{j=1}^J (K_j - 1) + (R - 1), \quad (2.6)$$

y, depende, por tanto, del número de clases latentes (R), del número de variables categóricas (J) y del número de modalidades de cada variable (K_j). Si el número definido en (2.6) supera el número total de observaciones, o uno menos que el número total de celdas en la tabla de clasificación cruzada de las variables manifiestas, habría un problema, ya que el modelo de clase latente no podría ser identificado (Linzer y Lewis, 2011).

Tal y como se indicó al principio de este epígrafe, el análisis de clases latentes tiene la ventaja de que existen herramientas para evaluar el ajuste del modelo y determinar un número apropiado de clases latentes R , dado un conjunto de datos. No obstante, hay que tener en cuenta que cuando se añade una clase, el ajuste del modelo mejora, pero incorpora ruido y aumenta en número de parámetros a estimar (véase ecuación 2.6), por lo que es necesario

tener en consideración un criterio de parsimonia¹ que establezca un equilibrio entre la mejora del ajuste y la cantidad de parámetros que se incorporan al modelo al aumentar una clase. Las dos medidas de parsimonia más utilizadas son el criterio de información de Akaike, o AIC (Akaike, 1973),

$$AIC = -2\Lambda + 2k$$

y el criterio de información bayesiano, o BIC (Schwartz, 1978),

$$BIC = -2\Lambda + k \ln N$$

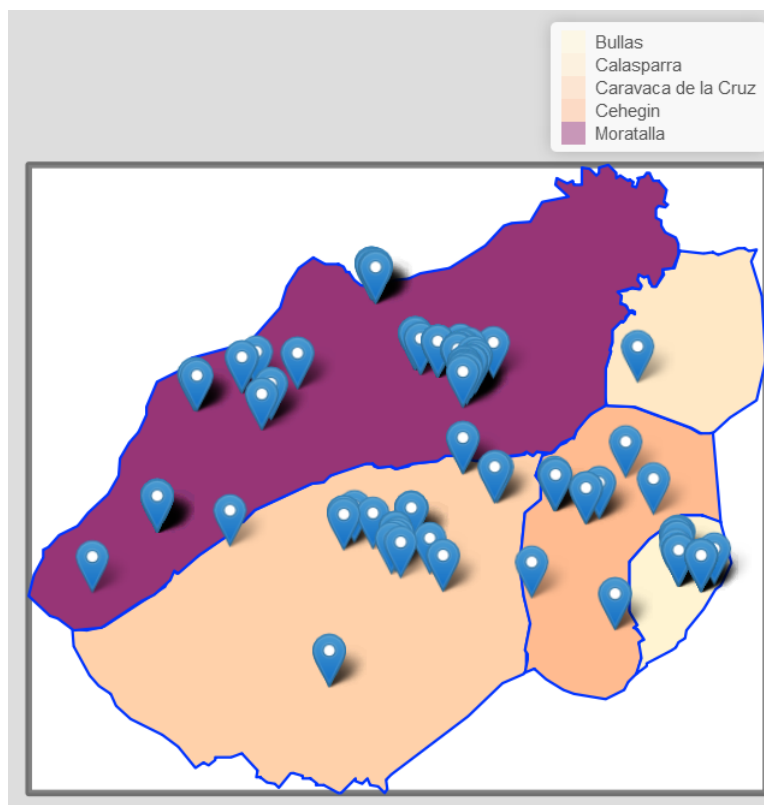
donde Λ es el logaritmo de la función de verosimilitud del modelo donde se alcanza el máximo, k el número total de parámetros estimados y N el número de observaciones. Los modelos preferidos con ambos criterios son aquellos que minimizan los valores del BIC y/o del AIC. En el contexto de los modelos de clase latente básicos, es decir, aquellos que no incluyen covariables, el BIC suele ser más apropiado debido a su relativa simplicidad (Lin y Dayton, 1997; Forster, 2000; Greene y Hensher, 2003).

2.3. Datos

Para llevar a cabo el estudio aplicado de este capítulo se van a utilizar dos conjuntos de datos. El primero contiene la información de las características socio-económicas y de comportamiento de viaje de los turistas y del alojamiento elegido para hospedarse durante ese fin de semana. Estos datos proceden de la encuesta que se hizo en la Comarca Noroeste de la Región de Murcia durante varios fines de semana de otoño de 2003 y ya explicada en el epígrafe 1.5 de esta memoria. El segundo conjunto de datos contiene la información de la oferta de establecimientos de alojamiento de turismo rural existente en la Comarca del estudio, así como de sus atributos.

Para recoger información de los alojamientos de turismo rural, se utilizó Guías de turismo rural en papel y distintas páginas web donde estos alojamientos se publicaban, como la de NORATUR o la página web de la Comunidad Autónoma de la Región de Murcia. Pero, también, en alguna ocasión hubo que completar la información de las casas llamando por teléfono a los propietarios o encargados de los establecimientos. De esta forma se recogieron datos sobre 199 alojamientos, cerca del 89 % de la oferta total de la

¹Los criterios de parsimonia tratan de encontrar un equilibrio entre el exceso y la falta de adecuación del modelo a los datos, penalizando la log-verosimilitud en función del número de parámetros estimados.

Figura 2.1: Distribución de los alojamientos rurales estudiados

†Elaboración propia utilizando la librería *leaflet* de R

Municipio	Casa Rural	Hospedería rural	Campings
Bullas	14	1 con 8 habitaciones	8
Calasparra	1	—	—
Caravaca	34	—	—
Cehegín	15	—	—
Moratalla	101	1 con 9 habitaciones	9

zona en 2003, año de implementación de la encuesta. Estos establecimientos se encontraban distribuidos por toda la comarca (véase Figura 2.1).

Los datos obtenidos de los alojamientos nos daban información sobre su precio, su tamaño, el tipo de construcción, su ubicación, de las instalaciones que disponía, el tipo de alquiler, sobre su equipamiento y prestaciones, de los servicios que ofrecía y si contaba con la ‘Q’ de calidad. Con esta información, construimos un total de 13 variables que utilizaremos para la clasificación de los alojamientos rurales en grupos. Estas variables junto con sus diferentes niveles se muestran en la Tabla 2.1. También se indica en esta tabla el

porcentaje de casas que presenta cada uno de los niveles de las variables consideradas (tercera columna).

2.4. Clasificación de los alojamientos rurales

Para clasificar los alojamientos de la Comarca del Noroeste de la Región de Murcia en grupos homogéneos, de manera que los alojamientos del mismo grupo tengan características similares, utilizamos el análisis cluster con datos categóricos y el análisis de clases latentes. Las variables descritas en la Tabla 2.1 constituyen el input para estos dos métodos.

2.4.1. Análisis cluster

Previo a aplicar el análisis cluster, vamos a transformar los atributos que describen los alojamientos (véase Tabla 2.1) en otras variables, de tipo cuantitativo (*dimensiones*), realizando un análisis de correspondencia múltiple (ACM) sobre la matriz 199x13, que recoge la información de las trece variables definidas en la Tabla 2.1 para cada uno de los 199 alojamientos rurales considerados. El ACM se ha realizado usando la librería *FactoMineR* de R (Lê et al., 2008).

Con el ACM se obtienen 21 dimensiones, que son el resultado de la diferencia entre el número de categorías ($K = 34$) menos el número de variables cualitativas ($p = 13$). De estas 21 dimensiones, hemos considerado aquellas cuya inercia es superior a la inercia media ($\frac{IT}{p} = 0,12$), y que, además, recogen un porcentaje adecuado de la inercia total ($IT = 1,62$) del conjunto completo de datos. En este caso, la información sobre las variables nominales se reduce a un conjunto de cuatro ejes o dimensiones cuantitativas cuyas principales inercias son 0,31, 0,23, 0,15 y 0,12², que explican un 50,48 % de la inercia total.

La Tabla 2.2 muestra la contribución de cada una de las variables cualitativas iniciales a la inercia principal de las cuatro dimensiones elegidas. Por lo tanto, la primera dimensión, cubre las variables que indican el número de habitaciones, el tipo de alquiler, si el uso del cuarto de baño es compartido o no, el precio de la vivienda, el tipo de construcción y la posibilidad de

²La dimensión 4 se ha incluido para garantizar que con las cuatro dimensiones al menos se explique el 50 % de la inercia total y porque, además, la inercia explicada por esta dimensión está cercana a la inercia media.

Tabla 2.1: Descripción de los alojamientos

Características	Niveles	%
Precio	0,6-0,9616	20,10
	0,9616-1,6828	34,67
	1,6828-2,3139	25,63
	2,3139-4,3273	19,60
Número de habitaciones	1	17,09
	2	35,68
	3	26,13
	4 ó más	21,11
Tipo de construcción	Original	56,28
	Nueva	24,12
	Madera	9,55
	Molino	10,05
Ubicación	Aldea	7,54
	Huerta	32,16
	Campo-montaña	41,21
	Pueblo	19,10
Cuartos de baño	Compartido	77,39
	Sin compartir ^a	22,61
Alquiler de caballos	Sí	36,68
	No	63,32
Parque infantil	Sí	36,68
	No	63,32
Servicio de comidas	Sí	42,71
	No	57,29
Piscina	Sí	71,86
	No	28,14
Instalaciones deportivas	Sí	36,68
	No	63,32
Minigranja	Sí	14,57
	No	85,43
Tipo de alquiler	Casa completa	88,44
	Por habitaciones	11,56
Posee la 'Q' de calidad	Sí	3,02
	No	96,98

^aSe considera "sin compartir" si el alojamiento dispone de habitaciones con cuarto de baño integrado o el número de cuartos de baño coincide con el número de habitaciones de la vivienda manera que no sea necesario que los turistas compartan el servicio.

Tabla 2.2: Análisis de correspondencia múltiple

Variab les	Dim 1	Dim 2	Dim 3	Dim 4
Precio	0,6033	0,0423	0,7173	0,4373
Número de habitaciones	0,8344	0,0157	0,6031	0,4134
Tipo de construcción	0,5557	0,2985	0,2210	0,3193
Ubicación	0,0228	0,5509	0,0952	0,2096
Aseos	0,6742	0,0423	0,0002	0,0092
Parque infantil	0,0151	0,5208	0,0007	0,0004
Servicio de comidas	0,2349	0,1211	0,0463	0,0086
Piscina	0,0420	0,2690	0,0764	0,0892
Instalaciones deportivas	0,1287	0,5333	0,0080	0,0002
Minigranja	0,0316	0,3614	0,0566	0,0174
Caballos	0,1034	0,2761	0,0580	0,0081
Alquiler	0,7737	0,0102	0,0271	0,0123
Posee la "Q" de Calidad	0,0190	0,0067	0,0163	0,0631
<i>Inercia</i>	0,3107	0,2345	0,1482	0,1222
<i>% Inercia explicada</i>	19,23 %	14,52 %	9,17 %	7,56 %
<i>% Inercia explicada acumulada</i>	19,23 %	33,75 %	42,92 %	50,48 %

contratar un servicio de comidas. La segunda dimensión está constituida por las variables que hacen referencia a la ubicación del alojamiento rural y a los equipamientos de que dispone, esto es, la disponibilidad de instalaciones deportivas, parque infantil y minigranja. También influyen en esta dimensión el tipo de construcción, la posibilidad de montar a caballo y la disponibilidad de piscina. Esta conclusión también se puede obtener si observamos el primer gráfico de la Figura 2.2 donde las variables que están más próximas al eje de abscisas y más alejadas del origen de coordenadas son las que mejor están representadas por la dimensión 1 y, por tanto, la definen; mientras que las que están más próximas al eje de ordenadas y más alejadas del origen definirían a la dimensión 2. En este gráfico también puede observarse que la variable "Posee la 'Q' de calidad" no está bien representada por ninguna de estas dos dimensiones al estar situada próxima al origen de coordenadas.

Continuando con la descripción de las dimensiones (véase Tabla 2.2), la tercera dimensión está definida básicamente por el precio y el número de habitaciones, aunque también influye, en menor medida, el tipo de construcción; mientras que las variables de mayor importancia en la cuarta dimensión son el precio, el número de habitaciones, el tipo de construcción y la ubicación del alojamiento rural. El resto de variables aparecen cerca del origen de coordenadas (véase segundo gráfico de la Figura 2.2) indicando que no

Figura 2.2: Representación gráfica de las variables nominales en el espacio

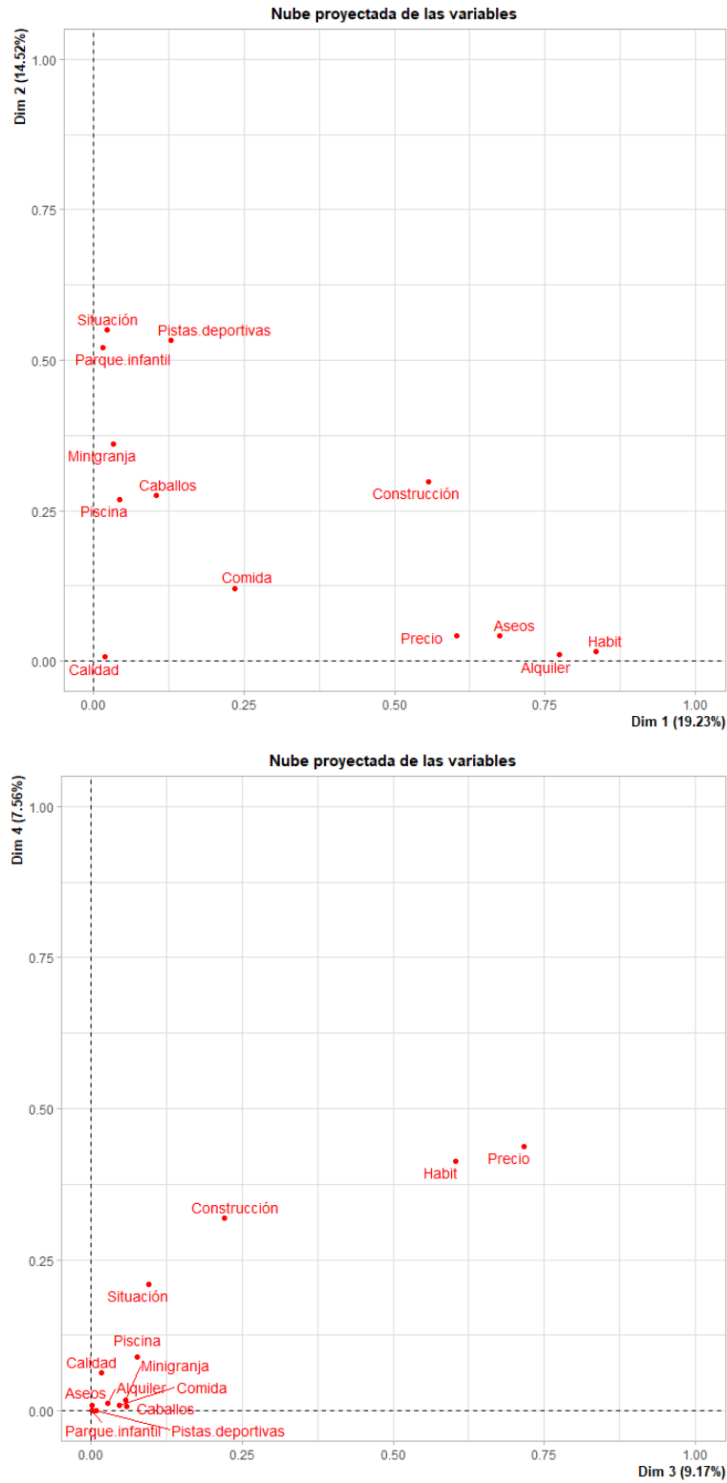
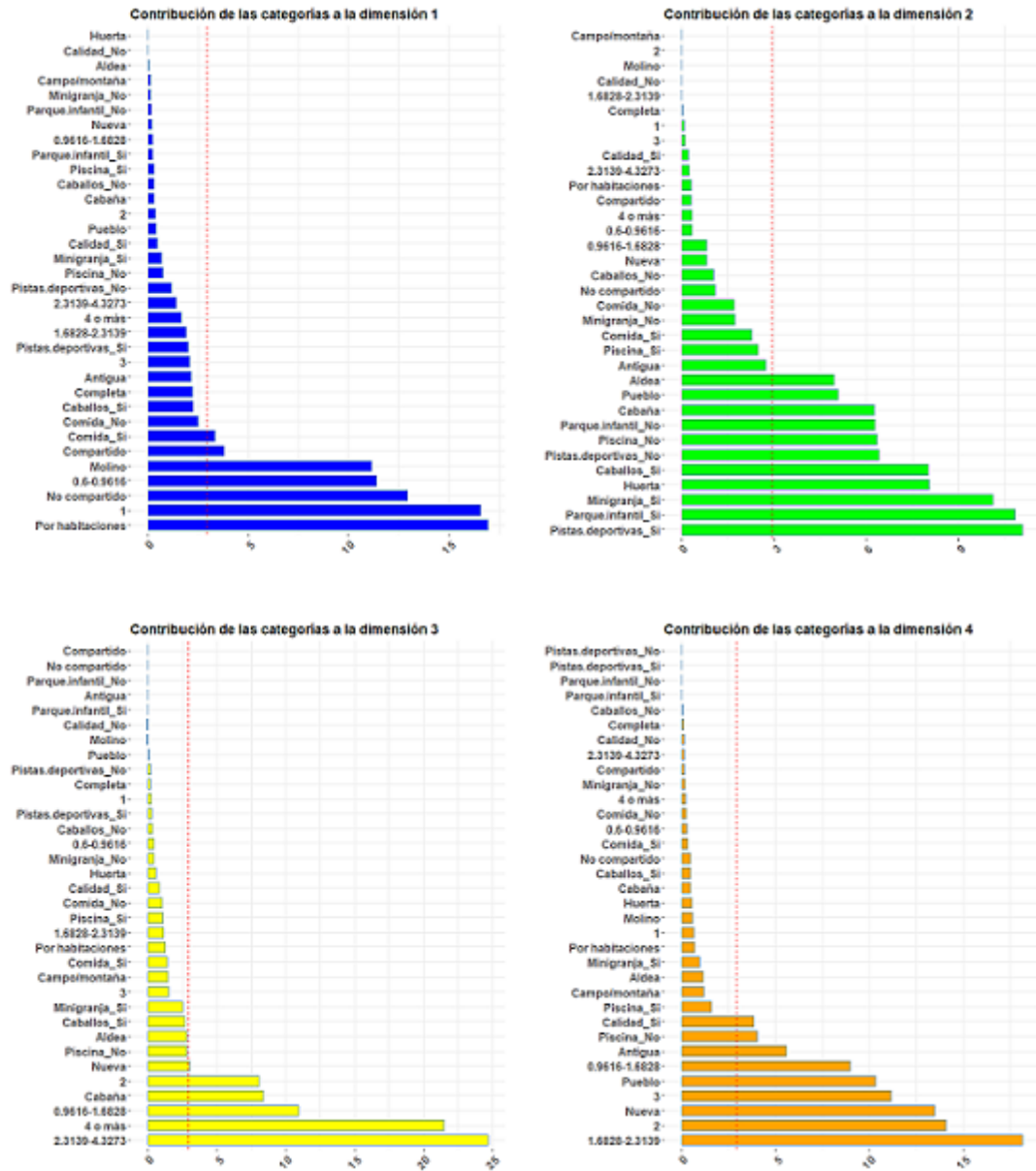


Figura 2.3: Contribuciones de las categorías de las variables a las dimensiones



influyen en las dimensiones 3 y 4. No obstante, aunque ninguna de las cuatro dimensiones representa bien la posibilidad de poseer la ‘Q’ de calidad, es la dimensión 4 la que la representa mejor. Se puede ver que está cerca de este

eje y más alejado del origen de coordenadas.

En la Figura 2.3 podemos ver qué categorías de las variables contribuyen más a cada dimensión, es decir, explican más porcentaje de inercia de esa dimensión. Por ejemplo, las categorías ‘tipo de alquiler por habitaciones’, ‘1 habitación’, ‘aseo sin compartir’, ‘precio por día de la vivienda entre 60 y 96,16 euros’ y ‘tipo de construcción molino’ definen la dimensión 1. Por otro lado, que el alojamiento rural disponga de alquiler de caballos, minigranja, parque infantil y pistas deportivas, esté situada en la huerta y sea una cabaña son determinantes para la dimensión 2. En la dimensión 3 que el precio del alojamiento esté comprendido entre 231,39 y 432,73 euros explica aproximadamente el 25% de la inercia de esa dimensión. Otras categorías con peso en la definición de la dimensión 3 son que el alojamiento disponga de 4 o más habitaciones y el tipo de construcción sea cabaña. Finalmente, en la dimensión 4, las categorías con mayor peso son que el precio oscile entre 168,28 y 231,39 euros, disponga de dos o tres habitaciones, la vivienda sea de nueva construcción y esté situada en un pueblo.

Para la segunda etapa de este análisis, que busca clasificar los alojamientos en grupos internamente homogéneos y mutuamente heterogéneos, se ha llevado a cabo un análisis cluster no jerárquico de K-medias, utilizando la función *kmeans* disponible en R (R Core Team, 2019). El análisis cluster se realizó utilizando, como variables “input”, las puntuaciones (coordenadas) de cada alojamiento en el sistema formado por las cuatro dimensiones cuantitativas obtenidas a partir del análisis de correspondencia múltiple, y probando con diferentes números de grupos (2-7). Esto último se hizo con el fin de encontrar el número de grupos óptimo para clasificar estos 199 alojamientos rurales. Para ayudarnos con la identificación del número adecuado de grupos o conglomerados, nos hemos apoyado en la librería *Nbclust* de R (Charrad et al., 2014). Esta librería proporciona 30 índices que sirven para determinar el número óptimo de conglomerados y propone al usuario el mejor esquema de agrupación a partir de los diferentes resultados obtenidos. Según la información proporcionada, la solución que más indicadores recomiendan es la división en cuatro grupos. Además, esta es la agrupación que produce los grupos más claros y más distinguibles de alojamientos. Finalmente, los tamaños de los grupos obtenidos fueron, respectivamente, de 32, 57, 86 y 24 alojamientos para los cluster del 1 al 4. De este modo, el grupo 3 representa la proporción mayor de establecimientos de alojamiento (43,2%), seguido por el grupo 2 (28,6%), repartiéndose el resto de establecimientos entre los grupos 1 (16,1%) y 4 (12,1%).

Tabla 2.3: Porcentaje de casas por cluster según el nivel de cada variable

Variable	Niveles	Grupo 1	Grupo 2	Grupo 3	Grupo 4
Precio	0,6-0,9616		18,97	6,98	95,83
	0,9616-1,6828		34,48	55,81	4,17
	1,6828-2,3139	3,23	32,76	36,05	
	2,3139-4,3273	96,77	13,79	1,16	
Número de habitaciones	1		1,72	10,47	100,00
	2		50,00	48,84	
	3	6,45	32,76	36,05	
	4 o más	93,55	15,52	4,65	
Tipo de Construcción	Antigua	74,19	36,21	79,07	
	Cabaña		32,76		
	Molino			5,81	62,50
	Nueva	25,81	31,03	15,12	37,50
Situación	Aldea			17,44	
	Campo/montaña	58,06	29,31	36,05	66,67
	Huerta	12,90	70,69	12,79	33,33
	Pueblo	29,03		33,72	
Aseos	Compartido	93,55	98,28	79,07	
	No compartido	6,45	1,72	20,93	100,00
Parque Infantil	No	74,19	12,07	94,19	62,50
	Sí	25,81	87,93	5,81	37,50
Servicio de comidas	No	87,10	43,10	72,09	
	Sí	12,90	56,90	27,91	100,00
Piscina	No	22,58	10,34	50,00	
	Sí	77,42	89,66	50,00	100,00
Instalaciones deportivas	No	58,06	8,62	91,86	100,00
	Sí	41,94	91,38	8,14	
Minigranja	No	93,55	56,90	97,67	100,00
	Sí	6,45	43,10	2,33	
Caballos	No	100,00	75,86	98,84	66,67
	Sí		24,14	1,16	33,33
Alquiler	Completa	96,77	100,00	100,00	8,33
	Por habitaciones	3,23			91,67
Calidad	No	93,55	94,83	98,84	100,00
	Sí	6,45	5,17	1,16	

La Tabla 2.3 se ha preparado para facilitar la lectura de los resultados y muestra el porcentaje de alojamientos en cada grupo que pertenecen a cada una de las categorías definidas para cada una de las trece variables nominales de este estudio. Así, el Grupo 1 incluye principalmente los establecimientos

de alojamiento más grandes (el 93,6 % tienen 4 o más habitaciones), es decir, aquellos con cuatro o más habitaciones y cuyo precio oscila entre 231,39 y 432,73 euros (96,8 %), de construcción original (74,2 %), localizados en el campo o en la montaña (58 %), y que poseen piscina (77,4 %). En este grupo no hay casas de madera ni molinos reconvertidos en vivienda. Tampoco hay establecimientos situados en aldeas.

El Grupo 2 abarca los alojamientos de dos o tres dormitorios que están situados fuera de un núcleo de población, principalmente en la huerta y algunos en el campo, y que poseen equipamientos adicionales tales como instalaciones deportivas, parque infantil y piscina. Aunque los alojamientos de este grupo no se definen por el tipo de construcción, en este grupo están incluidas todas las casas de madera y todos los alojamientos que se encuentran en campings. En este grupo también están incluidos el 86,2 % de los alojamientos que disponen de minigranja y 61,9 % de los que ofertan el servicio de montar a caballo. También disponen de servicio de comidas.

El Grupo 3, que es el más numeroso, se compone de los establecimientos de alojamiento de dos o tres dormitorios que carecen de equipamientos. Es decir, no disponen de los servicios de parque infantil, pistas deportivas y minigranja. No ofertan tampoco servicio de alquiler de caballos ni de comidas. Poseen la arquitectura típica de la zona. Todos los alojamientos situados en aldeas y el 76,3 % de los situados en un pueblo se encuentran en este grupo.

Por último, el Grupo 4 incluye aquellos alojamientos que son molinos que han sido reconvertidos en establecimientos de alojamiento y los cuales se alquilan por habitaciones individuales cada una con su propio cuarto de baño. Están situados a cierta distancia de zonas urbanizadas, y disponen de servicio de comida y tienen piscina. En este grupo están incluidos todos los alojamientos con una habitación y de precio más bajo, entre 60 y 96,16 euros.

2.4.2. Análisis de clases latentes

El análisis de clases latentes (ACL) es el otro método de agrupación elegido para clasificar los alojamientos en grupos (o clases latentes) homogéneos, de manera que los objetos de un mismo grupo sean más similares entre sí que los de otros grupos. Este método se aplica directamente sobre la matriz 199x13 que recoge la información de las trece variables definidas en la Tabla 2.1 para cada uno de los 199 alojamientos rurales considerados. La estimación del ACL se ha realizado con la librería *poLCA* de R (Linzer y Lewis, 2011).

El ACL se ha realizado comenzando un modelo completamente independiente con únicamente una clase latente ($R = 1$), y luego se ha ido aumentando iterativamente el número de clases latentes hasta conseguir un modelo adecuado según el Criterio de Información Bayesiano (BIC). La solución adoptada ha sido la división en cuatro clases (véase Tabla 2.4). Además, de nuevo, esta es la agrupación que produce las clases más claras y más distinguibles de alojamientos. Los tamaños de las clases numeradas del 1 al 4 son, respectivamente, 47, 59, 67 y 26. De este modo, la clase 3 representa la proporción mayor de establecimientos de alojamiento (33,7%), seguida por la clase 2 (29,6%), la clase 1 (23,6%) y, por último, la clase 4 (13,1%).

Tabla 2.4: BIC según número de clases latentes

Número de grupos	BIC
3	3366,2045
4	3347,0788
5	3383,6905
6	3417,2790
7	3452,7701

En la Tabla 2.5 se describen los tipos de alojamientos utilizando la probabilidad condicional estimada de cada una de las categorías definidas para cada una de las trece variables nominales de este estudio, dada la permanencia a una clase ($\hat{\pi}_{jk/r}$). Observando esta tabla podemos establecer que la Clase 1 incluye principalmente los establecimientos de alojamiento más grandes, es decir, aquellos con tres o más habitaciones, cuyo precio es superior a 168,28 euros, de construcción original, localizados en el campo o en la montaña y que poseen piscina.

La Clase 2 abarca los alojamientos de dos o tres dormitorios que están situados principalmente en la huerta, y que poseen equipamientos tales como parque infantil, servicio de comidas, instalaciones deportivas y piscina. Aunque los alojamientos de este grupo no se definen por el tipo de construcción, en él están incluidas todas las casas de madera y todos los alojamientos de la zona que se encuentran en campings. A grupo también pertenecen el 86,2% de los alojamientos que disponen de minigranja y 56,5% de los que ofertan el servicio de montar a caballo.

En la Clase 3, que es la más numerosa, predominan establecimientos con dos dormitorios, de construcción original y que el único equipamiento con el

Tabla 2.5: Porcentaje (estimado) de casas por clase latente según el nivel de cada variable

Variable	Niveles	Clase 1	Clase 2	Clase 3	Clase 4
Precio	0,6-0,9616		18,17	8,96	89,17
	0,9616-1,6828	2,65	33,76	68,02	6,95
	1,6828-2,3139	36,42	31,29	23,02	
	2,3139-4,3273	60,93	16,79		3,88
Número de habitaciones	1			13,75	96,12
	2		49,50	61,25	
	3	35,37	31,55	25,00	
	4 o más	64,63	18,95		3,88
Tipo de Construcción	Antigua	78,95	41,67	74,21	3,88
	Cabaña		31,38		
	Molino			7,77	57,36
	Nueva	21,05	26,95	18,02	38,76
Situación	Aldea			22,40	
	Campo/montaña	50,90	33,79	32,10	65,12
	Huerta	23,42	66,21	6,29	34,88
	Pueblo	25,68		39,22	
Aseos	Compartido	97,80	100,00	72,82	
	No compartido	2,20		27,18	100,00
Parque Infantil	No	82,23	11,08	96,95	65,10
	Sí	17,77	88,92	3,05	34,90
Servicio de comidas	No	91,35	46,50	65,86	
	Sí	8,65	53,50	34,14	100,00
Piscina	No	38,31	10,38	48,10	
	Sí	61,69	89,62	51,90	100,00
Instalaciones deportivas	No	67,68	9,59	96,27	96,14
	Sí	32,32	90,41	3,73	3,86
Minigranja	No	100,00	58,77	96,95	92,26
	Sí		41,23	3,05	7,74
Caballos	No	100,00	78,53	98,05	65,12
	Sí		21,47	1,50	34,88
Alquiler	Completa	100,00	100,00	100,00	10,82
	Por habitaciones				89,18
Calidad	No	93,50	95,00	100,00	100,00
	Sí	6,50	5,00		

que cuentan es la piscina. Tampoco ofertan servicio de alquiler de caballos ni de comidas. Todos los alojamientos situados en aldeas y el 68,4% de los situados cerca de un pueblo pertenecen a este grupo.

Por último, la Clase 4 incluye aquellos alojamientos que antiguamente eran molinos y que actualmente han sido reconvertidos en establecimientos

de alojamiento, que se alquilan por habitaciones individuales cada una con su propio baño. Están situados a cierta distancia de zonas urbanizadas, en el campo o en la huerta, y todos disponen de servicio de comida y tienen piscina. Dado que fundamentalmente se trata de alojamientos que se alquilan por habitaciones son el grupo de precio más bajo, entre 60 y 96,16 euros.

2.4.3. Comparación de los resultados obtenidos con ambos métodos de clasificación

Con el fin de realizar un análisis comparativo entre los resultados obtenidos con ambas técnicas se ha elaborado la Tabla 2.6. Si nos fijamos en la diagonal principal de esta tabla de contingencia, podemos observar que prácticamente existe una correspondencia casi perfecta³ entre los grupos de alojamientos obtenidos mediante análisis cluster y la técnica de clases latentes, poniendo de relieve la robustez de la clasificación de los alojamientos rurales del área de estudio en cuatro tipos.

Tabla 2.6: Número de alojamientos por cluster según clase latente

Cluster / Clase Latente	Clase 1	Clase 2	Clase 3	Clase 4	Total
Grupo 1	28	3		1	32
Grupo 2	1	55		1	59
Grupo 3	18	1	67		96
Grupo 4				24	24
Total	47	59	67	26	199

Finalmente, una vez clasificados los 199 alojamientos rurales en estos cuatro tipos diferentes, el siguiente paso es determinar, mediante un análisis logit, las características socio-económicas y/o condiciones de viaje que son relevantes en la elección de un tipo concreto de alojamiento. Dado que los resultados obtenidos con el análisis de clases latentes son similares⁴ a los obtenidos con el análisis cluster⁵, en el epígrafe 2.5 presentamos los resultados

³El 87,4% de los alojamientos rurales considerados en este estudio han sido asignados al mismo grupo por ambas técnicas.

⁴La distribución de las observaciones muestrales (elección de alojamiento observado) entre los grupos del AC y las clases latentes del ACL coinciden en un 91,2%.

⁵En el Anexo 4.8 se puede consultar los resultados obtenidos con el modelo logit donde las alternativas de elección son los cuatro grupos obtenidos mediante el análisis cluster (modelo LAC).

obtenidos utilizando, como alternativas de elección, la agrupación obtenida con el análisis de clases latentes.

2.5. Perfil del turista por tipo de alojamiento

Como ya se ha indicado previamente, el perfil del turista que es más probable que opte por uno de los cuatro tipos de alojamientos disponibles en la zona Noroeste de la Región de Murcia se determinará dentro del marco de un modelo logit multinomial. Por lo tanto, la variable endógena es discreta y toma valores entre 1 y 4, que simbolizan cada uno de los cuatro tipos de alojamientos determinados con el análisis de clases latente⁶.

Las variables independientes son factores socio-económicos y relacionados con el viaje de los turistas que se han definido a partir de la información recabada en la encuesta realizada ad-hoc para este trabajo (véase Anexo I). Además de las variables recogidas en la Tabla 1.2 del epígrafe 1.5 de esta memoria, se han definido cuatro variables adicionales⁷, que son “actividades deportivas”, “actividades asociadas con el entorno”, “actividades gastronómicas” y, por último, “actividades culturales”. La variable “actividades deportivas” recoge la intención del turista por realizar cualquiera de las siguientes actividades: senderismo, hacer el descenso del río Segura en canoa, paseos a caballo o en bicicleta, mientras que la variable “actividades asociadas con el entorno” incluiría actividades como caminar o visitar lugares cercanos por su relevancia paisajística u orográfica. La variable “actividades gastronómicas” hace referencia al propósito del turista por degustar platos típicos de la zona o por aprender a cocinarlos. También estarían incluidas, actividades como hacer pan y dulces. Para finalizar, la variable “actividades culturales” incluyen la visita a lugares de interés arquitectónico o cultural o participar en cualquier evento de interés cultural organizado en la zona. Estas variables junto al resto, con sus distintas modalidades y la modalidad de referencia⁸ en la estimación del modelo se muestran en la Tabla 2.7.

⁶Para probar la robustez de los resultados obtenidos, se ha procedido también a estimar un logit multinomial utilizando como conjunto de elección los cuatro grupos de alternativas obtenidos aplicando el análisis cluster. Como se puede comprobar en la Tabla 8 del Anexo 4.8 los resultados resultan muy similares.

⁷Estas variables se han definido a partir la información aportada por la pregunta P19 del cuestionario incluido en el Anexo 4.8

⁸La modalidad de referencia es el nivel que no se usa en la estimación del modelo para evitar problemas de multicolinealidad (Greene, 1999).

Con los 307 datos obtenidos en las entrevistas que se realizaron, durante los fines de semana de otoño de 2003, a turistas que pasaban al menos una noche en un alojamiento rural de la Comarca de Noroeste se han estimado varios modelos logit, usando la librería *mlogit* de R (Croissant, 2020). Teniendo en cuenta que no hay observaciones repetidas, el método utilizado es el de máxima verosimilitud. En primer lugar, se ha estimado un modelo logit tomando como alternativa base el grupo 1 de alojamientos y considerando que la parte determinista de la utilidad que el individuo i obtiene por la elección de la alternativa j viene dada por la expresión,

$$\begin{aligned}
V_{ij} = & ACS_j + \beta_{1j}GEN_i + \beta_{2j}PROC_i + \beta_{3j}EDAD_{25i} \\
& + \beta_{4j}EDAD_{2640i} + \beta_{5j}ESTSUP_i + \beta_{6j}TIPFAM_{1i} \\
& + \beta_{7j}TIPFAM_{2i} + \beta_{8j}INGFAM_i + \beta_{9j}ALQ_i \\
& + \beta_{10j}REFLUG_i + \beta_{11j}PERS_{1i} + \beta_{12j}PERS_{2i} \\
& + \beta_{13j}GAST_{1i} + \beta_{14j}DEPOR_i \\
& + \beta_{15j}COMIDA_i + \beta_{16j}CULTURA_i \quad j = 1, 2, 3, 4, \quad \forall i
\end{aligned} \tag{2.7}$$

Después de esta primera estimación, se han eliminado de modelo aquellas variables menos significativas según la prueba de t-student. Una vez eliminadas, se ha realizado una nueva estimación con el resto de variables explicativas, teniendo en cuenta diferentes combinaciones. Se ha repetido el proceso hasta que todas las variables explicativas incluidas en el modelo ha tenido un nivel de significación de 10 % o al eliminar esa variable, el modelo ha empeorado su ajuste. Finalmente, de todas las estimaciones realizadas, hemos elegido el modelo que mejor se ajusta a los datos según el test de cociente de verosimilitud⁹, el índice Pseudo- R^2 de McFadden¹⁰, el criterio de información Akaike y el porcentaje de casos bien clasificados.

⁹El test de cociente de verosimilitud confirma la significación global de un modelo. Bajo la hipótesis nula de que los coeficientes de todas las variables explicativas consideradas en el modelo son iguales a cero, el estadístico $-2(L(0) - L(\hat{\beta}))$ se distribuye como una distribución chi-cuadrado con k grados de libertad, donde $L(0)$ es el valor máximo de la función de log-verosimilitud bajo la hipótesis nula y $L(\hat{\beta})$ el máximo de la función de log-verosimilitud no restringida.

¹⁰El Pseudo- R^2 de McFadden (McFadden, 1974) o índice ρ^2 es otra medida para comparar la bondad del ajuste los modelos de elección discreta y se usa de forma similar al R^2 del modelo de regresión. Este estadístico se define como

$$\rho^2 = 1 - \frac{L(\hat{\beta})}{L(0)}$$

Louviere et al. (2000) indican, basándose en las simulaciones realizadas por Domencich y McFadden (1975), que valores de ρ^2 entre 0,2 y 0,4 equivaldrían al rango 0.7 a 0.9 para

En la ecuación 2.7 se han incluido todas las variables explicativas consideradas en la Tabla 2.7, excepto SLAB, TRURAL y AREA, excluidas por la falta de variabilidad entre las observaciones asociadas con la alternativa 4. La Tabla 2.8 muestra las variables y los valores estimados de los parámetros para el modelo que da la probabilidad de elección de cada uno de los cuatro tipos diferentes de alojamientos disponibles en la zona Noroeste de la Región de Murcia, determinados con el análisis de clases latentes, según las características socio-económicas y las relacionadas con el viaje de los individuos.

Una observación inicial de la Tabla 2.8 revela que ninguna de las modalidades de la variable COMIDA es explicativa en la elección del alojamiento rural. Por lo tanto, que el individuo participe en actividades gastronómicas no influye en la probabilidad de elección de un tipo de alojamiento rural. Además, tampoco resulta determinante en la elección del tipo de alojamiento si el turista dispone de estudios superiores. El resto de variables, sin embargo, son significativas, en una de las modalidades y, por lo tanto, influyen en la probabilidad de elegir los tipos de alojamientos.

Por ejemplo, que el turista proceda de Murcia es significativo para los establecimientos de la Clase 3. El valor negativo del coeficiente para el tercer grupo quiere decir que el hecho de que un individuo resida en la provincia de Murcia y no en otra provincia, tiene mayor probabilidad de elegir un alojamiento de la Clase 1 que de la 3. Esto mismo sucede con la modalidad hombre de la variable “género” que también es significativa para los alojamientos de la Clase 3, indicando que la preferencia de elección son las casas de la Clase 1 frente a las de la Clase 3.

Con respecto a la variable EDAD, que el turista tenga una edad inferior o igual a 25 años es significativo para los establecimientos de la Clase 3. El valor negativo del coeficiente para esta clase quiere decir que el hecho de que un individuo tenga esa edad, a diferencia de los que son mayores de 40 años, disminuye la probabilidad de elección de los alojamientos de tamaño medio sin equipamientos adicionales (Clase 3) frente a los alojamientos de gran tamaño, con 3 o más habitaciones y que disponen de piscina (Clase 1). Esto mismo sucede con los individuos de edad comprendida entre 26 y 40 años, aunque su efecto es menor. Este tramo de edad también es significativo para los alojamientos incluidos en la Clase 4. De hecho, la probabilidad de elección de los alojamientos de un dormitorio en un molino (Clase 4) disminuye

una función lineal, por tanto, se consideran indicativos de un buen ajuste. Hensher et al. (2015) señalan que valores del pseudo- R^2 entre 0,3 y 0,4 equivalen a los valores del R^2 en el modelo de regresión lineal entre 0,6 y 0,8.

considerablemente respecto a los de gran tamaño (Clase 1). Por tanto, los individuos con edad comprendida entre 25 y 40 años respecto de los turistas con edad superior a 40 años muestran mayor preferencia por los alojamientos de la Clase 1 frente a los de las Clases 3 y 4.

Tabla 2.7: Variables y sus modalidades

Tipo	VARIABLES	Denominación	Modalidad de referencia	Modalidades restantes
Características Socio-económicas	GEN	Género del encuestado	Mujer	Hombre
	PROC	Provincia de origen	Resto de España	Murcia
	EDAD	Edad del turista	Más de 40 años	Hasta 25 años De 26 a 40 años
	ESTSUP	Educación superior	No	Sí
	SLAB	Situación laboral	Desempleado, tareas del hogar y jubilados	Ocupado Estudiante
	TIPFAM	Tipología de unidad familiar	Individuo independiente sin descendientes a su cargo	Soltero que vive con sus padres Individuo con descendientes a su cargo
	INGFAM	Ingresos familiares mensuales	Hasta 2405 euros	Más de 2405 euros
Condiciones de viaje	ALQ	Forma de alquiler del alojamiento	A través de NORATUR o de agencias	Directamente del propietario
	REFLUG	Referencia del lugar	Internet, guías turísticas u otros viajes	A través de conocidos
	PERS	Compañeros de viaje	Únicamente con la familia	Únicamente amigos Familia y amigos
	GASTO	Gastos diarios por persona y día (excluyendo alojamiento)	Hasta 60 euros	Más de 60 euros
	TRURAL	Experiencia previa en turismo rural	No	Sí
	DEPOR	Actividades deportivas	No	Sí
	AREA	Actividades asociadas con el entorno	No	Sí
	COMIDA	Actividades gastronómicas	No	Sí
	CULTURA	Actividades culturales	No	Sí

En referencia a la variable “tipología de la unidad familiar”, si el individuo es soltero y vive con sus padres será más propenso a elegir casas de

tamaño medio con o sin instalaciones (Clases 2 y 3, respectivamente), antes que casas de gran tamaño (Clase 1). Mientras que si los ingresos familiares mensuales son superiores a 2405 euros se incrementa la probabilidad de elegir un alojamiento perteneciente a la Clase 4.

En cuanto a la forma de alquiler del alojamiento, aquellos que lo han hecho a través de NORATUR o de otra agencia tienen mayor probabilidad de elegir las casas tradicionales de dos dormitorios sin equipaciones (Clase 3). Los alojamientos de las Clases 2 y 3 aumentan su probabilidad de elección respecto a los de la Clase 1, si el alojamiento ha sido recomendado por conocidos frente a haber obtenido información sobre él por internet, guías turísticas u otros viajes.

En relación a la variable “compañeros de viaje” se puede observar que presenta un coeficiente negativo para todas aquellas modalidades en las que la compañía son los amigos (viaja solo con ellos o con ellos y la familia). Así, viajar con amigos, en lugar de únicamente con la familia, aumenta la probabilidad de elegir un alojamiento de la Clase 1, que es el grupo al que pertenecen los establecimientos más grandes. Además, aquellos turistas que tengan pensado gastarse más de 60 euros diarios prefieren un alojamiento de la Clase 1 antes que de la Clase 3.

Por último, si un individuo participa en actividades deportivas, por ejemplo, senderismo, descenso del río Segura, montar a caballo o ciclismo, etc., entonces la probabilidad de elección de un alojamiento de la Clase 2 se ve afectada positivamente, ya que estos alojamientos disponen de equipamientos adicionales, entre los que están las pistas deportivas. Al igual que si el individuo participa en actividades culturales tiene un efecto positivo sobre la probabilidad de elegir un alojamiento de las Clases 3 o 4 antes que uno de la Clase 1.

Para concluir con el análisis y con el fin de testar si las variables SLAB, TRURAL y AREA son o no determinantes en la elección de alojamiento, se ha repetido la estimación del modelo logit, pero únicamente con las tres primeras clases detectadas con ACL¹¹. Es decir, se ha omitido la alternativa que incluye aquellos alojamientos con únicamente una habitación o que

¹¹Estas variables no se pudieron incluir en el modelo LACL con 4 alternativas de elección, porque el porcentaje de observaciones correspondientes a establecimientos de la alternativa 4 es bastante pequeño (de 2.9% para la clase 4 y del 2.6% en el grupo 4). Esto dificulta la estimación de algunas variables que describen a los individuos, al no existir variabilidad suficiente. La estimación del modelo LCA eliminando la cuarta alternativa de elección también se han incluido en el Anexo 4.8. Por último, en el anexo 4.8 se puede consultar un cuadro resumen de los cuatro modelos logit estimados, donde se especifica las variables

en origen eran molinos y ahora son establecimientos de alojamiento que se alquilan por habitaciones individuales con baño privado. En la Tabla 2.9 se muestran los resultados de la estimación del modelo LACL que toma como alternativa base la clase 1 de alojamientos y cuya utilidad determinística, esta vez sí, incluye todas las variables explicativas consideradas en Tabla 2.7,

$$\begin{aligned}
V_{ij} = & ACS_j + \beta_{1j}GEN_i + \beta_{2j}PROC_i + \beta_{3j}EDAD_{25i} \\
& + \beta_{4j}EDAD_{2640i} + \beta_{5j}ESTSUP_i + \beta_{6j}SLAB_{1i} \\
& + \beta_{7j}SLAB_{2i} + \beta_{8j}TIPFAM_{1i} + \beta_{9j}TIPFAM_{2i} \\
& + \beta_{10j}INGFAM_i + \beta_{11j}ALQ_i + \beta_{12j}REFLUG_i \\
& + \beta_{13j}PERS_{1i} + \beta_{14j}PERS_{2i} + \beta_{15j}GAST_{1i} \\
& + \beta_{16j}GAST_{2i} + \beta_{17j}TRURAL_i + \beta_{18j}DEPOR_i \\
& + \beta_{19j}AREA_i + \beta_{20j}COMIDA_i + \beta_{21j}CULTURA_i \quad j = 1, 2, 3 \quad \forall i
\end{aligned} \tag{2.8}$$

Al explorar la Tabla 2.9 se observa que las variables explicativas significativas para los alojamientos de las clases 2 y 3 en LACL con cuatro alternativas (véase Tabla 2.8), siguen siéndolo en el modelo LACL con 3 alternativas y, además, con el mismo signo. También se puede apreciar que la variable “ingresos familiares mensuales” sigue sin ser determinante en la elección de alojamiento de las Clases 1, 2 y 3. En cuanto a las tres nuevas variables introducidas en (2.8), la variable TRURAL no es determinante en la elección del alojamiento rural, mientras que el turista sea estudiante o tenga interés en probar la gastronomía de la zona es significativo (con coeficiente negativo) para los establecimientos de la Clase 2, disminuyendo la probabilidad de ser elegidos frente a un alojamiento de la Clase 1. Por último, que el individuo esté interesado en caminar o visitar lugares de interés por la zona aumenta la probabilidad de elegir los alojamientos de la Clase 2 frente a los de la Clase 1.

explicativas incluidas, cuales han sido significativas para alguno de los tipos de alojamientos y el signo de su coeficiente.

Tabla 2.8: Modelo estimado con las cuatro clases obtenidas con el análisis de clases latentes (modelo LACL)

Atributos	Denominación	Modalidades	Clase 1	Clase 2	Clase 3	Clase 4
ACS	Constante		0	0,6677	4,0320***	0,3299
GEN	Género del turista	Hombre	0	-0,4087	-0,9960***	-0,4541
PROC	Provincia de procedencia	Murcia	0	-0,2709	-0,7409*	-1,6555
EDAD	Edad del turista	Hasta 25 años De 26 a 40 años	0	0,4201 -0,2958	-2,0040*** -1,5971***	1,3137 -2,2740*
TIPFAM	Tipología de unidad	Soltero que vive con sus padres	0	1,1498**	1,7264***	0,0146
INGFAM	Ingresos familiares	Más de 2405 euros	0	-0,0666	-0,1200	2,3712**
ALQ	Forma de alquiler del alojamiento	Directamente al propietario	0	0,2860	-1,2502***	-1,2478
REFLUG	Referencia del lugar	A través de conocidos	0	0,6635**	0,9640***	0,4570
PERS	Compañeros de viaje	Únicamente amigos Familia y amigos	0	-2,7210*** -1,3443***	-2,5338*** -1,1754**	-5,1193*** -3,0263*
GASTO	Gasto diario por persona y día	Más de 60 euros	0	0,3801	-0,9144**	1,1204
DEPOR	Actividades deportivas	Sí	0	0,6004*	-0,3175	-1,1935
CULTURA	Actividades culturales	Sí	0	0,1869	0,8815**	2,5588**
			307			
<i>Observaciones</i>						
Número de parámetros			42			
Función Log-verosimilitud			-293,12			
Pseudo-R ² de McFadden			0,2016			
AIC			670,23			
BIC			826,76			
% casos bien clasificados			54,07%			

Significatividad: ****: $p < 0,001$; ***: $p < 0,01$; **: $p < 0,05$; *: $p < 0,1$

Tabla 2.9: Modelo LACL con las Clases 1, 2 y 3 obtenidas con el análisis de clases latentes (modelo LACL)

Atributos	Denominación	Modalidades	Clase 1	Clase 2	Clase 3
ACS	Constante		0	0,1610	3,8706***
GEN	Género del turista	Hombre	0	-0,3658	-1,0041***
PROC	Provincia de origen	Murcia	0	-0,2533	-0,7793**
EDAD	Edad del turista	Hasta 25 años De 26 a 40 años	0	1,0199	-1,5513*
SLAB	Situación laboral	Estudiante	0	-1,4323**	-0,8918
TIPPAM	Tipología familiar	Soltero que vive con sus padres	0	1,2913**	1,7457***
ALQ	Forma de alquiler del alojamiento	Directamente propietario	0	0,3235	-1,2497***
REFLUG	Referencia del lugar	A través de conocidos	0	0,7018**	1,0094***
PERS	Compañeros de viaje	Únicamente amigos Familia y Amigos	0	-2,7707***	-2,4940***
GASTO	Gastos diarios por persona y día	Más de 60 euros	0	0,4319	-0,8543**
DEPOR	Actividades deportivas	Sí	0	0,5654	-0,3369
AREA	Actividades asociadas con el entorno	Sí	0	0,6646*	0,2525
COMIDA	Actividades gastronómicas	Sí	0	-0,6511*	-0,5422
CULTURA	Actividades culturales	Sí	0	0,4260	1,0712***
			298		
			<i>Observaciones</i>		
			32	<i>Número de parámetros</i>	
			-265,72	<i>Función Log-verosimilitud</i>	
			0,1861	<i>Pseudo-R² de McFadden</i>	
			595,44	<i>AIC</i>	
			713,75	<i>BIC</i>	
			54,36 %	<i>% casos bien clasificados</i>	

Significatividad: ***: $p < 0,001$; **: $p < 0,01$; *: $p < 0,05$; +: $p < 0,1$

2.6. Conclusiones

En este capítulo se muestra una metodología para determinar el perfil de un turista para cada tipo de alojamiento ofertado en un destino turístico rural. Una herramienta clave en este estudio ha sido la clasificación de establecimientos de alojamiento rural en grupos homogéneos. Para ello se han utilizado dos métodos de agrupación, el primero combina un análisis de correspondencias múltiple y un análisis cluster y el segundo es un análisis de clases latentes. A pesar de las dificultades que implican las técnicas basadas en variables cualitativas, el estudio ha demostrado su utilidad en la obtención de diferentes tipologías de alojamiento que existen en un destino de turismo rural.

Usando el modelo logit multinomial se han estimado las probabilidades de elección para cada uno de los cuatro tipos de alojamiento en función de las características de los turistas que han decidido pasar sus vacaciones en ese destino. Los alojamientos del Grupo 1, que son establecimientos principalmente grandes, ubicados en el campo o en la montaña y que poseen piscina, son preferidos por aquellos turistas que viven en pareja sin hijos, que suelen viajar con amigos, que prefieren alquilar directamente al propietario y que buscan información del lugar a través de internet y guías turísticas. Los alojamientos pertenecientes al Grupo 2, que son los de tamaño medio, situados en zonas agrícolas y que poseen equipamientos adicionales, como instalaciones deportivas, son los preferidos por aquellos turistas que realizan un gasto medio-alto en el destino y buscan realizar actividades deportivas o de senderismo o visitar lugares de interés cercanos a la vivienda. Los turistas con más de 40 años que prefieren realizar actividades culturales optan por los alojamientos del Grupo 3, que son aquellos con una arquitectura típica de la zona y situados en aldeas o en los alrededores de un pueblo. Por último, los establecimientos que no se alquilan al completo, sino por habitaciones individuales con baño privado, son el objetivo de los turistas mayores de 40 años, con ingresos altos y que se decantan por realizar actividades culturales.

Aunque hemos hecho un estudio aplicado a una región en un tiempo concreto, los resultados del estudio ponen de manifiesto que existen diferencias significativas entre los turistas que eligen cada tipo de alojamiento en un destino. Por tanto, a modo individual, cada alojamiento puede hacer uso de los perfiles turísticos aquí obtenidos como base para realizar políticas de inversión y marketing. Se trataría de tener en cuenta aquellas variables que proporcionan una mayor probabilidad de elegir ese tipo de alojamiento turístico para saber a qué turistas dirigir sus promociones y cómo satisfacer de una manera óptima sus preferencias de viaje y actividades.

A nivel de destino, este estudio muestra que los turistas rurales, aunque sean homogéneos en cuanto a su motivaciones para ir a un alojamiento rural, muestran diferencias en cuanto a sus características socio-económicas, sus comportamientos de viaje y sus preferencias por las actividades. En consecuencia, en la planificación de la oferta de un destino rural es necesario tener en cuenta los perfiles de los turistas que lo visitan, así como sus preferencias. Especialmente, en aquellos destinos con objetivos de inversión y promoción.

Capítulo 3

Preferencias heterógenas de los turistas rurales ante la elección de alojamiento

3.1. Introducción

Conocer los atributos o características de los alojamientos más valorados por los turistas, constituye una importante herramienta para comprender las decisiones de estos turistas ante la elección de alojamiento. En el caso de los hoteles, existen diferentes estudios que han analizado los principales factores que determinan su elección. Estos estudios muestran que ciertas características como la ubicación, el precio, las instalaciones, el tamaño de las habitaciones, la estancia, la limpieza, el silencio o la disponibilidad de aire acondicionado tienen mucha influencia en la elección de un hotel determinado (Lockyer, 2005; Stringam et al., 2010; Merlo y Joao, 2011; Chen et al., 2017; Masiero et al., 2015).

La elección de un alojamiento rural tradicionalmente se ha asociado a su ubicación en un entorno natural y a sus características intrínsecamente rurales y, no tanto, a los atributos físicos que definen e identifican el alojamiento (Uysal y Hagan, 1993 ; Park y Yoon, 2009; Yoo et al., 2018). Sin embargo, los alojamientos rurales, al igual que los hoteles, son también muy variados en cuanto a su tamaño, al tipo de edificación, a su localización, a los servicios y actividades que ofertan o a la calidad del mobiliario de que disponen. En este capítulo mostramos que estas características físicas de los alojamientos rurales, así como las actividades y servicios que ofertan, son determinantes para su elección.

Ahora bien, la percepción y evaluación de las características de los alojamientos no tiene porque ser igual para todos los individuos. Ésta puede variar en función del turista que esté haciendo la elección. Por ejemplo, puede haber turistas que prefieran una casa rústica, típica de la zona, frente a otros a los que les sea indiferente el tipo de construcción. También pueden existir turistas que busquen una casa cercana a un pueblo, mientras que otros prefieran una casa alejada de cualquier zona habitada. En este capítulo tratamos de modelizar estos gustos desiguales de los turistas en la elección de alojamiento rural, relajando el supuesto de preferencias homogéneas característico del modelo logit multinomial.

Los primeros Modelos de Elección Discreta se construyeron sobre la base de que los individuos eran homogéneos en cuanto a sus gustos por las características de las alternativas de elección. Sin embargo, dada la amplia variedad de patrones de comportamiento de los individuos ante situaciones de elección (Ortúzar y Willumsen, 2001), se han desarrollado diferentes propuestas con la intención de modelizar esa heterogeneidad de preferencias de los individuos (Boxall et al., 1996; Hensher et al., 2015; van Oel y van Den Berkhof, 2013). En este capítulo valoramos la existencia de preferencias heterogéneas de los turistas rurales ante la elección de alojamiento, suponiendo la existencia de dos tipos de variaciones en las preferencias: una sistemática y otra aleatoria. La primera especificación explica la heterogeneidad de las preferencias en función de las características socio-demográficas observables de los turistas. Este enfoque supone, por ejemplo, que aquellos turistas que viajan con niños se comportan de una forma similar entre ellos y diferenciada de aquellos otros grupos que viajan sin niños. La especificación de preferencias aleatorias considera que la heterogeneidad se debe a variaciones del gusto que no son observadas. Ambos modelos se estiman con los datos de la encuesta realizada en la Comarca Noroeste de la Región de Murcia durante varios fines de semana de otoño de 2003. Los resultados de ambas estimaciones se comparan con la estimación de un logit multinomial homogéneo.

Una vez estimados los modelos logit, se puede realizar de una forma muy sencilla una estimación de la disposición a pagar por los atributos de una cierta alternativa. Estas estimaciones permiten hacerse una idea de la cantidad media de dinero que los turistas estarían dispuestos a pagar para mantener sus niveles de utilidad ante un cambio en uno de los atributos de una casa. De esta forma, se pueden tener valoraciones del efecto que un atributo tendría sobre la elección de una casa, por ejemplo, como se valoraría el que la casa elegida fuese una vivienda de nueva construcción frente a que sea de construcción tradicional, o que una casa sin la certificación “Q” de calidad, la

adquiriese. Además, veremos que estos efectos pueden variar para diferentes segmentos de turistas.

El enfoque conceptual y empírico, que se ofrece en este capítulo, es una vía para valorar la idoneidad de los modelos de elección discreta para la elección de alojamiento rural. En este capítulo también se propone el uso de modelos que proporcionan ajustes mejores que el logit y que muestran gran heterogeneidad en las preferencias de los turistas. Estos resultados podrían ayudar a los promotores y propietarios de alojamientos rurales para mejorar la oferta del destino cuando planifican campañas promocionales y de marketing o para tomar decisiones de inversión efectivas (Dávila et al., 2002; Sharpley y Sharpley, 1997).

Este capítulo se estructura en 5 epígrafes. En el segundo epígrafe se desarrollan las dos especificaciones logit que nos permiten incorporar variaciones en el gusto de los individuos (heterogeneidad del gusto). La primera incorpora la heterogeneidad de las preferencias sistemática y la segunda modeliza la heterogeneidad aleatoria usando el modelo logit mixto. En el tercer epígrafe se explica que son y cómo se calculan las disposiciones a pagar. En el cuarto se detallan los datos que se usan, se especifican los modelos logit a estimar y se presentan los resultados obtenidos con la estimación de estos modelos y de las disposiciones a pagar. Por último, en el epígrafe quinto, se exponen las conclusiones derivadas de la aplicación de los modelos logit que incorporan la heterogeneidad del gusto.

3.2. Modelos de elección con preferencias heterogéneas

Según Palma et al. (2018) el uso de los modelos de elección discreta tiene dos importantes ventajas en el estudio de las preferencias de los individuos. En primer lugar, se trata de un método indirecto para evaluar el valor de los atributos de una alternativa. Esto es, los individuos no tienen que hacer una valoración directa de cada atributo o característica. En segundo lugar, estos modelos no requieren que los turistas hagan nada diferente de lo que normalmente hacen para reservar un alojamiento, es decir, simplemente tienen que elegir un alojamiento entre los diferentes existentes.

En los modelos de elección discreta, la heterogeneidad de preferencias se ha introducido utilizando diferentes mecanismos. En este estudio analizamos

dos posibles enfoques que utilizan especificaciones logit. El primero supone que la heterogeneidad procede o se puede determinar a través de ciertas características observables de los individuos (*heterogeneidad condicional*), mientras que el segundo se basa en que la heterogeneidad de gustos tiene un carácter aleatorio y difícilmente observable (*heterogeneidad aleatoria*). A continuación, desarrollamos ambas opciones.

3.2.1. Heterogeneidad condicional

En el epígrafe 1.3 del capítulo 1 desarrollamos los modelos de elección discreta. Estos modelos se basan en la teoría de Maximización de la Utilidad Aleatoria, que supone que cada individuo asigna una utilidad a cada una de las alternativas disponibles en el contexto de elección y elige aquella alternativa que más utilidad le reporta. Esta utilidad, U_{ij} , no conocida con certeza por los analistas, se descompone, según la ecuación (1.9), en una parte observable, V_{ij} , y otra aleatoria, ε_{ij} , que recoge los factores desconocidos de los individuos o los errores de medida. Además, si los ε_{ij} , son independientes e idénticamente distribuidos entre sí, con distribución del valor extremo del tipo I o Gumbel se obtiene el modelo Logit Multinomial, que tiene una forma analítica sencilla para expresar la probabilidad de elección de una alternativa.

El modelo Logit Multinomial, que también estudiamos en el capítulo 1, considera que la utilidad determinística, V_{ij} , depende linealmente de los atributos de la alternativa observada, x_{ijk} , según la expresión

$$V_{ij} = \sum_k \beta_k x_{ijk} \quad (3.1)$$

donde los parámetros a estimar, β_k , pueden o no depender de la alternativa¹, pero no varían con los individuos. Así, esta suposición de linealidad implica que todos los individuales valoran las características que definen las alternativas de igual forma, esto es, todos los individuos tienen un comportamiento homogéneo ante las características de cada alternativa.

Ortúzar y Willumsen (2011) proponen introducir la heterogeneidad de preferencias explicándola a través de alguna variable adicional que sea observable, como pueden ser las características socio-económicas de los turistas. Su propuesta, conocida como “Heterogeneidad condicional” o “Preferencias

¹En el supuesto de que los β_k varíen por alternativa habría que incorporar en la notación el subíndice j , β_{kj} .

heterogéneas sistemáticas”, consiste en agregar nuevos términos a la utilidad, que no son más que interacciones entre los atributos de los productos y las características de los consumidores. Bajo este supuesto, la utilidad determinística tendría la siguiente expresión

$$V_{ij} = \sum_k \beta_k x_{ijk} + \sum_r \sum_k \beta_{rk} x_{ijk} z_{ir} \quad (3.2)$$

donde z_{ir} indica el valor de la característica r del individuo i y β_{rk} el parámetro a estimar para la característica r del individuo y el atributo k de la alternativa.

Considerar la expresión (3.2) para la utilidad determinística implica suponer que pueden existir coeficientes de los atributos de las alternativas distintos para diferentes tipos o segmentos de individuos. Por tanto, se puede establecer heterogeneidad en los gustos en función de las características de los individuos. Además, este método, que incorpora la heterogeneidad del gusto de una forma bastante sencilla, tiene también la ventaja de que puede mantener la forma analítica de la probabilidad del modelo logit multinomial. Así, suponiendo que los ε_{ij} , son independientes e idénticamente distribuidos entre sí con distribución del valor extremo del tipo I o Gumbel, la probabilidad de que un individuo i elija la alternativa j , P_{ij} , vendrá dada por la expresión

$$P_{ij} = \frac{\exp(\sum_k \beta_k x_{ijk} + \sum_r \sum_k \beta_{rk} x_{ijk} z_{ir})}{\sum_{l \in C} \exp(\sum_k \beta_k x_{ilk} + \sum_r \sum_k \beta_{rk} x_{ilk} z_{ir})} \quad \forall j \in C_i \quad (3.3)$$

En consecuencia, el método de preferencias heterogéneas sistemáticas es una forma simple y fácil de relajar el supuesto de preferencias iguales para todos los individuos. La principal limitación de este enfoque es que las características de los consumidores deben ser observables, por lo que a menudo se utilizan los factores socio-económicos de los individuos (Palma et al., 2018).

3.2.2. Heterogeneidad aleatoria

La heterogeneidad aleatoria asume que las variaciones de gusto de los individuos son aleatorias y no observables. Un modelo que permite incorporar este tipo de variaciones en el gusto es el modelo logit mixto. Además, este modelo también admite cualquier patrón de sustitución y correlación en los factores inobservados sobre el tiempo, lo que es especialmente útil cuando se trabaja con datos de preferencias declaradas. Otra ventaja es su flexibilidad,

ya que se puede aproximar a cualquiera de los modelos de utilidad aleatoria (McFadden y Train, 2000). La principal dificultad de este modelo es que consume mucho tiempo de computación. Aunque a partir del año 2000, gracias a las mejoras en la velocidad de los ordenadores y en los algoritmos de simulación, se ha utilizado en diferentes estudios (Alfnes, 2004, Brownstone y Train, 1999; Brownstone y Train, 1999; Brownstone et al., 2000 y Espino et al., 2008, entre otros).

El modelo logit mixto puede derivarse bajo una variedad de especificaciones de comportamiento diferentes (Train, 2003). La más utilizada, en las aplicaciones que estudian las variaciones en el gusto de los individuos, es la versión de coeficientes aleatorios, de ahí que este modelo también sea conocido como *logit de parámetros o coeficientes aleatorios*. En esta versión, al igual que sucede con el logit multinomial, la utilidad asociada a cada una de las alternativas se considera que es una especificación lineal respecto a los parámetros. Pero, a diferencia de lo que ocurre en el logit, estos parámetros varían entre los individuos según una distribución. Esto es, en el logit mixto la utilidad de la alternativa j para un individuo i viene descrita por

$$U_{ij} = \beta'_i X_{ij} + \varepsilon_{ij} \quad (3.4)$$

donde $X_{ij} = (x_{ij1}, x_{ij2}, \dots, x_{ijK})$ es el vector de variables observadas asociadas a la alternativa j y al individuo i y $\beta_i = (\beta_{i1}, \beta_{i2}, \dots, \beta_{iK})$ es un vector de coeficientes aleatorios que representa los gustos de los individuos con una función de densidad $f(\beta_i/\theta)$, caracterizada por un conjunto de parámetros θ , que generalmente son la media y la matriz de varianzas–covarianzas.

El logit mixto también supone que los ε_{ij} son independientes e idénticamente distribuidos entre sí con una distribución del valor extremo de tipo I o Gumbel. Sin embargo, debido a que los β_i no están dados, la probabilidad de que un individuo elija una alternativa no tendrá una forma cerrada como en el caso del logit multinomial. En este caso, la probabilidad se obtiene como el valor esperado de las probabilidades logit en la población. Suponiendo que los coeficientes β_i sean aleatorios con una función de densidad² $f(\beta_i/\theta)$, la probabilidad, P_{ij} , de que un individuo i elija la alternativa j , viene dada como la integral de la probabilidad del logit multinomial sobre todos los valores de estos parámetros:

$$P_{ij} = \int_{\beta_i} L_{ij}(\beta_i) f(\beta_i/\theta) d\beta_i \quad (3.5)$$

²Las distribuciones normal, log-normal, uniforme y triangular son las más utilizadas para $f(\beta_i/\theta)$ en el modelo logit de coeficientes aleatorios.

donde

$$L_{ij}(\beta_i) = \frac{e^{\beta_i' X_{ij}}}{\sum_{k \in C} e^{\beta_i' X_{ik}}} \quad \forall j \in C. \quad (3.6)$$

Las probabilidades P_{ij} definidas por la ecuación (3.6) no tienen una expresión cerrada, por lo que se aproximan mediante un procedimiento de simulación para cualquier valor dado de θ , que son los parámetros que caracterizan la distribución de los β_i . El proceso de simulación consiste en dado un valor de θ , se extrae un valor para el vector β_i^r de la función de densidad $f(\beta_i/\theta)$, indicando el superíndice r el número de extracción. Usando este valor de β_i^r , se calcula $L_{ij}(\beta_i^r)$ mediante la fórmula (3.6). Este proceso se repite para un número R de extracciones. La probabilidad de elección aproximada, \tilde{P}_{ij} , se toma como la media de los valores resultantes $L_{ij}(\beta_i^r)$, esto es,

$$\tilde{P}_{ij} = \frac{1}{R} \sum_{r=1}^R P_{ij}(\beta_i^r) \quad (3.7)$$

Como señala Train (2003), el motivo de considerar probabilidad de elección simulada, \tilde{P}_{ij} , como el valor medio descrito en (3.7) se debe a que:

- \tilde{P}_{ij} es un estimador insesgado de P_{ij} .
- La varianza de \tilde{P}_{ij} decrece cuando se incrementa R .
- El valor de \tilde{P}_{ij} es estrictamente positivo, de manera que el $\ln(\tilde{P}_{ij})$ está definido, lo que es útil cuando se usa \tilde{P}_{ij} para aproximar la función log-verosimilitud.
- \tilde{P}_{ij} es doblemente diferenciable respecto los parámetros θ y las variables X , lo que facilita la búsqueda numérica del máximo de la función de verosimilitud y el cálculo de las elasticidades.
- La suma de las probabilidades simuladas, \tilde{P}_{ij} , sobre las alternativas es uno, lo cual es útil cuando se quiere predecir.

Actualmente el método estimación del modelo logit mixto utiliza el enfoque de maximización de la función de verosimilitud simulada utilizando extracciones “inteligentes” de la distribución, como son las extracciones Halton. Este tipo de extracciones se ha demostrado que son más eficaces que

las extracciones aleatorias³, ya que reducen tiempo de ejecución y, además, pueden reducir el error de simulación que está asociado con un número de extracciones dado (Bhat, 1999; Train, 2003).

3.3. Disposiciones a pagar

La disposición a pagar por un atributo es una estimación de la media de la cantidad de dinero que los turistas individuales están dispuestos a pagar para retener sus niveles originales de utilidad ante un cambio en uno de los atributos de una alternativa (Apostolakis y Jaffry, 2005a, 2005b; Brey et al., 2007; Espino et al., 2008; Huybers, 2003). Proporcionar un valor monetario a los cambios en la utilidad que obtiene un individuo ante el cambio de uno de los atributos o características de la alternativa que ha elegido, es de vital importancia para la formulación de políticas y para la economía del bienestar (Apostolakis y Jaffry, 2005a). Una forma de obtenerlo es equipararlo al aumento en la utilidad que el consumidor individual obtendría como resultado de este cambio con lo que habría ganado en términos de una variación “equivalente” en sus ingresos.

La disposición a pagar por un cambio en un cierto atributo es lo que se denomina, normalmente, el valor subjetivo del atributo o característica (Espino et al., 2004). Según la teoría microeconómica, este valor se puede computar como la tasa marginal de sustitución entre el ingreso y la cantidad del atributo suponiendo niveles de utilidad constantes. Esto es, como el cociente negativo entre la utilidad marginal de dicha característica y la utilidad marginal de la renta. Como bien demuestran Espino et al. (2008), la utilidad marginal de la renta es igual a menos la utilidad marginal del coste de la alternativa. Por tanto, cuando las utilidades tienen una expresión lineal para las variables que definen los atributos, como es en el caso de los modelos logit empleados en este estudio, esta medida se puede calcular como el cociente entre el coeficiente del atributo y el del coste de la alternativa, esto es,

$$\text{Disposición a pagar por el atributo } k = \frac{\partial V / \partial x_k}{\partial V / \partial P} = \frac{\beta_k}{\beta_P} \quad (3.8)$$

donde β_k es el coeficiente del atributo k y β_P es el coeficiente del precio.

³En la estimación del modelo logit mixto, tradicionalmente, se han utilizado extracciones aleatorias de la distribución, lo que requería que el número de extracciones, R , fuese bastante grande si se quería alcanzar niveles de precisión aceptables. Con las extracciones “inteligentes” este número se ha reducido drásticamente (Hensher et al., 2005).

La expresión (3.8) es válida para los atributos de carácter cuantitativo. En el caso de que el atributo sea cualitativo, la disposición a pagar por ese atributo será medida, tal y como lo hacen Espino et al. (2008), por

$$\text{Disposición a pagar por el atributo } k = -\frac{V_1 - V_0}{\partial V / \partial P} \quad (3.9)$$

donde V_1 es la utilidad cuando el nivel del atributo k es 1 y V_0 es la utilidad cuando el nivel del atributo k es 0.

El método descrito por las ecuaciones (3.8) y (3.9) es únicamente correcto si los parámetros son fijos. En el caso de que los parámetros sean aleatorios, se sugieren otros métodos, ya que la relación de dos parámetros aleatorios no es fácil de analizar. Entre los métodos propuestos destaca el de obtener el espacio de preferencia de las disposiciones a pagar realizando simulaciones o el de fijar el coeficiente de coste (Hensher et al., 2005).

3.4. Investigación empírica

3.4.1. Datos y variables

Los datos usados para estimar los modelos logit homogéneo, logit con heterogeneidad sistemática y logit mixto propuestos en este epígrafe son los obtenidos con el experimento de preferencias declaradas que se llevó a cabo en la Comarca Noroeste de la Región de Murcia en 2003 y que se ha descrito en el epígrafe 1.5 del capítulo 1 de esta memoria. También hemos empleado la información disponible de los individuos en cuanto a sus características socio-económicas y de comportamiento de viaje. En la Tabla 3.1 se recogen todas las variables utilizadas, tanto las que definen los atributos de las casas como las que hacen referencia a las características socio-económicas y de comportamiento de viaje de los turistas.

3.4.2. Modelos estimados

Con objeto de determinar los atributos que son determinantes para la elección de los alojamientos rurales y estudiar si existe heterogeneidad en cuanto a las preferencias de los turistas ante estos atributos, estimamos dos especificaciones para el logit multinomial y dos para el logit mixto. Las diferencias entre los dos modelos logit y los dos logit mixtos se encuentran en la especificación de la utilidad determinística. En los cuatro modelos se usan las

Tabla 3.1: Variables del modelo

Denominación	Tipos
Precio de la casa	Variable continua
Tamaño de la casa	Variable continua
Construcción original	1 si el alojamiento es una casa típica de la zona
De nueva construcción	1 si el alojamiento es una casa de nueva construcción en la zona
Construcción de madera	1 si el alojamiento es una cabaña de madera
Casa Pueblo	1 si el alojamiento está situado a las afueras de un pueblo
Casa Huerta	1 si el alojamiento está situado en zona de huerta
Casa Campo	1 si el alojamiento está situado en el campo o la montaña
Minigranja	1 si el alojamiento tiene minigranja
Alquiler de caballos	1 si el alojamiento ofrece el servicio de alquiler de caballos
Casa alquilada completa	1 si el alojamiento se alquila al completo
Baño compartido	1 si las habitaciones no tienen su propio cuarto de baño
Calidad	1 si el alojamiento posee la “Q” de calidad
Reserva por internet	1 si el alojamiento se puede reserva a través de internet
<i>Características socioeconómicas</i>	
Tamaño del grupo de viaje	1 si el individuo viaja con 6 o más personas
Viaja con niños	1 si el individuo viaja con niños
Frecuencia de viaje a las zonas rurales	1 si el individuo hace más de dos veces al año
Título universitario	1 si el individuo tiene una diplomatura, licenciatura o similar

variables que definen los atributos de las casas (véase Tabla 3.1), pero únicamente en dos de ellos se consideran preferencias sistemáticas, mediante la interacción de los atributos de las casas con las características socio-económicas de los turistas, también incluidas en la Tabla 3.1.

El primer modelo es un logit multinomial con preferencias homogéneas entre los turistas. En este modelo, que denominamos *logit homogéneo*, la utilidad determinística es una función únicamente de las características que definen los alojamientos rurales introducidas en forma lineal, esto es,

$$\begin{aligned}
V_{ij} = & \beta_1 \text{Precio}_j + \beta_2 \text{Tamaño}_j + \beta_3 \text{Original}_j + \beta_4 \text{Nueva}_j \\
& + \beta_5 \text{Madera}_j + \beta_6 \text{Pueblo}_j + \beta_7 \text{Huerta}_j + \beta_8 \text{Campo}_j \\
& + \beta_9 \text{Minigranja}_j + \beta_{10} \text{Caballos}_j + \beta_{11} \text{Alquiler}_j + \beta_{12} \text{Baño}_j \\
& + \beta_{13} \text{Calidad}_j + \beta_{14} \text{Internet}_j \quad j = 1, 2, 3
\end{aligned} \tag{3.10}$$

El segundo modelo también es un logit multinomial pero con heterogeneidad sistemática entre los individuos. Para este modelo, que denominamos *logit heterogéneo*, hemos considerado que la anterior utilidad determinística se expande para incluir algunas interacciones entre las características socio-económicas y de comportamiento de viaje de los turistas con algunos niveles de los atributos de las casas rurales. Para este modelo la utilidad determinística tiene la siguiente expresión,

$$\begin{aligned}
V_{ij} = & \beta_1 \text{Precio}_j + \beta_2 \text{Tamaño}_j + \beta_3 \text{Original}_j + \beta_4 \text{Nueva}_j + \beta_5 \text{Madera}_j \\
& + \beta_6 \text{Pueblo}_j + \beta_7 \text{Huerta}_j + \beta_8 \text{Campo}_j + \beta_9 \text{Minigranja}_j \\
& + \beta_{10} \text{Caballos}_j + \beta_{11} \text{Alquiler}_j + \beta_{12} \text{Baño}_j + \beta_{13} \text{Calidad}_j \\
& + \beta_{14} \text{Internet}_j + \beta_{15} \text{Tamaño}_j \times \text{Tamaño del grupo de viaje}_i \\
& + \beta_{16} \text{Minigranja}_j \times \text{Viaja con niños}_i + \beta_{17} \text{Tamaño}_j \times \text{Título universitario}_i \\
& + \beta_{18} \text{Calidad}_j \times \text{Título universitario}_i + \beta_{19} \text{Completa}_j \times \text{Frecuencia viajes}_i \\
& + \beta_{20} \text{Precio}_j \times \text{Frecuencia viajes}_i \quad j = 1, 2, 3
\end{aligned} \tag{3.11}$$

En el tercer modelo incluimos la posibilidad de preferencias heterogéneas aleatorias en los gustos individuales al considerar un modelo logit mixto. En este modelo, denominado *logit mixto 1*, suponemos que la utilidad determinística viene dada por la ecuación (3.10), pero donde algunos de los parámetros, β_j , son considerados variables aleatorias que siguen una distribución normal, con media y varianza desconocidas.

El cuarto modelo también es un logit mixto que, además de considerar heterogeneidad con preferencias aleatorias, incorpora también heterogeneidad sistemática. En este modelo, denominado *logit mixto 2*, a parte de tener algunos de los parámetros, β_j , en la función de utilidad que son variables aleatorias, se han introducido las mismas interacciones, entre las características socio-económicas y de comportamiento de viaje de los encuestados con los atributos de las casas rurales, que se incluyeron en la expresión (3.11).

Para construir el modelo definido en cada especificación se ha aplicado un procedimiento de selección secuencial clásico de variables predictoras. Esto implica estimar modelos logit multinomiales o mixtos alternativos, según el caso, y elegir el mejor según el test de la razón de verosimilitud, el índice pseudo- R^2 de McFadden y el criterio de información de Akaike. Todas las especificaciones del logit multinomial se han estimado utilizando el software econométrico LIMDEP 8.0 (Greene, 2002).

3.4.3. Resultados de las estimaciones

Los resultados de la estimación de los cuatro modelos se presentan en la Tabla 3.2. La especificación del logit homogéneo (primera columna), que supone coeficientes constantes a través de los individuos, tiene un valor de 0,13 del índice pseudo- R^2 de McFadden ajustado⁴, que podemos considerar aceptable⁵. Como era de esperar, el coeficiente de la variable “Precio” tiene signo negativo. Por tanto, a mayor precio por la casa, menor probabilidad de ser elegida, *ceteris paribus*. Por el contrario, la variable “Tamaño” tiene un coeficiente positivo, indicando que cuanto mayor sea el tamaño de la casa, más probabilidad tiene de ser elegida, *ceteris paribus*. En cuanto al tipo de construcción, las casas originales son positivamente percibidas por los turistas en la zona Noroeste de la Región de Murcia, mientras que las casas construidas con el propósito expreso de alquilarlas tienen una menor probabilidad de ser elegidas. En cuanto a la localización de la casa, los turistas valoran positivamente el hecho de que la casa rural esté en el campo o la montaña y negativamente su localización a las afueras de un núcleo urbano. La probabilidad de alojarse en una casa rural en la Comarca del Noroeste de la Región de Murcia también está positivamente afectada si la casa tiene la certificación de calidad “Q” y ofrece la posibilidad de alquilar caballos.

⁴El Pseudo- R^2 de McFadden (McFadden, 1974) tiende a aumentar cuando se añade una variable explicativa, aunque ésta no tenga relevancia estadística, tal y como le sucede al R^2 del modelo de regresión. Para corregir este problema se introduce en la expresión del pseudo- R^2 el número de variables explicativas, k , esto es,

$$\rho_C^2 = 1 - \frac{L(\hat{\beta}) - k}{L(0)}$$

dando lugar al índice denominado pseudo- R^2 de McFadden ajustado, ρ_C^2 .

⁵Valores del pseudo- R^2 de McFadden entre 0,2–0,4 indican un excelente ajuste del modelo (McFadden, 1974).

Tabla 3.2: Resultados de las estimaciones de los modelos logit

Variable	Logit homogéneo		Logit heterogéneo		Logit mixto 1		Logit mixto 2	
	Coefficiente	t-test	Coefficiente	t-test	Coefficiente	t-test	Coefficiente	t-test
Precio de la casa	-0,3382***	-4,7070	-0,2541**	-2,9860	-0,8911**	-3,2990	-0,6226**	-2,6590
Tamaño de la casa	0,7223***	8,7090	0,7337***	8,6790	1,7082***	4,4250	1,5453***	4,7300
Contrucción original	0,2714***	3,6840	0,2717***	3,6230	1,0811**	3,3920	0,9943***	3,6320
De nueva construcción	-0,5783***	-7,1810	-0,5948***	-7,2900	-2,0254***	-3,7770	-1,8483***	-4,0110
media					3,4126**	3,3470	3,2299***	3,5900
St. D.					0,8408**	2,7920	0,7887**	2,9180
Construcción de madera	0,1852*	2,4290	0,2025**	2,6190	-0,8697***	-3,8570	-0,7754***	-4,1240
Casa pueblo	-0,4261***	-5,3250	-0,4407***	-5,4300	-0,0719	-0,3040	-0,1291	-0,5870
Casa huerta	0,1118	1,4600	0,1073	1,3770	2,2709**	2,7170	2,0361**	2,7720
media					0,8301**	3,2850	0,7672***	3,5290
St.D.					0,0885	0,6880	0,0994	0,9230
Casa campo	0,3097***	4,1430	0,3051***	4,0070	1,2973**	2,6210	0,2210+	1,7040
Minigranja	0,0559	1,2880	0,0845+	1,8740	0,2222	1,5910	-0,5025+	-1,8910
media					-0,0697	-0,5460	1,5445***	3,4890
St. D.					1,2871**	2,6040	0,4495**	3,0650
Alquiler de caballos	0,1153**	2,6260	0,1313**	2,9370	0,5005**	3,0610	0,0071	0,0530
Casa alquilada completa	-0,0823+	-1,8760	-0,0823**	-2,8140	0,0020	0,0130	2,1602***	3,8620
media					2,4451***	3,8580	0,2617***	3,7840
St. D.							0,3100**	2,6290
Calidad	0,1718***	3,9140	0,1656***	3,7110			0,4097*	2,3140
Baño compartido							-0,2527*	-1,9970
Tamaño de la casa × tamaño del grupo de viaje							-0,4672+	-1,7720
Minigranja × viajar con niños							0,1939+	1,7540
Construcción original × título universitario								
Calidad × título universitario								
Casa alquilada completa × frecuencia de viaje								
Precio de la casa × frecuencia de viaje								
<i>Observaciones</i>	1228		1228		1228		1228	
<i>Individuos</i>	307		307		307		307	
<i>Número de parámetros</i>	12		18		18		23	
<i>Función Log-verosimilitud</i>	-1158,98		-1130,21		-1110,61		-1088,12	
<i>Pseudo-R² de McFadden</i>	0,1409		0,1622		0,1768		0,1935	
<i>Pseudo-R² de McFadden ajustado</i>	0,1320		0,1489		0,1634		0,1764	
<i>AIC/N</i>	1,9071		1,8701		1,8381		1,8096	

Significatividad: ****; $p < 0,001$; ***; $p < 0,01$; **; $p < 0,05$; +; $p < 0,1$

En la segunda columna de la Tabla 3.2 se muestran las estimaciones para la especificación del logit heterogéneo. Este modelo es estadísticamente mejor que el primer modelo según el test del cociente de verosimilitud⁶, que indica que el poder explicativo del modelo mejora como resultado de los términos de interacción significativos. Esto sugiere que una parte substancial de la heterogeneidad de las preferencias entre los encuestados se debe a que algunos coeficientes varían de una forma específica para algunos segmentos de turistas. Por ejemplo, para el segmento de turistas que viajan con hijos, la posibilidad de que la casa rural tenga una minigranja ejerce un efecto positivo mayor sobre la probabilidad de elección de la casa que para los turistas que no viajan con niños. El segmento de turistas que viaja con seis o más personas incrementa positivamente sus preferencias por casas rurales de mayor tamaño. Los turistas que habitualmente hacen turismo rural en la zona (dos o más veces al año) valoran el precio de la vivienda menos negativamente que aquellos que visitan la zona por primera o segunda vez. Sin embargo, la valoran más negativamente si ha de alquilarse la casa completa. Por último, el segmento de turistas con un título universitario, grado o similar, tiene mayor preferencia por las casas rurales tradicionales y valoran la certificación de calidad “Q”, pero menos que los turistas con niveles de formación inferiores.

La tercera columna de la Tabla 3.2 muestra las estimaciones de la especificación del logit mixto 1, que resulta ser también estadísticamente mejor que el primer modelo según el test del cociente de verosimilitud. Así, el poder explicativo del modelo mejora como resultado de permitir preferencias heterogéneas aleatorias entre los individuos para algunos atributos de los alojamientos. En este modelo, cinco de las variables tienen coeficientes aleatorios, “Nueva”, “Huerta”, “Minigranja”, “Completa” y “Baño”. Todos estos coeficientes siguen una distribución normal. La media y la desviación típica para el coeficiente de la variable “Nueva” son significativos. Su media tiene un valor negativo y, por tanto, la variable “Nueva” ejerce un efecto disuasorio a la hora de elegir este tipo de casa. Sin embargo, la desviación típica de este coeficiente es bastante grande, por lo que no todos los individuos la valoran igual de negativamente. De hecho, los valores que presentan la media y la desviación típica para esta variable indican que el 72,4% de los individuos considera negativamente que la casa sea de nueva construcción en relación a que no lo sea, mientras que un 27,6% lo valoran positivamente. Las medias de los otros cuatro coeficientes que se consideran aleatorios no son significativamente diferentes de cero, mientras que sus desviaciones típicas son bastante grandes y muy significativas. Por lo que los turistas tienen una amplia variedad de puntos de vista acerca de compartir un cuarto de baño,

⁶También indican lo mismo el resto de estadísticos de bondad del ajuste calculados.

sobre si alquilar la casa en su totalidad o por habitaciones, en cuanto a que se sitúe en la huerta o a que disponga de minigranja.

La última columna de la Tabla 3.2 presenta las estimaciones del logit mixto 2. Este modelo considera el análisis de la heterogeneidad sistemática y permite que cuatro parámetros varíen aleatoriamente sobre los encuestados. Estos parámetros son los mismos que en el modelo logit mixto anterior, excepto para el coeficiente de la variable “Minigranja”. Además, en el modelo logit mixto 2, la media y la desviación típica de la variable “Casa alquilada completa” son significativos, indicando que el 62,75 % de los individuos consideran negativamente que la casa únicamente se pueda alquilar en su totalidad, mientras que el 37,25 % piensan que es mejor alquilar la casa completa. Con respecto a los atributos de las casas, los resultados derivados de los modelos logit mixtos (1 y 2) son consistentes con los de los modelos logit homogéneo y logit heterogéneo, pero el valor de los parámetros estimados con los primeros (logit mixtos) son más altas, cuando éstas son significativamente distintas de cero. Esto es algo esperado, como se explica en Brownstone y Train (1999) y Espino et al. (2008), debido a que los coeficientes aleatorios disminuyen la varianza del término de error aleatorio de la utilidad.

Dado que el logit multinomial se puede considerar como una especificación del logit mixto (cuando ninguno de los coeficientes es aleatorio), podemos hacer una comparación formal entre ellos sobre la base de los estadísticos del cociente de verosimilitud (Brownstone et al., 2000). De acuerdo a los valores de las funciones de verosimilitud obtenidas, los dos modelos logit mixtos son estadísticamente mejor que los modelos logit multinomiales. Además, el modelo logit mixto es más informativo, porque tiene en cuenta cierta variación aleatoria en los gustos y puede ayudar a comprender mejor la toma de decisiones de los individuos.

Una vez estimados los modelos logit y los dos logit mixtos, procedemos a calcular las disposiciones a pagar para los diferentes atributos de las casas rurales. En la Tabla 3.3 se muestran estas disposiciones a pagar, que se han calculado usando las expresiones (3.8) y (3.9) definidas en el apartado teórico, según sea el atributo cuantitativo o cualitativo, respectivamente. Así, suponiendo un modelo de utilidad lineal y un atributo cualitativo, la disposición a pagar por la diferencia entre niveles de atributos será el negativo del cociente entre la diferencia de estos coeficientes de cada atributo y el coeficiente del precio del atributo. Por ejemplo, para el modelo logit homogéneo, se obtiene una disposición a pagar de 251,24 euros porque la casa sea de construcción tradicional frente a una de nueva construcción. Esto significa que la cantidad

Tabla 3.3: Disponibilidades a pagar

	Logit homogéneo	Logit heterogéneo	Logit mixto 1	Logit mixto 2
Construcción original versus de nueva construcción	251,24	341,01	348,61	456,57
Construcción original versus casa de madera	25,49	27,23	26,67	33,02
Casa de madera versus de nueva construcción	225,75	313,77	321,65	423,55
Casa de campo versus casa en un pueblo	217,56	293,51	190,75	247,77
Tamaño de la casa	213,57	296,73 ^a	191,70	254,91 ^a
Viajar con seis o más personas		338,76		290,23
Viajar con menos de seis personas		238,72		206,17
Casa con mini-granja	16,53	23,63 ^a	9,93	6,40 ^a
Viajar con niños		83,35		65,76
Viajar sin niños		-16,84		-33,83
Posibilidad de alquilar caballos	34,09	51,67	24,94	35,50
Casa alquilada completa	-24,34	-35,13 ^a	-7,82	-16,91 ^a
Viaja más de dos veces al año		-192,84		-155,75
Viaja dos o menos veces al año		-22,35		-5,67
Casa con el certificado de calidad “Q”	50,80	65,17	56,17	72,20

^a valor medio

máxima que la población encuestada estaría dispuesta a pagar por una casa de construcción antigua versus una de nueva construcción para pasar un fin de semana en una casa rural en la Comarca del Noroeste de la Región de Murcia sería de 251,24 euros.

De acuerdo a la Tabla 3.3, para los cuatro modelos estimados, los atributos más valorados son el tipo de construcción, la ubicación y el tamaño de la casa, que indica que la población encuestada está dispuesta a pagar una suma importante por una construcción original, situada en el campo o la montaña y una casa de tamaño grande. Aunque, una habitación adicional es más valorada cuando el turista viaja con muchas personas. Además, los turistas están dispuestos a pagar entre 50 y 72 euros más por un casa rural que posea el certificado de calidad “Q” frente a una que no lo tenga. Los turistas están dispuestos a pagar entre 25 y 52 euros por tener la posibilidad de alquilar un caballo. Con respecto a la forma de alquiler de la vivienda,

los individuos están dispuestos a pagar más por una casa que se puede alquilar por habitaciones en oposición a una casa que sólo se puede alquilar en su totalidad, aunque las cantidades que están dispuestos a pagar difieren dependiendo de si son viajeros frecuentes o no. Finalmente, cuando el viaje es con niños, los turistas están dispuestos a pagar más si la casa dispone de una mini-granja.

3.5. Conclusiones

En este capítulo se analizan las preferencias de los turistas ante los atributos que definen e identifican las casas rurales de la Comarca del Noroeste de la Región de Murcia. Para la estimación de los modelos logit planteados, se han utilizado los datos obtenidos de un experimento de preferencias declaradas. Este experimento se contextualizó para que los turistas tuvieran que elegir entre dos alternativas de casas rurales, no etiquetadas y definidas en base a un conjunto de atributos y la opción de “quedarse en casa”.

En los modelos estimados, se analizan variaciones del gusto, sistemáticas y aleatorias, de los turistas que visitan esta zona. Los resultados obtenidos se comparan con el modelo logit homogéneo, donde la utilidad indirecta de los individuos depende sólo de los atributos de las casas rurales. La variación sistemática en el gusto se ha analizado empleando un modelo logit multinomial que incluye, en la función de utilidad, interacciones de características socio-económicas y de comportamiento de viaje de los turistas con los niveles de los atributos de las casas rurales. Por otro lado, las variaciones del gusto aleatorias se han estudiado usando especificaciones del modelo logit mixto, que permiten que los parámetros varíen aleatoriamente entre los turistas.

Los resultados de los modelos muestran que el atractivo de una casa rural para los turistas rurales de la Región de Murcia mejora – en diferentes grados – si es una casa de construcción tradicional remodelada para ser usada como alojamiento turístico, localizada en el campo o montaña, con un amplio número de habitaciones, que dispone de la certificación de calidad “Q” y ofrece la posibilidad de alquilar caballos. Hay otros atributos que también ejercen un efecto positivo sobre el atractivo de las casas rurales, pero sólo para algunos segmentos de turistas. Por ejemplo, la oferta de una minigranja podría ser atractiva para aquellos turistas que viajan con niños, o la posibilidad de reservar una casa por habitaciones puede atraer a los turistas que hacen turismo rural con frecuencia. Por último, otros atributos de las casas pueden tener tanto un efecto positivo como negativo para su elección,

estos son, compartir el cuarto de baño, ser una casa de nueva construcción, estar situada en un área de “huerta” o si la casa se alquila por habitaciones o al completo.

En el mercado de turismo rural, donde los establecimientos de alojamiento y negocios son pequeños, el estudio desarrollado en este capítulo proporciona una herramienta valiosa para los planificadores y administradores de establecimientos rurales. La evaluación de las preferencias de los turistas sobre las instalaciones y los servicios que ofrece cada establecimiento puede conducir al logro de un desarrollo mayor de su actividad económica. Por ejemplo, los gerentes de una casa rural pueden decidir tomar las medidas necesarias para conseguir ganar la certificación de calidad “Q”, o simplemente ofrecer la posibilidad de montar a caballo (a través de establos locales) como medio para hacer su establecimiento más atractivo. También pueden proponer ofertas de alquilar la casa en su totalidad o por habitaciones para atraer a los turistas frecuentes, o la creación de una minigranja como un imán para los turistas que viajan con niños. El conocimiento de las preferencias del consumidor para los atributos como el tipo de construcción, la situación o el tamaño, que no son fáciles de modificar, es también una herramienta importante para el desarrollo del turismo rural, especialmente en el caso de las decisiones de inversión en nuevas casas rurales.

Por otra parte, la información de este estudio podría ser una herramienta para el desarrollo de estrategias de marketing y campañas de publicidad para los establecimientos rurales de la Comarca del Noroeste de la Región de Murcia. La comparación de los atributos de un alojamiento rural existente con aquellos preferidos por los turistas pueden mejorar el conocimiento sobre la casa y proporcionar la base para promociones específicas.

Dado que los resultados de este estudio se basan en las respuestas de los turistas en la Región de Murcia, no es posible sacar conclusiones de los resultados más allá de esta región. Sin embargo, esta investigación muestra la viabilidad conceptual y práctica de los modelos de elección discreta con preferencias declaradas cuando se aplican para determinar las preferencias sobre los alojamientos turísticos. Esto también confirma el éxito de otros modelos que proporcionan ajustes mejorados sobre el logit y son más explicativos en términos de heterogeneidad de las preferencias de los turistas.

Capítulo 4

Los efectos de las motivaciones de los turistas rurales sobre la elección de alojamiento

4.1. Introducción

Los factores físicos que identifican a los hoteles y sus habitaciones influyen en su elección. Existen varios trabajos que señalan atributos tales como la ubicación, el precio, las instalaciones, el tamaño de las habitaciones de huéspedes, la estancia, la limpieza, el silencio o el aire acondicionado, como determinantes para la elección de un hotel (Lockyer, 2005; Stringam et al., 2010; Merlo y Joao, 2011; Chen et al., 2017; Masiero et al., 2015). La mayoría de estos atributos son meros descriptores de las características o cualidades físicas que dan al alojamiento un valor y lo identifican frente al resto, pero también existen otros factores que pueden ser determinantes cuando se trata de decidir el alojamiento donde pasar las vacaciones (Kim y Perdue, 2013; Martín et al., 2018). Recientemente, Kim y Park (2017) afirmaron que el contexto en el que los turistas toman sus decisiones, en su caso dependiendo de si era por motivo de ocio o de negocios, influye en la elección de un hotel. Esto es, mostraron que el criterio de elección del hotel puede cambiar dependiendo de los objetivos, las restricciones, y las características que los turistas tienen en el momento de tomar la decisión.

En el contexto de elección de un alojamiento rural, en el capítulo anterior hemos mostrado que atributos como el tamaño o el tipo de construcción, la calidad del alojamiento y los servicios y actividades ofrecidos eran decisivos en la elección de alojamiento rural. Pero al igual que sucede en la elección de

un hotel, otros factores podrían influir en la elección del alojamiento rural. Aunque muchos turistas que se deciden por un alojamiento rural, lo hacen motivados por el medio ambiente, la naturaleza y la cultura rural, como sugirió Lane (1994), varios estudios (Kastenholz et al., 1999; Frochot, 2005 y Molera y Albaladejo, 2007; entre otros) han probado la existencia de diferentes segmentos de turistas rurales en cuanto a las motivaciones que buscan o a los objetivos que se plantean cuando van al entorno rural. Hay turistas que buscan una imagen tradicional y auténtica de la vida en el campo, otros relajarse, otros simplemente quieren pasar tiempo con la familia y los amigos, mientras que otros se sienten atraídos por practicar actividades o deportes relacionados con la naturaleza. Así pues, los distintos grupos de turistas, dependiendo de las motivaciones o de los intereses que busquen, probablemente tendrán distintas preferencias en cuanto a los atributos que debe tener el alojamiento, tal y como probaron Kim y Park (2017) en cuanto a la distinción entre los viajeros de ocio y de negocios a la hora de elegir un hotel.

En este capítulo proponemos un modelo para testar si los turistas tienen preferencias diferentes en función de sus motivaciones o intereses para ir campo. Este modelo supone que la elección del alojamiento rural depende no sólo de los atributos observables del alojamiento, como el tamaño, la ubicación, el precio o las actividades realizadas, sino también de las motivaciones o intereses que los turistas quieren satisfacer yendo a un entorno rural. Las motivaciones de los turistas no son observables, resultan difíciles de medir y tienen muchas dimensiones, por lo que deben tratarse como variables latentes. Así, para poder trabajar con variables explicativas de ambos tipos, atributos observables del alojamiento y variables latentes que definen los motivos de los turistas en la elección de alojamiento, usamos el modelo híbrido de elección discreta (HDC). Este modelo es una extensión de los modelos de elección discreta (DC) que incorporan el efecto de las variables latentes a través de un modelo de ecuaciones estructurales de múltiples indicadores y múltiples causas (MIMIC) (Marquez et al., 2018). Para estimar las especificaciones propuestas del modelo HDC, utilizamos los datos de la encuesta realizada en la Comarca del Noroeste de la Región de Murcia durante varios fines de semana en otoño de 2003 y descrita anteriormente en el capítulo 1.

Con este estudio, nos proponemos dos objetivos. Por un lado, comprobar que las motivaciones o intereses que buscan los turistas para ir al campo son decisivos en la elección de su alojamiento y en la determinación de los atributos que un establecimiento debe tener. Para la consecución de este objetivo, como ya se ha comentado anteriormente, es necesario utilizar un modelo híbrido de elección discreta (HDC). Hasta donde sabemos, esta es la

primera vez que se utiliza en un estudio de turismo u hospitalidad. Un segundo objetivo es constatar que las motivaciones implican ciertas restricciones o limitaciones a la hora de elegir el alojamiento rural. De esta manera, conseguimos extender la hipótesis de Kim y Park (2017) sobre que el contexto de elección influye en la decisión de elección de hotel al contexto de elección de un alojamiento en las zonas rurales.

Además, el conocimiento de las motivaciones de los turistas que afectan a la elección de un alojamiento puede ser un apoyo para comprender las preferencias de un determinado tipo de turistas y, por lo tanto, ayudar a los promotores y propietarios de alojamientos rurales a diseñar una promoción de alojamiento adecuada para ellos. Asimismo, nuestros resultados también muestran que las motivaciones definen preferencias heterogéneas en cuanto a los atributos de los alojamientos rurales. Esto también puede ser una ayuda en las decisiones de inversión en la oferta de alojamientos destinadas a mejorar las características favoritas de determinados tipos de turistas.

Este capítulo se estructura en 8 epígrafes. El segundo epígrafe se dedica a justificar por qué creemos que las motivaciones o intereses que los turistas quieren satisfacer cuando deciden ir al entorno rural pueden influir en la elección de un alojamiento donde pasar sus vacaciones. En el tercer epígrafe desarrollamos la metodología que nos permite incluir en los modelos de elección discreta, como variables explicativas, estas motivaciones (latentes), esto es, los modelos híbridos de elección discreta. En el cuarto epígrafe presentamos un modelo especial de ecuaciones estructurales, llamado MIMIC, donde las motivaciones (latentes) vienen explicadas por las características socio-económicas y las condiciones de viaje de los turistas y éstas, a su vez, explican una serie de indicadores motivacionales. En el epígrafe quinto se detallan los datos y las variables que se usarán para estimar el modelo MIMIC que se desarrolla en el sexto epígrafe. En el séptimo epígrafe se estiman, con los datos mencionados anteriormente, un modelo MNL y dos especificaciones de HDC, lo que nos permite constatar que los dos modelos HDC estimados ofrecen mejoras significativas en la bondad del ajuste respecto del modelo MNL y evidencian que las motivaciones pueden afectar a la probabilidad de elección de alojamiento rural. El capítulo finaliza exponiendo las principales conclusiones del estudio realizado.

4.2. Las motivaciones de los turistas y la elección del alojamiento rural

En la literatura de la hospitalidad, los atributos que influyen en la elección de alojamiento son básicamente descriptores de sus características físicas (tamaño de las habitaciones de los huéspedes, aire acondicionado, ...) o prestaciones ofrecidas por estos alojamientos (servicio y calidad de los alimentos, limpieza, ...), tal y como mostraron Lockyer (2005), Stringam et al. (2010) y Chen et al. (2017), entre otros. Estos modelos de elección tradicionales suponen que los consumidores valoran sus elecciones basándose sólo en atributos cognitivos (Kim y Perdue, 2013). Sin embargo, otros paradigmas teóricos dentro de la investigación de mercado y del proceso de toma de decisiones sugieren la existencia de más sistemas para valorar las alternativas en una elección. En la elección del alojamiento, Kim y Perdue (2013) utilizan atributos cognitivos, afectivos y sensoriales para analizar la elección del hotel. Relatan que dos dimensiones básicas operan en la elección de los turistas: la cognitiva, relativa a las dimensiones físicas, y la experiencial, referida a los atributos afectivos y sensoriales. Su estudio empírico se basó en la auto-teoría cognitiva-experiencial (CEST), una de las teorías de procesamiento dual analizadas en Epstein (2003), que proporcionó una justificación teórica de la utilidad de los sistemas cognitivos y experienciales en el procesamiento de la información. Más tarde, Kim y Park (2017), basándose en la teoría del procesamiento de decisiones constructivo (Bettman et al., 1998), propusieron que la elección del hotel podía variar en función del contexto de elección o de los objetivos de elección. En particular, estos autores examinan las diferencias en la elección del hotel entre dos tipos de turistas con contextos diferentes (ocio versus negocios).

En el contexto rural, los alojamientos suelen ser negocios pequeños, tales como hoteles rurales, B&B, casas de alquiler y casas de huéspedes (Ye et al, 2019), que normalmente son administrados por las familias locales como medida de supervivencia económica y para la revitalización de estas áreas. En el capítulo anterior de esta memoria mostramos que los atributos cognitivos de los alojamientos, como el tamaño, la ubicación, el precio o las actividades que se ofrecen son decisivos en la elección del alojamiento rural. Pero, siguiendo la propuesta de Kim y Park (2017), la importancia de los atributos cognitivos en la elección del alojamiento rural también puede variar dependiendo de los diferentes objetivos y las limitaciones que los turistas tengan en el momento de tomar la decisión de visitar el entorno rural.

En sus comienzos, el turismo rural estaba relacionado con las personas

que visitaban el campo en busca de un entorno natural y de la cultura rural (Lane, 1994). Sin embargo, varios estudios han demostrado que el mercado del turismo rural es muy amplio y que las motivaciones para ir al campo son muy variadas. Kastenholz et al. (1999), Frochot (2005) y Molera y Albaladejo (2007), entre otros, comprobaron la existencia de varios segmentos de turistas rurales con diferentes motivaciones o que tienen diferentes intereses para ir a un entorno rural.

Las motivaciones o intereses que los turistas buscan satisfacer cuando deciden ir al campo podría implicar ciertas restricciones o limitaciones a la hora de elegir el alojamiento donde pasar sus vacaciones. Si los turistas buscan pasar tiempo con la familia y los amigos, tenderán a viajar con los niños y probablemente a poner más interés en los alojamientos que ofrezcan actividades deportivas u otro tipo de actividad, que les permita pasar tiempo en común. Sin embargo, si los turistas están muy interesados en la naturaleza y en relajarse, ellos tenderán a dar un valor bajo a las actividades y presumiblemente elegirán alojamientos alejados y tranquilos.

Teniendo en cuenta lo anterior, nuestra propuesta es que la elección de alojamiento rural no sólo depende de los atributos cognitivos sino también de otras variables que nos permiten definir las motivaciones o intereses de estos turistas rurales. Pero la determinación de estas motivaciones no es una tarea fácil, ya que al igual que sucede con las percepciones o actitudes (Frochot, 2005), no son variables directamente observables (Walker y Ben-Akiva, 2002) y deben medirse a través de su relación con otras variables (indicadores) que suponemos que están vinculadas con las características socio-demográficas de los turistas. La inclusión de las motivaciones como variables explicativas en los modelos de elección discreta nos conduce a trabajar con los modelos híbridos de elección discreta (HDC).

4.3. Modelos híbridos de elección discreta

Tradicionalmente, los modelos de elección discreta consideran como variables explicativas sólo características observables y tangibles de las alternativas y los turistas. Sin embargo, en las elecciones también pueden influir otros factores que no son fácilmente observables, denominados *variables latentes*, y que son inferidas (a través de un modelo matemático) a partir de otras variables que se observan, conocidas como *variables manifiestas o indicadores* (Park y Yoon, 2009). En las últimas décadas, se han desarrollado diferentes propuestas con el objetivo de incluir el efecto tanto de las

variables observables como de las latentes en los modelos de elección discreta (Bahamonde-Birke y Ortuzar, 2014b; Keane, 1997; Bollen, 1989; Green, 1984). Hoy en día, el más popular de estos modelos es el modelo híbrido de elección discreta (HDC). Las ventajas de usar un modelo HDC, tanto desde una perspectiva empírica como teórica, han sido demostradas en diferentes trabajos, como el de Ashok et al. (2002), Ben-Akiva et al. (2002), Vredin-Johansson et al. (2006), Tam et al. (2010) o Daziano y Bolduc (2013), entre otros. Además, hay muchas aplicaciones de los modelos HDC en la demanda del ciclismo (Motoaki y Daziano, 2015; Maldonado-Hinarejos et al., 2014).

La mayoría de los estudios que utilizan el marco de los modelos HDC consideran las percepciones y actitudes¹ como variables latentes (Walker y Ben-Akiva 2002; Raveau et al., 2010; Yañez et al., 2010; Dały et al., 2012; Bahamonde-Birke et al., 2015), aunque también se han estudiado otros efectos, como han señalado Thorhauge et al. (2019). Sin embargo, ninguno de estos estudios ha estudiado el efecto de las motivaciones o los intereses buscados.

Los modelos de elección discreta tradicionales suponen que sólo atributos tangibles y objetivos definen la utilidad determinística, V_{ij} . Los modelos HDC también incorporan elementos intangibles asociados a las actitudes, percepciones u otras características psicológicas de los individuos para definir V_{ij} . Por lo tanto, si suponemos una especificación lineal, como es habitual, de la utilidad determinística, V_{ij} , la función de utilidad U_{ij} se puede expresar como

$$U_{ij} = V_{ij} + \varepsilon_{ij} = \sum_k \beta_{jk} x_{ijk} + \sum_l \theta_{jl} \eta_{ijl} + \varepsilon_{ij} \quad (4.1)$$

donde x_{ijk} representa los atributos observados de las alternativas, η_{ijl} las variables latentes desconocidas, β_{jk} y θ_{jl} son los parámetros a estimar y ε_{ij} el término de error.

Siguiendo la teoría RUM, que supone que los individuos seleccionan la alternativa con el máximo valor de utilidad percibida de entre las alternativas de un conjunto de opciones, la probabilidad de que el individuo i seleccione la

¹Como señalan Bahamonde-Birke et al. (2015), las percepciones y actitudes son expresiones de diferentes juicios de valor. Las percepciones se relacionan con la forma de percibir determinadas alternativas y las actitudes expresan un estado de ánimo de los individuos hacia la vida, la sociedad etc. Así, la actitud se asemeja a una característica socio-económica del individuo, mientras que una percepción está intrínsecamente asociada a una alternativa.

alternativa j en el conjunto de opciones B condicionada al vector de variables latentes η_{ij} viene dada por la expresión

$$\begin{aligned} P(y_{ij} = 1/\eta_{ij}) &= Pr[U_{ij} \geq U_{ik} \quad \forall k \in B, k \neq j] \\ &= Pr[\varepsilon_{ik} \leq \varepsilon_{ij} + (V_{ij} - V_{ik}) \quad \forall k \in B, k \neq j] \end{aligned} \quad (4.2)$$

Dependiendo de las suposiciones que se hagan sobre la distribución de los componentes aleatorios de la utilidad se pueden derivar modelos HDC diferentes. Lo más habitual es suponer que este componente aleatorio es independiente e idénticamente distribuido según la distribución Valor Extremo Tipo I con parámetro de localización 0 y parámetro de escala 1, que conduce al modelo logit multinomial,

$$P(y_{ij} = 1/\eta_{ij}) = \frac{\exp(\sum_k \beta_{jk} x_{ijk} + \sum_l \theta_{jl} \eta_{ijl})}{\sum_{q=1}^J \exp(\sum_k \beta_{qk} x_{ijk} + \sum_l \theta_{ql} \eta_{iql})} \quad (4.3)$$

Como las variables η_{ijl} son desconocidas, es necesario estimarlas también. La estimación de las variables η_{ijl} puede realizarse previa a la estimación del modelo de elección discreta (estimación secuencial del modelo HDC) o conjuntamente con las ecuaciones estructurales y de medida del modelo MIMIC (estimación simultánea del modelo HDC), según como se utilice la información disponible (Raveau et al., 2010).

La estimación secuencial del modelo HDC se realiza como un proceso en dos etapas. En la primera se estiman las variables latentes usando el modelo MIMIC de ecuaciones estructurales (véase epígrafe 4.4) y, en una segunda etapa, se añaden los valores esperados de las variables latentes al conjunto de variables explicativas observadas del modelo de elección discreta y se procede a estimar los parámetros asociados a unas y otras variables de la forma estándar.

El método simultáneo implica integrar la expresión (4.3) sobre todo el espacio de las variables latentes, ya que las variables latentes no se pueden observar directamente. De ahí que la expresión de la probabilidad de que el individuo i seleccione la alternativa j en el conjunto de opciones B sea

$$P(y_{ij} = 1) = \int_{\eta_i} \frac{\exp(\sum_k \beta_{jk} x_{ijk} + \sum_l \theta_{jl} \eta_{ijl})}{\sum_{q=1}^J \exp(\sum_k \beta_{qk} x_{ijk} + \sum_l \theta_{ql} \eta_{iql})} g(\eta_i) d\eta_i \quad (4.4)$$

donde $g(\cdot)$ es la función de densidad de las variables latentes. La dimensión de la integral coincide con el número de variables latentes, lo que causa

una mayor complejidad en la estimación y un mayor coste computacional. Además, para identificar las variables latentes es necesario introducir la información proporcionada por los indicadores motivacionales en la estimación, ya que los indicadores no son variables explicativas, sino que son endógenas a las variables latentes (Raveau et al., 2010). Por eso, la probabilidad conjunta de observar el conjunto de indicadores I_i y la elección de la alternativa j para el individuo i sea

$$P(y_{ij} = 1, I_i) = \int_{\eta_i} \frac{\exp(\sum_k \beta_{jk} x_{ijk} + \sum_l \theta_{jl} \eta_{ijl})}{\sum_{q=1}^J \exp(\sum_k \beta_{qk} x_{ijk} + \sum_l \theta_{ql} \eta_{iql})} f(I_i/\eta_i) g(\eta_i) d\eta_i \quad (4.5)$$

donde $f(\cdot)$ es la función de densidad de los indicadores.

Ambos métodos de estimación aseguran la consistencia de los parámetros y si bien es cierto que con el método simultáneo se obtienen estimaciones más eficientes de los parámetros (Palma et al., 2018), éste se ha usado muy poco por la gran complejidad del proceso de estimación y la dificultad en encontrar software comercial que lo incluya. Por otro lado, el método secuencial, que es el más empleado, permite, bajo ciertas condiciones, obtener estimaciones de los parámetros insesgadas (Ben Akiva et al., 2002; Raveau et al., 2010) y, además, las estimaciones obtenidas de los parámetros no son significativamente diferentes de las obtenidas con el método simultáneo para el caso del modelo logit multinomial (Maldonado-Hinarejos et al., 2014), que será el empleado en la parte empírica de este estudio.

4.4. Variables latentes y ecuaciones estructurales

Las variables inobservables o latentes se identifican estimándolas como función de otras variables observadas, conocidas como indicadores o manifiestas. Estos indicadores, que miden el nivel de acuerdo de los consumidores con un conjunto de cuestiones que les conciernen, pueden ser continuas o categóricas y, generalmente, se obtienen mediante una encuesta sobre preferencias (Muthen, 1993; Bahamonde Birke et al., 2015). Los indicadores categóricos pueden ser, a su vez, dicotómicos, esto es, con dos niveles en escala nominal, u ordinales con más de dos niveles. En las ciencias educativas, sociales y del comportamiento lo habitual es que las variables observadas

(indicadores o variables manifiestas) sean de tipo categórico con varias categorías ordenadas (por ejemplo, utilizando escalas de tipo Likert). El enfoque más popular para estimar las variables latentes como función de los indicadores ordinales es a través de un tipo de modelo de ecuaciones estructurales (SEM) denominado modelo MIMIC (Bahamonde-Birke et al, 2015). El SEM ha sido ampliamente utilizado en la literatura turística² (Nunkoo et al., 2013). Un ejemplo de su uso en el contexto del destino turístico español es el trabajo de Perles et al. (2011).

El modelo MIMIC, como todo modelo de ecuaciones estructurales, tiene dos partes, las ecuaciones estructurales y las ecuaciones de medida. En el caso del modelo MIMIC, las variables latentes se explican por características de los individuos y/o de las alternativas por medio de las ecuaciones estructurales y, su vez, las ecuaciones de medida explican los indicadores (actitudinales, de percepción o motivacionales) en función de las variables latentes. Por lo tanto, este modelo se define por

$$\eta_i = \Gamma\omega_i + \zeta_i \quad (4.6)$$

$$I_i = \Lambda\eta_i + \delta_i \quad (4.7)$$

donde I_i corresponde a un vector ($P \times 1$) de indicadores asociados con el individuo i , con P indicando el número de indicadores; Λ es una matriz ($P \times L$) de parámetros desconocidos que relaciona las L variables latentes con los P indicadores (llamados *cargas factoriales*); η_i es a un vector ($L \times 1$) de variables latentes asociadas con el individuo i , donde L indica el número de variables latentes y δ_i es un vector ($P \times 1$) de errores de medida normalmente distribuidos con media cero y con Θ como matriz de varianzas-covarianzas, habitualmente diagonal. Mientras que en la ecuación estructural (4.6), Γ es una matriz ($L \times M$) de parámetros desconocidos que relaciona las L variables latentes con las M características de los individuos y/o de las alternativas; ω_i es un vector ($M \times 1$) de variables explicativas que causan las variables latentes, donde M es el número de variables explicativas que causan las variables latentes y ζ_i es un vector ($L \times 1$) de errores estructurales o residuos que típicamente sigue una distribución normal multivariante con media cero y matriz de covarianza Ψ .

El método de estimación más utilizado en la modelización de ecuaciones estructurales es el método de máxima verosimilitud (ML) sobre la matriz de

²En Nunkoo et al. (2013) los autores realizan una revisión del uso de modelos de ecuaciones estructurales basados en la matriz de covarianzas, publicados en nueve revistas de turismo entre 2000 y 2011.

covarianzas de la muestra. Este método parte de la suposición de que los indicadores son variables continuas que siguen una distribución normal multivariante en la población y, por tanto, no es apropiado cuando las variables observadas no son continuas o aún siendo continuas, la suposición de que siguen una distribución normal multivariante no es plausible (Li, 2006). Tradicionalmente y bajo ciertas circunstancias, la no continuidad de las variables observadas ha sido considerada, en la literatura especializada, compatible con variables observadas categóricas en escala ordinal.

Cuando los indicadores son variables ordinales, se debe adoptar otros enfoques. El primero pasa por tratar a las variables ordinales observadas como si fueran aproximadamente continuas y el segundo, tratarlas propiamente como categóricas utilizando el enfoque de la variable de respuesta subyacente (URV) (Joreskog, 1994; Joreskog y Moustaki, 2001; Muthen, 1984). Este segundo enfoque supone que cada una de las variables ordinales observadas I_i está generada por una variable continua inobservada subyacente, I_i^* , de manera que la variable categórica I_i con k_r categorías de respuesta, toma el valor k , con $k = 1, 2, \dots, k_r$ sí y sólo si se verifica que existe un conjunto de parámetros $\tau_{r,k}$, denominados *umbrales*, tales que

$$I_{ri} = \begin{cases} 1 & \text{si } -\infty = \tau_{r,0} < I_{ri}^* \leq \tau_{r,1} \\ 2 & \text{si } \tau_{r,1} < I_{ri}^* \leq \tau_{r,2} \\ \vdots & \vdots \\ k_r & \text{si } \tau_{r,k_r-1} < I_{ri}^* \leq \tau_{r,k_r} = +\infty \end{cases} \quad (4.8)$$

En este caso, la ecuación de medida (4.7) sería

$$I_i^* = \Lambda\eta_i + \delta_i \quad (4.9)$$

donde se ha reemplazado I_i por I_i^* . Bajo este enfoque, las ecuaciones de medida para indicadores ordinales vienen definidas por las ecuaciones (4.9) y (4.8) y representan un sistema de modelos probit ordenados.

Si nos centramos en el primer enfoque para indicadores ordinales, el método de estimación es un método de ML con correcciones robustas (ML robusto) basado en la matriz de covarianza de la muestra y fue propuesto por Satorra y Bentler (1994) y Yuan y Bentler (1998). Este método se desarrolló para variables observadas aproximadamente continuas pero potencialmente no normales, que podrían ser compatibles con las variables observadas ordinales con al menos cinco categorías de respuesta (Green et al., 1997).

El segundo método, propuesto por Muthén (1984; véase también Joreskog, 1994), utiliza un estimador por mínimos cuadrados no ponderado (ULS) o ponderado diagonalmente (DWLS), con libre distribución asintótica sobre la matriz de correlación policórica. Este enfoque se considera superior al ML convencional cuando se emplean datos de variables de respuesta ordinales en la modelización de las variables latentes (Li, 2016). No obstante, las estimaciones de los parámetros obtenidas por ULS y DWLS, al no trabajar con la matriz de pesos completa, no son asintóticamente eficientes, lo que provoca que las estimaciones de los errores estándar sean potencialmente inexactas (Savalei, 2014) con ambos estimadores, aunque resultan ligeramente mejores las estimaciones con DWLS (Bolt, 2005). Esto se puede superar implementando correcciones robustas sobre los errores estándar en la estimación de la matriz de covarianza asintótica de las estimaciones de los parámetros (véase Satorra y Bentler, 1994 para ULS y Muthén, du Toit y Spisic, 1997, para DWLS). Además, también puede realizarse un tipo de corrección robusta sobre el estadístico chi-cuadrado, lo que implica ajustar tanto la media como la varianza del test estadístico para hacer su forma aproximadamente a la de la distribución chi-cuadrado de referencia que se asocia con esos grados de libertad (Asparouhov y Muthén, 2010).

4.5. Datos y variables

Los datos usados en este estudio provienen de la encuesta realizada en la Comarca del Noroeste de la Región de Murcia durante varios fines de semana en el otoño de 2003 y que ya ha sido comentada en el capítulo 1. Esta encuesta incluye datos individuales de los turistas sobre: (1) sus características socio-económicas y condiciones de viaje; (2) sus valoraciones, en una escala Likert de 5, de un conjunto de indicadores destinados a la identificación de las motivaciones más importantes para visitar una zona rural; y (3) los datos de preferencias declaradas, necesarios para averiguar qué atributos de los alojamientos resultan relevantes en la elección de un establecimiento rural.

Las características socio-económicas y las condiciones de viaje de los turistas, mostradas en la Tabla 4.1, junto con las valoraciones realizadas por los turistas sobre los indicadores motivacionales, que se muestran en la Figura 4.1 con un gráfico de barras apilado divergente³, se han utilizado en la

³El gráfico de barras divergentes permite visualizar para cada ítem, de una forma sencilla y rápida, aquellas respuestas que son favorables o positivas de aquellas que son desfavorables o negativas, al eliminar de la representación el porcentaje de respuestas de la categoría “Ni de acuerdo ni en desacuerdo”. Este tipo de gráficos son de gran utilidad para presentar los resultados de cuestionarios que utilizan una escala Likert.

construcción y estimación del modelo MIMIC, mientras que los atributos que definen los alojamientos (véase Tabla 4.2) se usarán en la estimación de los modelos híbridos de elección junto con la variables latentes identificadas con el modelo MIMIC.

Aunque en el experimento de preferencias declaradas se completaron 307 cuestionarios, todos ellos válidos para el estudio, únicamente 290 individuos completaron todos los indicadores motivacionales, por lo que son estos últimos los que constituyen la muestra utilizada para la estimación de las especificaciones de los modelos híbridos de elección y del modelo logit multinomial.

Tabla 4.1: Características de los individuos incluidas en el modelo MIMIC

Tipo	Variable	Modalidad de referencia	Modalidades restantes
Características Socioeconómicas	Género del encuestado	Mujer	Hombre
	Provincia de origen	Resto de España	Murcia
	Edad del turista	Más de 40 años	Hasta 25 años De 26 a 40 años
	Nivel de formación alcanzada	Educación primaria	Educación secundaria o similar Educación superior
	Situación laboral	Desempleado, estudiante, tareas del hogar y jubilado	Ocupado
	Estado civil	Divorciado/separado	Soltero Casado
Condiciones de viaje	Viaja con niños	No	Sí
	Viaja con amigos	No	Sí
	Viaja con la familia	No	Sí
	Usa internet para organizar el viaje	No	Sí
	Gastos diarios por persona y día es superior a 60 euros (excluyendo alojamiento)	No	Sí
	Viaja en grupos de 6 o más personas	No	Sí
	Viaja a áreas rurales más de dos veces al año	No	Sí
	Frecuencia de viaje	Cinco o menos veces al año	Más de 5 veces al año

Figura 4.1: Valoración de los indicadores

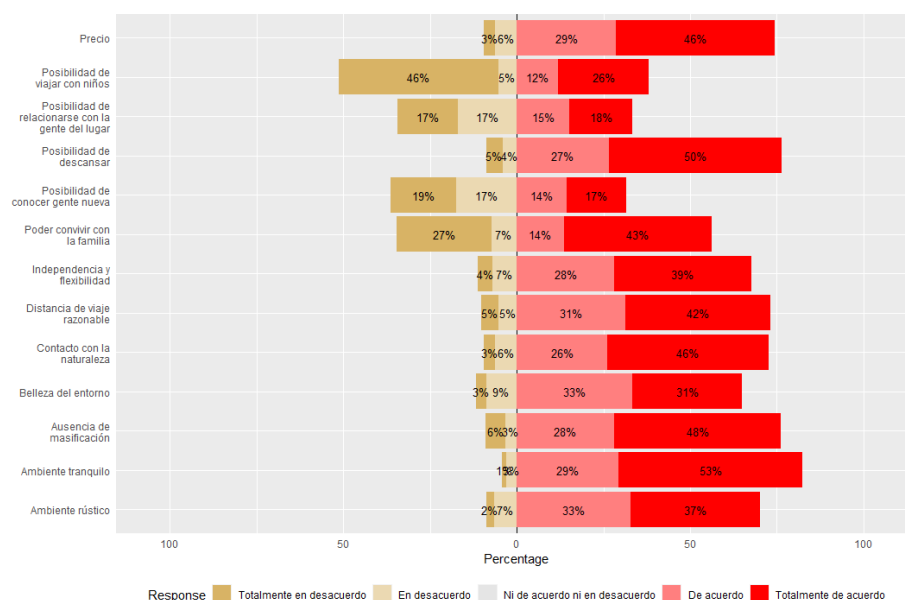


Tabla 4.2: Variables de los modelos HDC y del logit multinomial

Atributo	Tipo de variable	Definición
Precio	continua	Precio del alojamiento en euros
Tamaño	continua	Número de habitaciones del alojamiento
Construcción: Original	dicotómica	1 si el alojamiento es una casa típica de la zona
Construcción: Nueva	dicotómica	1 si el alojamiento es una casa de nueva construcción
Construcción: Madera	dicotómica	1 si el alojamiento es una cabaña de madera
Ubicación: Pueblo	dicotómica	1 si el alojamiento está situado a las afueras de un pueblo
Ubicación: Huerta	dicotómica	1 si el alojamiento está situado en zona de huerta
Ubicación: Campo-montaña	dicotómica	1 si el alojamiento está situado en el campo o la montaña
Mini-granja	dicotómica	1 si el alojamiento tiene minigranja
Alquiler de caballos	dicotómica	1 si el alojamiento ofrece el servicio de alquiler de caballos
Parque infantil	dicotómica	1 si el alojamiento dispone de un área delimitada con elementos de juego destinada a niños
Servicio de comidas	dicotómica	1 si el alojamiento dispone de servicio de comida
Piscina	dicotómica	1 si el alojamiento dispone de piscina
Instalaciones deportivas	dicotómica	1 si el alojamiento dispone de alguna pista para practicar deportes como fútbol, baloncesto, etc.
Baño compartido	dicotómica	1 si las habitaciones no tienen su propio cuarto de baño
Casa alquilada completa	dicotómica	1 si el alojamiento se alquila al completo
Calidad	dicotómica	1 si el alojamiento posee la "Q" de calidad
Reserva por internet	dicotómica	1 si el alojamiento se puede reserva a través de internet

4.6. Definición y estimación del modelo MI-MIC

Como se indicó en el epígrafe 4.3, la estimación secuencial de un modelo HDC, que es la que utilizamos en este estudio, implica un proceso en dos etapas, que comienza con la estimación de las variables latentes por medio de un modelo MIMIC y finaliza con la inclusión del valor esperado de estas variables latentes, junto con los atributos de un alojamiento rural, como variables explicativas en un modelo de elección discreta. En este epígrafe nos centramos en la definición y estimación de las variables latentes.

4.6.1. Definición de las variables latentes

El primer paso en la construcción de un modelo híbrido de elección es definir las variables latentes. Con el fin de identificar las variables latentes subyacentes, se realizó un análisis factorial exploratorio (AFE) sobre las respuestas a las preguntas relativas a los indicadores motivacionales. Este análisis se utiliza a veces para guiar la construcción del instrumento de medida en los modelos de ecuaciones estructurales. En la Tabla 4.3 se muestran los resultados obtenidos con el AFE realizado sobre la matriz de correlación policórica usando la librería *psych* de R (Revelle, 2021). Según estos resultados, parece razonable elegir una solución con cuatro factores, que explica alrededor del 73 % de la varianza de la información contenida en las trece variables observadas o indicadores. El factor 1 lo denominamos *Medio Ambiente* porque se refiere al entorno que se disfruta en un contexto rural. El factor 2, *Individualidad*, alude a las condiciones individuales de viaje y búsqueda de paz. *Familia* es el nombre designado para factor 3, ya que los indicadores están relacionados con los viajes en familia. Finalmente, *Sociabilidad* es el factor 4, que explica los viajes motivados principalmente con el objetivo de entablar contacto con otras personas.

Después del AFE se llevó a cabo un análisis factorial confirmatorio (AFC) para ratificar las variables latentes definidas. En la Tabla 4.4 se muestran las cargas factoriales estandarizadas (λ) de los indicadores para las cuatro variables latentes. Todas ellas son significativas. Además, los estadísticos

de bondad del ajuste⁴ indican un ajuste global aceptable⁵ para los datos (CFI=0,975, TLI=0,965, RMSEA=0,052, SRMR=0,057, WRMR=0,775).

Tabla 4.3: Análisis factorial exploratorio (AFE)

Item	Factor 1	Factor 2	Factor 3	Factor 4
Ambiente tranquilo	0,7055	0,1387	0,0500	0,0526
Belleza del entorno	0,8052	-0,0647	0,0682	0,02039
Contacto con la naturaleza	0,7229	-0,0803	0,2790	0,0488
Ausencia de masificación	0,8117	0,0820	-0,1601	-0,0807
Ambiente rústico	0,7186	0,0661	-0,1208	0,0582
Poder convivir con la familia	0,0167	-0,0020	0,9221	0,05210
Posibilidad de relacionarse con la gente del lugar	0,0622	-0,0131	0,0248	0,9205
Posibilidad de conocer gente nueva	-0,0500	0,0235	-0,0198	0,9593
Posibilidad de viajar con niños	0,0018	0,0425	0,9348	-0,0188
Posibilidad de descansar	0,1494	0,7118	0,1952	-0,0396
Independencia y flexibilidad	0,0570	0,7871	-0,0977	0,1349
Distancia de viaje razonable	-0,0732	0,7768	0,1346	0,0296
Precio	-0,0095	0,8350	-0,0880	-0,0626
<i>Valor propio</i>	4,9807	1,8752	1,4849	1,0950
<i>Porcentaje de varianza explicada</i>	0,2321	0,1985	0,1524	0,1429
<i>Porcentaje acumulado de varianza explicada</i>	0,2321	0,4306	0,5829	0,7258

Método de extracción: Análisis de componentes principales sobre la matriz de correlación policórica

Método de rotación: Oblimin

4.6.2. Validación del instrumento de medida

Una vez identificadas y verificadas las variables latentes subyacentes, el siguiente paso es examinar las interrelaciones entre las variables latentes y las variables explicativas (socio-económicas y de condiciones del viaje) mediante un modelo MIMIC. El modelo MIMIC estima simultáneamente las

⁴Los estadísticos de bondad del ajuste nos informan hasta qué punto la matriz de varianzas-covarianzas poblacional estimada se parece a la matriz varianzas-covarianzas muestral derivada de las puntuaciones en los indicadores o, en el caso de que los datos estén estandarizados, la semejanza entre la matriz de correlación poblacional estimada y la muestral.

⁵Los estudios de simulación en Yu y Muthen (2002) sugieren que los valores de corte para los índices de bondad del ajuste propuestos por Hu y Bentler (1999) para ajustes buenos con indicadores continuos (CFI > 0,95, TLI > 0,95, RMSEA < 0,06, SRMR < 0,08) también son razonables con indicadores categóricos, excepto que el punto de corte para SRMR de 0,08 no funciona bien cuando el tamaño muestral es pequeño (≤ 250). Yu y Muthen (2002) sugieren WRMR < 0,90 para ajustes buenos tanto con indicadores continuos como categóricos.

ecuaciones de medida que relacionan cada variable latente con sus indicadores y las ecuaciones estructurales que especifican las relaciones entre las variables latentes y las características socio-económicas y condiciones de viaje de los turistas. No obstante, antes de estimar el modelo MIMIC, es conveniente asegurarse de que el instrumento de medida incorporado en él, y que se va utilizar para medir cada una de las variables latentes, tiene la “calidad” suficiente para estimar el modelo estructural. A esto se le conoce como validación del instrumento de medida e implica la comprobación, a partir de la información proporcionada con el análisis factorial confirmatorio, de que la escala de medida cumple las propiedades de fiabilidad y validez (convergente y discriminante).

La fiabilidad o consistencia interna tiene que ver con la idea de que las variables manifiestas o indicadoras que comparten una causa común, la variable latente que están midiendo, deben estar muy correlacionadas entre sí, ya que están diseñadas para medir lo mismo, pero de formas distintas. Para medir la fiabilidad del instrumento de medida se utilizan los coeficientes alfa de Cronbach (Cronbach, 1951) y Fiabilidad Compuesta (FR). El coeficiente Fiabilidad Compuesta (FR) para cada variable o factor latente se calcula utilizando la siguiente expresión

$$FR_i = \frac{\left(\sum_j \lambda_{ij}\right)^2}{\left(\sum_j \lambda_{ij}\right)^2 + \sum_j var(\epsilon_{ij})} \quad (4.10)$$

donde λ_{ij} es la carga factorial estandarizada del indicador j con la variable latente i y $var(\epsilon_{ij})$ es la varianza del término de error de cada indicador. Según Sharma (1996) valores de los coeficientes alfa de Cronbach y FR entre 0,70 y 0,90 son adecuados y por debajo de 0,50 son pobres.

Por otro lado, se entiende que una escala es válida cuando lo que está midiendo realmente es la variable latente que se supone que se quiere medir. La validez de la escala tiene varias dimensiones que es necesario comprobar, como son la convergente y discriminante, entre otras. La validez convergente existe cuando distintas medidas (variables manifiestas o indicadores) de un mismo concepto (variable latente) están muy correlacionadas entre sí, compartiendo (convergiendo) una proporción elevada de varianza (Aldás y Uriel, 2017). Para verificar si hay validez convergente debemos comprobar que las cargas factoriales que relacionan las variables manifiestas con la variable latente son significativas y toman valores altos⁶ y que el indicador Varianza

⁶Según Aldas y Uriel (2017) el nivel mínimo que debe exigirse a una carga es $>0,71$,

Extraída Promedio (AVE) para cada variable latente toma un valor superior a 0,5. El indicador AVE para la variable latente i viene dado por

$$AVE_i = \frac{\sum_j \lambda_{ij}^2}{\sum_j \lambda_{ij}^2 + \sum_j var(\epsilon_{ij})} = \frac{\sum_j \lambda_{ij}^2}{k}, \quad (4.11)$$

donde λ_{ij}^2 representa la varianza del indicador j explicada por la variable latente i y $var(\epsilon_{ij})$ la varianza no explicada de cada indicador por la variable latente i y k es el número de indicadores ⁷ que son explicados por la variable latente i . De la ecuación (4.11) se deduce que el indicador AVE no es más que la media de las varianzas de los indicadores que explica la variable latente que están midiendo. Por lo tanto, que AVE sea superior a 0,5 indica que la variable latente comparte más de la mitad de su varianza con sus indicadores (Fornell y Larcker, 1981), es decir, la variable latente explica, en promedio, al menos el 50 % de la varianza de sus indicadores. Esto es el mínimo exigible para que no haya problemas de validez convergente.

Con la validez discriminante comprobamos que una variable latente mide un concepto distinto del resto de las variables latentes. Para ello, tiene que cumplirse que la varianza que una variable latente comparte con sus variables manifiestas debe ser mayor que la que pueda compartir con otras variables latentes incluidas en el modelo. Para determinar la validez discriminante basta comparar que la varianza extraída media (AVE) de cada variable latente es superior al cuadrado de los coeficientes de correlación de esa variable latente con el resto de variables latentes (Fornell y Larcker, 1981).

En la Tabla 4.4, donde se muestran los resultados obtenidos en AFC, se puede observar que los coeficientes alpha de Cronbach y de Fiabilidad Compuesta (FC) para cada variable latente presentan valores superiores a 0,8 y están por encima del umbral recomendado de 0,60 (Bagozzi y Yi, 1988), verificando la fiabilidad del instrumento de medida incorporado en el modelo MIMIC. También se puede constatar la validez convergente de la escala, ya que las cargas factoriales estandarizadas son estadísticamente significativas y toman valores en un rango de 0,64 a 0,99. Además, la varianza media extraída (AVE) para cualquiera de las variables latentes es superior a 0,5. Por último, en la Tabla 4.5 se puede observar que los AVE de dos variables latentes cualesquiera es superior al cuadrado de la correlación entre ellos, indicando que no hay problema de validez discriminante.

aunque Bagozzi y Yi (1988) relaja este supuesto hasta un mínimo de 0,60.

⁷El denominador del indicador AVE es igual a k , ya que $\lambda_{ij}^2 + var(\epsilon_{ij}) = 1$ para cada indicador j .

Tabla 4.4: Fiabilidad y validez convergente del instrumento de medida

Indicador	Cargas factoriales (λ)	AC (α)	FC (ω)	AVE
<i>Medio ambiente</i>				
MA1. Ambiente tranquilo	0,846***			
MA2. Belleza del entorno	0,745***			
MA3. Contacto con la naturaleza	0,676***	0,840	0,842	0,518
MA4. Ausencia de masificación	0,675***			
MA5. Ambiente rústico	0,641***			
<i>Individualidad</i>				
I1. Posibilidad de descansar	0,872***			
I2. Independencia y flexibilidad	0,759***	0,831	0,835	0,562
I3. Distancia de viaje razonable	0,637***			
I4. Precio	0,709***			
<i>Familia</i>				
F1. Poder convivir con la familia	0,924***	0,906	0,906	0,828
F2. Posibilidad de viajar con niños	0,895***			
<i>Sociabilidad</i>				
S1. Posibilidad de relacionarse con la gente del lugar	0,993***	0,884	0,894	0,811
S2. Posibilidad de conocer gente nueva	0,797***			
p-valor (Chi -square)	0,000			
Comparative Fit Index (CFI)	0,975			
Tucker Lewis Index (TLI)	0,965			
Root mean square error of approximation (RMSEA)	0,065			
p-valor RMSEA $\leq 0,05$	0,052			
Standarized Root Mean Square Residual (SRMR)	0,057			
Weighted Root Mean Square Residual (WRMR)	0,775			

***p < 0,001; AC= Alpha de Cronbach; FC=Fiabilidad compuesta; AVE= Varianza extraída promedio

Tabla 4.5: Validez discriminante del instrumento de medida

	Medio Ambiente	Individualidad	Familia	Sociabilidad
Medio Ambiente	0.518			
Individualidad	0.393	0.562		
Familia	0.192	0.049	0.828	
Sociabilidad	0.130	0.123	0.113	0.811

AVE en la diagonal principal y el cuadrado de los coeficientes de correlación de las variables latentes (ρ^2) en el triángulo inferior

4.6.3. Estimación del modelo MIMIC

Una vez identificadas las variables latentes y validado el instrumento de medida, se estima un modelo MIMIC para relacionar las variables latentes con las variables explicativas y los indicadores. La estructura del modelo MIMIC se muestra en la Figura 4.2. Los vínculos entre las variables latentes y los indicadores (es decir, las ecuaciones de medida) siguen un modelo probit ordenado, mientras que los vínculos entre las características observables de los turistas y las variables latentes (es decir, las ecuaciones estructurales) se suponen lineales.

La estimación de este modelo MIMIC se ha realizado utilizando el estimador por mínimos cuadrados ponderados con media y varianza ajustados (WLSMV) implementado en la librería *lavaan* de R (versión 0.5-23.1097) para el análisis de las variables latentes (Rosseel, 2012). Nótese que el estimador del WLSMV implica una estimación de los mínimos cuadrados ponderados diagonalmente (DWLS), errores estándar robustos y estadístico ajustado en media y la varianza para compensar la pérdida de eficiencia cuando no se utiliza la matriz de ponderación completa (Li, 2016). La Tabla 4.6 reporta los resultados del modelo MIMIC⁸ que mide las motivaciones de los turistas.

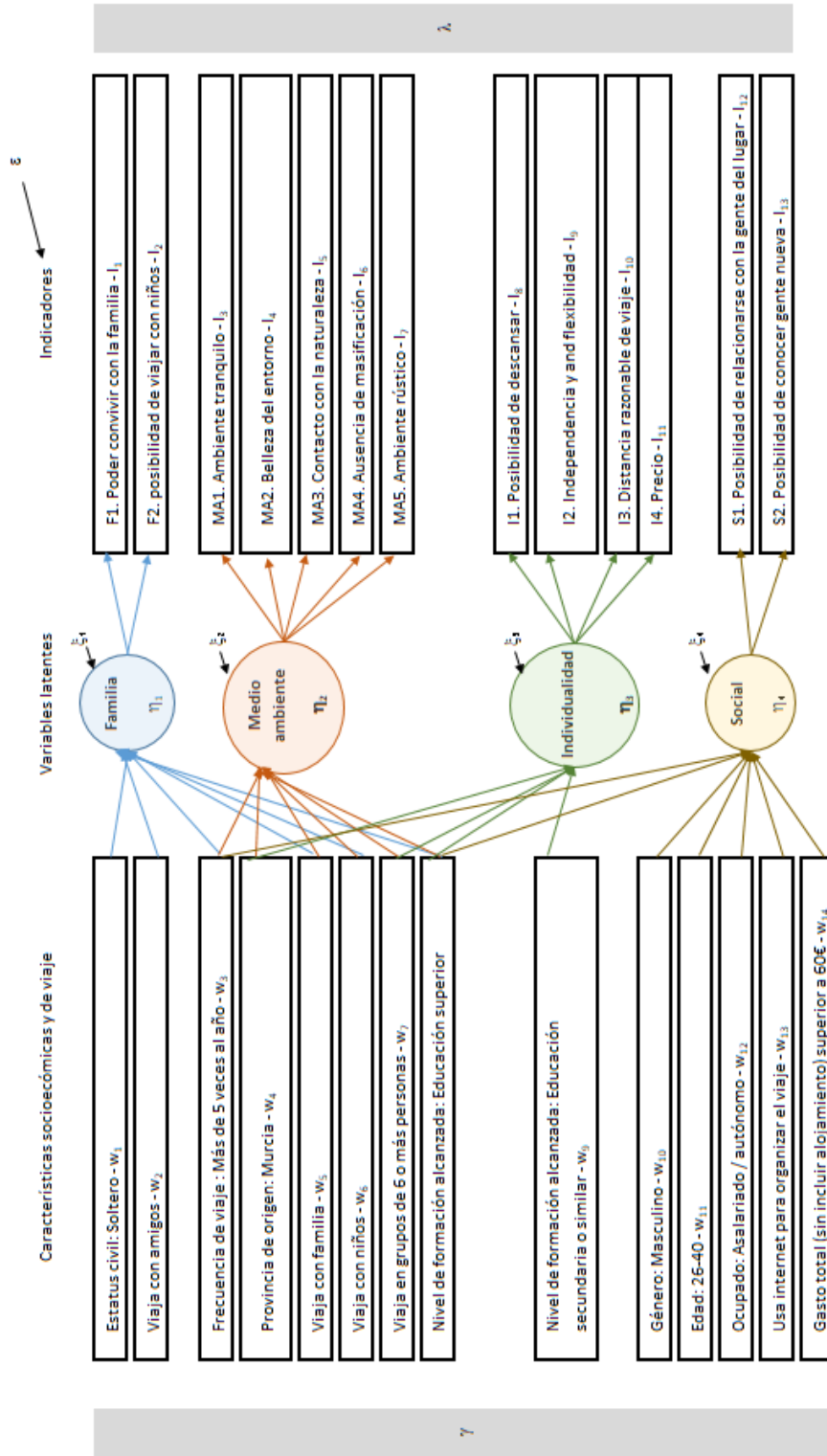
El signo de los coeficientes de los indicadores permite interpretar las variables latentes y las ecuaciones estructurales ayudan a explicar cómo se construyen las variables latentes mediante una relación causal (Motoaki y Daziano, 2015; Palma et al., 2018). Los coeficientes de los indicadores con signo positivo muestran que estos indicadores están positivamente correlacionados con la variable latente que los explica. Niveles elevados en los indicadores vinculados con una variable latente son causados por un nivel alto en dicha variable latente. Así, la variable latente *Medio Ambiente* recoge lo que sería la imagen tradicional y auténtica de la vida en el campo: el turista que busca estar en contacto con la naturaleza, disfrutar de la tranquilidad y la belleza del entorno natural y alejarse de los sitios concurridos. Niveles elevados de esta variable latente se relacionan con una frecuencia de viaje elevada (más de 5 veces al año), en grupos pequeños (5 personas máximo), con la familia e los hijos, con tener el lugar de residencia fuera de la Región de Murcia y no disponer de estudios universitarios.

⁸Aunque se han omitido las estimaciones de los parámetros de umbral, $\tau_{r,k}$, definidos en (4.8) en la tabla 4.6, es importante señalar que son significativos, lo que demuestra que los modelos ordenados propuestos representan adecuadamente la opinión de los individuos sobre las afirmaciones que captamos a través de las respuestas observadas en los indicadores (Soto et al., 2018).

Tabla 4.6: Estimación del modelo MIMIC

Ecuación de medida: $\Lambda (R^2)$	Estimación	s.e.	t	p-valor
<i>Medio Ambiente</i>	Fiabilidad: 0,859 (α), 0,860 (ω)			
MA1. Ambiente tranquilo (0,693)	1			
MA2. Belleza del entorno (0,488)	0,823***	0,063	13,025	0,000
MA3. Contacto con la naturaleza (0,576)	0,902***	0,072	12,553	0,000
MA4. Ausencia de masificación (0,521)	0,853***	0,066	13,016	0,000
MA5. Ambiente rústico (0,482)	0,818***	0,062	13,293	0,000
<i>Individualidad</i>	Fiabilidad: 0,871 (α), 0,873 (ω)			
I1. Posibilidad de descansar (0,747)	1			
I2. Independencia y flexibilidad (0,672)	0,935***	0,071	13,100	0,000
I3. Distancia de viaje razonable (0,653)	0,919***	0,076	12,139	0,000
I4. Precio (0,465)	0,750***	0,072	10,409	0,000
<i>Familia</i>	Fiabilidad: 0,899 (α), 0,899 (ω)			
F1. Poder convivir con la familia (0,810)	1			
F2. Posibilidad de viajar con niños (0,822)	1,016***	0,156	6,513	0,000
<i>Sociabilidad</i>	Fiabilidad: 0,905 (α), 0,905 (ω)			
S1. Posibilidad de relacionarse con la gente del lugar (0,805)	1			
S2. Posibilidad de conocer gente nueva (0,850)	1,034***	0,085	12,143	0,000
<i>Medio Ambiente (0,164)</i>				
Provincia de origen: Murcia	-0,239	0,131	-1,823	0,068
Nivel de formación alcanzada: Educación superior	-0,327	0,161	-2,033	0,042
Viaja con familia: Sí	0,337	0,204	1,651	0,099
Viaja con niños: Sí	0,308	0,185	1,663	0,096
Viaja en grupos de 6 o más personas: Sí	-0,389	0,133	-2,925	0,003
Frecuencia de viaje: Más de 5 veces al año	0,378	0,224	1,686	0,092
<i>Individualidad (0,294)</i>				
Provincia de origen: Murcia	-0,219	0,131	-1,678	0,093
Nivel de formación alcanzada: Educación secundaria o similar	-0,467	0,183	-2,559	0,010
Nivel de formación alcanzada: Educación superior	-0,685	0,181	-3,792	0,000
Viaja en grupos de 6 o más personas: Sí	-0,749	0,143	-5,253	0,000
<i>Familia (0,631)</i>				
Estado civil: Soltero	-0,669	0,194	-3,453	0,001
Nivel de formación alcanzada: Educación superior	-0,368	0,190	-1,940	0,052
Viaja con amigos: Sí	-0,445	0,119	-2,242	0,025
Viaja con familia: Sí	0,605	0,247	2,455	0,014
Viaja con niños: Sí	0,749	0,198	3,788	0,000
Frecuencia de viaje: Más de 5 veces al año	0,576	0,270	2,132	0,033
<i>Sociabilidad (0,224)</i>				
Género: Hombre	-0,205	0,124	-1,654	0,098
Edad: 26-40	0,482	0,196	2,460	0,014
Nivel de formación alcanzada: Educación superior	-0,661	0,162	-4,076	0,000
Situación laboral: Ocupado (asalariado /autónomo)	0,280	0,156	1,792	0,073
Usa internet para organizar su viaje: Sí	-0,235	0,132	-1,774	0,076
Gasto diario por persona y día superior a €60: Sí	0,238	0,128	1,851	0,064
Frecuencia de viaje: Más de 5 veces al año	0,543	0,251	2,169	0,030
p-valor (Chi -square)	0,048			
Comparative Fit Index (CFI)	0,984			
Tucker Lewis Index (TLI)	0,980			
Root mean square error of approximation (RMSEA)	0,023			

Figura 4.2: Estructura (gráfico de ruta) del modelo MIMIC



La variable latente *Individualidad* alude a la búsqueda de un alojamiento rural situado cerca de la vivienda habitual del individuo, con un precio asequible, que le permita descansar y que le proporcione una sensación de independencia y flexibilidad. Esta variable latente se explica a través del lugar de residencia, el nivel de formación alcanzado y el tamaño del grupo con el que viaja. Por tanto, niveles elevados de variable latente *Individualidad* se relacionan con turistas que viven fuera de la Región de Murcia, que prefieren viajar en grupos pequeños y tienen un nivel de educación bajo (como mucho han conseguido acabar la educación básica o primaria).

La variable latente *Familia* está relacionada con los individuos que están interesados en pasar un buen rato con la familia y la posibilidad de viajar con niños. Niveles elevados de la variable latente *Familia* se asocian con viajar más de 5 veces al año, con la familia y con niños, mientras que esta variable latente muestra niveles bajos para el individuo que viaja con amigos, es soltero y tiene estudios universitarios.

Por último, la variable latente *Sociabilidad* está asociada con personas que piensan que elegir un alojamiento rural es una oportunidad para conocer gente y relacionarse con los residentes locales. El modelo estructural, en este caso, indica que los turistas con edad entre 26 y 40 años muestran niveles elevados en la variable *Sociabilidad* en comparación con los turistas cuya edad está fuera de este rango. También se relacionan con valores elevados en la motivación *Sociabilidad* estar trabajando por cuenta propia o ajena, tener un presupuesto de viaje superior a 60 euros diarios y viajar más de 5 veces al año. El género del turista también ayuda a explicar los niveles en esta variable latente.

4.7. Definición y estimación de los modelos de elección

Una vez evaluado el modelo MIMIC, se estiman los valores esperados de las variables latentes para cada individuo utilizando un enfoque de regresión por mínimos cuadrados. Estas puntuaciones en las variables latentes, que son valores numéricos que indican la posición relativa de un individuo en la variable latente, se añaden posteriormente al conjunto de variables explicativas para estimar los modelos de elección discreta, completando así la estimación secuencial del modelo HDC.

Para detectar las preferencias de los turistas en cuanto a los atributos de los alojamientos rurales se han propuesto tres modelos de elección discreta, un MNL y dos especificaciones del modelo HDC. El modelo MNL sólo tiene en cuenta preferencias homogéneas entre los turistas, por lo que la función de utilidad indirecta únicamente depende de los atributos de los alojamientos rurales,

$$\begin{aligned}
V_{ij} = & \beta_1 \text{Precio}_j + \beta_2 \text{Tamaño}_j + \beta_3 \text{Original}_j + \beta_4 \text{Nueva}_j \\
& + \beta_5 \text{Madera}_j + \beta_6 \text{Pueblo}_j + \beta_7 \text{Huerta}_j + \beta_8 \text{Campo-montaña}_j \\
& + \beta_9 \text{Minigranja}_j + \beta_{10} \text{Caballos}_j + \beta_{11} \text{Pinfantil}_j + \beta_{12} \text{Comidas}_j \\
& + \beta_{13} \text{Piscina}_j + \beta_{14} \text{Deportivas}_j + \beta_{15} \text{Baño}_j + \beta_{16} \text{Alquiler}_j \\
& + \beta_{17} \text{Calidad}_j + \beta_{18} \text{Internet}_j \quad j = 1, 2, 3
\end{aligned} \tag{4.12}$$

En la primera especificación del modelo HDC, las variables latentes se introducen directamente en la función de utilidad y en forma lineal, tal y como se muestra en la siguiente expresión,

$$\begin{aligned}
V_{ij} = & \beta_1 \text{Precio}_j + \beta_2 \text{Tamaño}_j + \beta_3 \text{Original}_j + \beta_4 \text{Nueva}_j \\
& + \beta_5 \text{Madera}_j + \beta_6 \text{Pueblo}_j + \beta_7 \text{Huerta}_j + \beta_8 \text{Campo-montaña}_j \\
& + \beta_9 \text{Minigranja}_j + \beta_{10} \text{Caballos}_j + \beta_{11} \text{Pinfantil}_j + \beta_{12} \text{Comidas}_j \\
& + \beta_{13} \text{Piscina}_j + \beta_{14} \text{Deportivas}_j + \beta_{15} \text{Baño}_j + \beta_{16} \text{Alquiler}_j \\
& + \beta_{17} \text{Calidad}_j + \beta_{18} \text{Internet}_j + \beta_{19} \text{Ambiente}_i + \beta_{20} \text{Individualidad}_i \\
& + \beta_{21} \text{Familia}_i + \beta_{22} \text{Sociabilidad}_i \quad j = 1, 2, 3
\end{aligned} \tag{4.13}$$

Por último, en la segunda especificación⁹ de HDC, además de incorporar las variables latentes, también incluye interacciones entre las variables latentes y los atributos de las alternativas.

Con el fin de determinar las variables explicativas significativas en cada especificación se ha empleado un procedimiento clásico de selección secuencial. Este criterio implica estimar modelos alternativos eliminando las variables explicativas no significativas y elegir aquel modelo que es mejor según el test

⁹Para este modelo se omite incluir la expresión de la utilidad determinística, ya que supone incluir 72 términos más en la ecuación (4.13), que se corresponden con la interacción de cada variable latente con cada una de las variables que definen características de los alojamientos.

estimado utilizando el programa econométrico NLOGIT 5.0 (Greene, 2012) y los resultados se muestran en la Tabla 4.7 junto con las principales medidas de bondad del ajuste.

La especificación del MNL (primera columna de la Tabla 4.7), que supone coeficientes constantes entre individuos, parece ajustar los datos con un valor aceptable de 0,13 del coeficiente pseudo- R^2 de McFadden ajustado. Los coeficientes de las variables precio de la vivienda, vivienda de nueva construcción, situada en un pueblo y que se alquila completa son negativos, mientras que el resto son positivos. Así, los turistas de la Región de Murcia prefieren los alojamientos rurales de precio bajo y de gran tamaño, no construidos con el propósito expreso de ser alquilados, situados en el campo o en la montaña, que poseen la certificación “Q” de calidad y que ofertan la posibilidad de contratar excursiones en caballo. Esta especificación es similar a la realizada en el epígrafe 3.4 del capítulo 3.

La primera especificación del HDC, que incluye variables latentes en la utilidad indirecta en forma lineal (segunda columna de la Tabla 4.7), muestra un pseudo- R^2 ajustado de 0,16, un 26,5 % superior al del modelo MNL. Esto denota que la inclusión de las variables latentes identificadas por las ecuaciones estructurales mejoran la bondad del ajuste del modelo de elección. Los atributos de las alternativas significativos son los mismos que en el modelo MNL y, aunque los valores de sus coeficientes son diferentes, sus signos coinciden. Por lo tanto, la interpretación de los atributos de las alternativas es similar a la interpretación del modelo MNL. En cuanto a las motivaciones, *Individualidad*, *Familia* y *Sociabilidad* son significativas en la elección del alojamiento rural. El coeficiente de la motivación *Sociabilidad* tiene un signo positivo, lo que indica que a mayor motivación por establecer relaciones sociales, mayor será la probabilidad de elegir un alojamiento rural en esta zona, ceteris paribus. Sin embargo, los coeficientes de las motivaciones *Individualidad* y *Familia* son negativos, señalando que la probabilidad de elegir un alojamiento rural es menor para turistas con alta motivación para viajar con la familia o para hacer un viaje corto, tranquilo e individual.

Por último, la tercera columna de la Tabla 4.7 muestra los resultados del modelo HDC con interacciones entre las motivaciones latentes de los individuos y los atributos de los alojamientos rurales. Se estimaron las interacciones de cada uno de los atributos con las variables latentes; sin embargo, sólo se han incluido en el modelo final las interacciones significativas. La última especificación tiene el valor más alto para el pseudo- R^2 ajustado (0,18). Nuestros resultados muestran preferencias heterogéneas entre los encuestados, ya que

algunos coeficientes varían específicamente para ciertos segmentos de turistas según su motivación latente. Así, los individuos cuya motivación para elegir un alojamiento rural es el *Medio Ambiente*, valoran más negativamente que el resto de los turistas el atributo “Construcción: Nueva” de la casa rural, mientras que valoran más positivamente la posibilidad de poder alquilar la casa también por habitaciones.

Por otro lado, las casas de construcción original son percibidas más positivamente por los consumidores cuya motivación principal es la *Individualidad*, junto con la posibilidad de reservar alojamiento on-line¹⁰. El segmento de turistas motivados por la *Familia*, valora positivamente la posibilidad de que el alojamiento rural tenga una minigranja, ceteris paribus. Además, este colectivo muestra una preferencia por los alojamientos de gran tamaño superior a la que manifiestan el resto de los turistas y valoran negativamente que el alojamiento sea de nueva construcción, aunque en menor medida que el resto de turistas. Por último, para los turistas motivados principalmente por la *Sociabilidad*, el precio, el tamaño y la situación de la casa dentro de un pueblo son importantes, pero en menor grado que para otros turistas con motivación diferente, mientras que la construcción sea original resulta ser mucho más importante. Para este segmento, la probabilidad de elegir un alojamiento rural que oferte servicio de comidas es mayor, y es menor si dispone de instalaciones deportivas.

¹⁰Es importante señalar que cuando se hizo la encuesta, casi ningún establecimiento rural ofrecía este tipo de reserva.

Tabla 4.7: Resultados de las estimaciones de los modelos Logit

Atributos	MNL			Modelo HDC			Modelo HDC con interacciones		
	Estimación	SE	t-Stat	Estimación	SE	t-Stat	Estimación	SE	t-Stat
Precio	-0,3161***	0,0737	-4,29	-0,3055***	0,0730	-4,19	-0,4125***	0,0900	-4,58
Tamaño	0,6937***	0,0848	8,18	0,4982***	0,0867	0,75	0,7102***	0,1086	6,54
Construcción: Original	0,2873***	0,0758	3,79	0,2837***	0,0755	3,76	0,3737***	0,1266	2,95
Construcción: Nueva	-0,5738***	0,0830	-6,91	-0,5365***	0,0804	-6,67	-0,6526***	0,0965	-6,76
Construcción: Madera	0,1762**	0,0782	2,25	0,1606**	0,0768	2,09	0,1584**	0,0785	2,02
Ubicación: Pueblo	-0,4255***	0,0820	-5,19	-0,4094***	0,0801	-5,11	-0,4860***	0,0935	-5,2
Ubicación: Huerta	0,1132	0,0786	1,44	0,0938	0,0766	1,22	0,0993	0,0789	1,26
Ubicación: Campo-montaña	0,2796***	0,0770	3,63	0,2799***	0,0760	3,68	0,3170***	0,0861	3,68
Mini-granja	0,0683	0,0446	1,53	0,0680	0,0442	1,54			
Alquiler de caballos	0,1135**	0,0451	2,52	0,1081**	0,0444	2,44	0,1259***	0,0456	2,76
Casa alquilada completa	-0,0811*	0,0450	-1,8	-0,0883**	0,0443	-1,99	-0,0924	0,0700	-1,32
Instalaciones deportivas							0,1242	0,0845	1,47
Calidad							0,1620***	0,0455	3,56
Reserva del alojamiento por internet							0,1405**	0,0695	2,02
<i>Medio ambiente</i>				-0,1332	0,1583	-0,84	-0,1729	0,1636	-1,06
<i>Individuabilidad</i>				-0,3997***	0,1200	-3,33	-0,4678***	0,1350	-3,47
<i>Familia</i>				-0,1454*	0,0866	-1,68	-0,3601***	0,1297	-2,78
<i>Sociabilidad</i>				0,8771***	0,1175	7,46	0,9918***	0,1622	6,11
Nueva construcción x <i>Medio Ambiente</i>							-0,1852**	0,0891	-2,08
Ubicación: campo-montaña x <i>Medio Ambiente</i>							0,1165	0,0721	1,61
Piscina x <i>Medio Ambiente</i>							-0,0634	0,0432	-1,47
Tipo alquiler: casa completa x <i>Medio Ambiente</i>							0,1369*	0,0701	1,95
Tamaño x <i>Individuabilidad</i>							0,0913**	0,0374	2,44
Construcción original x <i>Individuabilidad</i>							0,1764**	0,0745	2,37
Instalaciones deportivas x <i>Individuabilidad</i>							0,0670	0,0510	1,31
Mini-granja x <i>Individuabilidad</i>							-0,0815**	0,0327	-2,49

Continúa en la siguiente página

Tabla 4.7 – Continúa de la página anterior

Atributos	MNL			Modelo HDC			Modelo HDC con interacciones		
	Estimación	SE	t-Stat	Estimación	SE	t-Stat	Estimación	SE	t-Stat
Tipo alquiler: casa completa x <i>Individualidad</i>							-0,08723	0,06517	-1,34
Reserva de internet x <i>Individualidad</i>							0,1053 **	0,0479	2,2
Tamaño x <i>Familia</i>							0,0967 ***	0,0368	2,63
Nueva construcción x <i>Familia</i>							0,1522 **	0,0616	2,47
Servicio de comidas x <i>Familia</i>							-0,0592	0,0403	-1,47
Mini-granja x <i>Familia</i>							0,1131 ***	0,0380	2,98
Precio x <i>Sociabilidad</i>							0,1561 *	0,0877	1,78
Tamaño x <i>Sociabilidad</i>							-0,2426 **	0,1072	-2,26
Construcción original x <i>Sociabilidad</i>							0,1803 **	0,0839	2,15
Ubicación: pueblo x <i>Sociabilidad</i>							0,1371 *	0,0805	1,7
Parque infantil x <i>Sociabilidad</i>							-0,0764	0,0475	-1,61
Servicio de comidas x <i>Sociabilidad</i>							0,0809 *	0,0479	1,69
Piscina x <i>Sociabilidad</i>							0,0699	0,04534	1,54
Instalaciones deportivas x <i>Sociabilidad</i>							-0,1220 **	0,0577	-2,12
<i>Fit índices</i>									
<i>Observaciones</i>	1160			1160				1160	
<i>Individuos</i>	290			290				290	
<i>Número de parámetros</i>	12			16				39	
<i>Función Log-verosimilitud</i>	-1099,66			-1053,10				-1013,11	
<i>Función Log-veros. restringida</i>	-1274,39			-1274,39				-1274,39	
<i>Pseudo R² de McFadden</i>	0,1371			0,1736				0,2050	
<i>Pseudo R² de McFadden ajustado</i>	0,1281			0,1621				0,1773	
<i>AIC/N</i>	1,9167			1,8433				1,8140	

4.8. Conclusiones

Basándonos en la teoría del proceso de elección constructiva del consumidor de Bettman et al (1998), en este capítulo proponemos que los objetivos motivacionales de los turistas para ir al entorno rural son determinantes en la elección del alojamiento e influyen en la importancia que los turistas asignan a los atributos del alojamiento. Para contrastar esta hipótesis, hemos estimado tres modelos de elección discreta tomando como muestra los turistas que visitaron la Comarca del Noroeste de la Región de Murcia, un MNL y dos especificaciones del modelo HDC. El modelo MNL considera como variables explicativas sólo los atributos objetivos y tangibles que definen a los alojamientos, como el tamaño, la ubicación, el precio o las actividades que ofrecen. Las dos especificaciones del modelo HDC, además de las anteriores variables, también introducen variables latentes que identifican las motivaciones de los turistas para visitar las zonas rurales.

La inclusión de las variables latentes en los modelos de elección discreta se realiza mediante el modelo de ecuaciones estructurales tipo MIMIC, que las vincula con indicadores motivacionales y con características socio-demográficas y condiciones de viaje de los turistas. En nuestro estudio, se han detectado cuatro variables latentes que identifican las motivaciones de los turistas rurales: *Medio ambiente*, *Individualidad*, *Familia* y *Sociabilidad*. Estas variables, incluidas en los modelos de elección discreta, nos permiten detectar preferencias heterogéneas entre los turistas. Nuestra estimación de los modelos de HDC muestra que los turistas con una mayor motivación por socializarse, tanto con residentes locales como con otros turistas, tienen una mayor probabilidad de elegir un alojamiento rural en la Región de Murcia que aquellos turistas motivados por pasar tiempo con la familia o motivados por hacer un viaje corto, tranquilo e individual. Pero las motivaciones también pueden denotar preferencias heterogéneas en cuanto a los atributos de los alojamientos rurales. Así, por ejemplo, el hecho de que el alojamiento tenga una minigranja influye positivamente en la elección del alojamiento para los turistas motivados por viajar con la familia y los niños, pero influye negativamente en los turistas que buscan hacer un viaje corto, tranquilo e individual.

Los dos modelos HDC estimados ofrecen mejoras significativas en la bondad del ajuste respecto el modelo MNL y evidencian que las motivaciones pueden afectar a la probabilidad de elección de alojamiento rural, y eso puede inclinar a los turistas hacia ciertos atributos de los alojamientos. Sin embargo, la estimación de los modelos HDC requiere mucha más información y

un proceso de estimación más complejo que un modelo MNL. Así que, de acuerdo con Palma et al. (2018) y Vij y Walker (2016), los modelos HDC deben usarse sólo para estudios en un contexto particular. Tal es el caso de nuestro estudio, que tenía como objetivo hacer un análisis de la elección del alojamiento rural, proponiendo que la motivación los objetivos de los turistas también influyen en sus elecciones.

El complejo análisis desarrollado en este estudio contribuye a la literatura al utilizar una modelización de elección discreta, los modelos HDC, que permite analizar el impacto de factores decisivos, y no directamente observables, en la elección del alojamiento, tales como las motivaciones de los turistas. Otros estudios previos sobre la elección de alojamiento utilizan modelos de elección discreta, pero hasta donde sabemos, éste es el primero en incorporar motivaciones para investigar el proceso de elección de alojamiento.

Los resultados de este estudio también apoyan el marco conceptual de Bettman et al. (1998) y muestran que la elección del alojamiento rural varía en función del contexto de decisión del turista (Medio Ambiente, Individualidad, Familia o Sociabilidad) y, además, que el contexto de decisión también condiciona sus preferencias en relación con los atributos del alojamiento. En consecuencia, el contexto de decisión es decisivo en la elección del alojamiento turístico, ya sea un hotel (Kim y Park, 2017) o un alojamiento rural (Albaladejo y Díaz, 2020).

Por último, nuestros resultados también indican que los turistas pueden inclinarse hacia ciertos atributos del alojamiento. Por consiguiente, este estudio podría ser una herramienta valiosa para los planificadores y administradores de establecimientos rurales al determinar las preferencias de los turistas por las instalaciones y servicios que ofrece un alojamiento rural. Proporciona información sobre los atributos de los alojamientos que son apreciados por los turistas, aquellos que no son valorados por ningún turista sea cual sea su motivación para ir al entorno rural, o los que sólo son atractivos para los turistas que tienen una motivación específica.

Conclusiones

La elección de alojamiento es una de las más importantes decisiones a la que se enfrenta un turista cuando está planeando un viaje (Sharpley, 2000; Chen et al., 2017), no solo por el coste económico, sino, porque el alojamiento constituye un elemento esencial para determinar la satisfacción y relevancia de la experiencia turística vivida por el turista y puede influir en su comportamiento futuro (Mcintosh y Siggs, 2005; Liu y Ma, 2017). Es, por tanto, crucial conocer y entender la elección de alojamiento, ya que esto puede resultar de gran ayuda para los gestores y propietarios de alojamientos en relación a conseguir una gestión más eficaz de los inmuebles y a realizar una política de inversiones adecuada (Kim y Park, 2017). Así, el propósito de este trabajo ha sido conocer las preferencias de los turistas en cuanto a los alojamientos rurales, el valor que le otorgan a las características y a los servicios que ofertan los alojamientos, así como la influencia que en estas elecciones tienen las características, sentimientos y motivaciones de los turistas. Para lograr este objetivo, hemos considerado el alojamiento rural como un producto caracterizado por diferentes atributos físicos existentes o potenciales, que le dan un valor y establecen una interconexión con el comportamiento del decisor, así como con sus sentimientos y motivaciones.

Para cuantificar la influencia que los atributos de los alojamientos y las características (observables o no) de los turistas tienen en la elección de alojamiento, las herramientas metodológicas empleadas han sido los Modelos de Elección Discreta, desde el modelo más simple y popular, el modelo logit multinomial, hasta modelos más complejos que incorporan algunos elementos del proceso cognitivo identificados como importantes para el proceso de elección, como son el modelo logit mixto (Train, 2003) y los modelos híbridos de elección discreta (Ben-Akiva, McFadden et al., 2002). Además, de profundizar en el conocimiento y utilización de modelos más ricos para describir el comportamiento humano, también hemos enriquecido la información usada para su estimación, utilizando diferentes fuentes de datos, los datos de preferencias reveladas (PR) y preferencias declaradas (SP), que difieren tanto

en su naturaleza como en el procedimiento de obtención. Los datos de PR proporcionan información sobre la elección real de alojamiento, mientras que los datos SP recogen lo que los turistas dicen que harían ante escenarios de elección hipotéticos, cuidadosamente diseñados por el investigador aplicando modelos estadísticos del diseño de experimentos. Por último, señalar que los datos RP y SP sobre la elección de alojamiento rural, las valoraciones sobre diversos indicadores motivacionales para pasar un periodo vacacional en el entorno rural y las características socio-económicas y de comportamiento de viaje de los turistas usadas en las tres aplicaciones prácticas de esta memoria, forman parte de una encuesta diseñada ad-hoc con el fin de determinar las preferencias de elección de los turistas que visitaban la Comarca del Noroeste de la Región de Murcia y fue realizada durante varios fines de semana en otoño de 2003.

El Capítulo 1 aglutina todo lo que es común en las aplicaciones prácticas propuestas y desarrolladas en los otros tres capítulos. Esto es, hemos comenzado con una breve descripción del surgimiento y evolución del turismo rural, haciendo hincapié en el tema de interés que son los alojamientos rurales. A continuación, hemos presentado los aspectos más relevantes relativos a la modelización de la elección de alojamiento rural basada en el empleo de modelos de elección discreta, analizando tanto sus fundamentos micro-económicos como econométricos, prestando especial atención al modelo logit multinomial, que es el más sencillo, pero constituye la base de los modelos que posteriormente se utilizarán en este trabajo. También hemos incluido en este capítulo la definición, ventajas y diferencias de las dos fuentes de datos empleadas para estimar estos modelos, en especial la modelización con preferencias declaradas. Los datos SP se obtuvieron mediante un experimento de elección diseñado y elaborado ad-hoc para los estudios incluidos en esta memoria. Este capítulo se cierra con la explicación del experimento de elección realizado junto con los datos sobre las elecciones reales de alojamiento rural (datos de preferencias reveladas), las características socio-económicas de los turistas y las condiciones de viaje y las valoraciones sobre una serie de indicadores para detectar la motivación que mueve a estos turistas a desplazarse al entorno rural.

En el Capítulo 2 hemos comprobado que el perfil del turista varía según el tipo de alojamiento que existe en un determinado destino de turismo rural. Para detectar qué factores (características socio-económicas y condiciones de viaje) de los turistas son relevantes en la elección de un alojamiento rural, hemos estimado un modelo logit multinomial con los datos de preferencias reveladas (PR) obtenidos en la encuesta. En nuestro contexto, las alternativas

que forman parte del conjunto de elección debían ser los distintos alojamientos situados en un destino rural, los cuales diferían entre sí en cuanto al número de habitaciones, número de camas y de baños, el tipo de construcción o de alquiler, la ubicación, las instalaciones, etc. No obstante, aunque la oferta de establecimientos era muy variada, fue posible clasificar todos los alojamientos de un determinado destino en grupos homogéneos y así reducir el tamaño del conjunto de elección del modelo logit multinomial propuesto. De esta forma, cada uno de los grupos obtenidos constituyó un tipo de alojamiento. Así, una pieza clave en el estudio propuesto en este capítulo ha sido la construcción del conjunto de elección. Para ello, hemos propuesto dos métodos de agrupación adecuados para trabajar con datos cualitativos, como son los atributos de los alojamientos. El primer método, más clásico, es un análisis cluster combinado con un análisis de correspondencias múltiple propuesto por Arimond y Elfessi (2001) y el segundo, más actual, es un análisis de clases latentes (Lazarsfeld, 1950; Agresti, 2003). Ambas técnicas de agrupación requieren de la especificación a priori del número de grupos, pero el análisis de clases latentes tiene la ventaja de que existen herramientas para evaluar la bondad del ajuste del modelo y también para determinar el número apropiado de clases latentes dado un conjunto de datos.

A pesar de las dificultades que implican las técnicas basadas en variables cualitativas, el estudio efectuado ha demostrado su utilidad para clasificar los alojamientos, que existen en un destino turístico rural, en grupos homogéneos o tipologías de alojamientos. En concreto, con ambas técnicas implementadas hemos agrupado los alojamientos rurales ubicados en la Comarca del Noroeste de la Región de Murcia en cuatro tipos o clases. Aunque las agrupaciones realizadas con cada técnica no contienen exactamente las mismas casas, sí han coincidido en su mayoría. Así, los alojamientos del grupo 1 han resultado ser establecimientos generalmente grandes, ubicados en el campo o en la montaña y con piscina. Los alojamientos pertenecientes al grupo 2 son de tamaño medio (dos o tres habitaciones), situados en zonas agrícolas y que poseen equipamientos adicionales como parque infantil e instalaciones deportivas. También disponen de servicio de comidas y piscina. Los alojamientos del grupo 3, que es el grupo más numeroso, son también de tamaño medio, mayoritariamente de dos habitaciones, con una arquitectura típica de la zona y situados en aldeas o en los alrededores de un pueblo y que carecen de otros equipamientos. Por último, los establecimientos del grupo 4 incluyen los alojamientos de tamaño más pequeño (1 habitación) o si son grandes, aquellos que no se alquilan al completo, sino por habitaciones individuales con baño privado.

Una vez clasificados los alojamientos, en una segunda fase del estudio, hemos estimado un modelo Logit Multinomial con el objetivo de determinar las características de los turistas, que pasan sus vacaciones en ese destino, que pudieran ser determinantes en la elección de cada tipo de alojamiento. Así, hemos detectado que los alojamientos del grupo 1 han resultado ser preferidos por aquellos turistas que viven en pareja sin hijos, que suelen viajar con amigos, que prefieren alquilar directamente al propietario y que buscan información del lugar a través de internet y guías turísticas. Los alojamientos pertenecientes al grupo 2 son los elegidos por aquellos turistas que realizan un gasto medio-alto en el destino y buscan realizar actividades deportivas o de senderismo o visitar lugares de interés cercanos a la vivienda. Los turistas con más de 40 años que prefieren realizar actividades culturales optan por los alojamientos del grupo 3. Por último, los establecimientos del grupo 4 son el objetivo de los turistas mayores de 40 años, con ingresos altos y que priorizan las actividades culturales.

Los resultados de este primer estudio, aunque ha sido realizado en una zona y en un periodo de tiempo concreto, ponen de manifiesto que existen diferencias significativas entre los turistas que eligen cada tipo de alojamiento en un destino. Esta información puede ser útil para los gestores de alojamientos rurales para orientar sus políticas de inversión y marketing para atraer perfiles de turistas concretos. Se trataría, por tanto, de tener en cuenta aquellas variables que proporcionan una mayor probabilidad de elegir ese tipo de alojamiento de turístico para saber a qué turistas dirigir sus promociones y cómo satisfacer de una manera óptima sus preferencias de viaje y actividades.

En el Capítulo 3 hemos presentado una metodología que ha resultado ser una herramienta eficaz y útil para evaluar las preferencias de los turistas sobre las instalaciones y los servicios que ofrece cada establecimiento. Además, con los resultados obtenidos hemos probado que, aunque la elección de un alojamiento rural tradicionalmente se haya asociado a su ubicación en un entorno natural y a características intrínsecamente rurales, existen otros factores físicos de los alojamientos que también son importantes en su elección, como el tamaño, el tipo de edificación, la calidad del equipamiento y los servicios y actividades que ofrece. También hemos detectado preferencias heterogéneas de los turistas respecto a estos atributos.

Para identificar las preferencias de los turistas sobre los atributos físicos, potenciales o efectivos, de las casas rurales situadas en la Comarca del Noroeste de la Región de Murcia, la metodología propuesta ha sido la estimación, con datos obtenidos del experimento de preferencias declaradas, de

varios modelos de elección discreta, todos ellos pertenecientes a la familia de los modelos “logit”. Este experimento de elección de alojamiento rural se contextualizó como un problema en el que los turistas tenían que elegir entre pasar un fin de semana en uno de los dos alojamientos rurales hipotéticos propuestos o la opción de “quedarse en casa”. Cada alojamiento fue descrito en términos de diferentes niveles de catorce atributos seleccionados de acuerdo con las características, equipamiento y servicios que los alojamientos rurales situados en las áreas donde el turismo rural tiene una larga tradición en España. Los conjuntos de elección (de tamaño tres) y las alternativas contenidas en ellos fueron creados simultáneamente usando técnicas del diseño de experimentos. En concreto, se aplicó un diseño factorial fraccional de efectos principales de resolución III para seleccionar 64 conjuntos de elección, que posteriormente fueron distribuidos en 16 grupos con 4 escenarios cada uno, usando un diseño de bloque D-eficiente. Así, se redujo el número de escenarios a examinar por cada encuestado a cuatro, asegurando, además, la reducción de los errores estándar de los parámetros estimados.

Con los datos de preferencias declaradas obtenidos con la implementación del diseño de elección descrito en el párrafo anterior e incluido en la encuesta efectuada en la Comarca Noroeste de la Región de Murcia, hemos estimado cuatro especificaciones logits, siendo una de ellas un modelo logit homogéneo. Este modelo presentaba coeficientes constantes de los atributos de los alojamientos a través de los individuos, y se tomó de referencia para compararlo con las otras tres especificaciones. Las estimaciones realizadas abalan la “calidad” de los datos de preferencias declaradas frente a los de preferencias reveladas en la detección de los atributos, físicos y tangibles, de los alojamientos que resultan ser determinantes en la elección de alojamiento rural. Pero también permiten contrastar el potencial de otros modelos logit, más ricos desde el punto de vista comportamental, aunque más complejos, como son el heterogéneo y el mixto, para descubrir variaciones en los gustos, sistemáticas y aleatorias, dentro de las preferencias de los turistas. Con el logit heterogéneo, hemos modelizado las variaciones del gusto sistemáticas, incluyendo en la función de utilidad interacciones de características socio-económicas y de comportamiento de viaje de los turistas con los niveles de los atributos de las casas rurales, mientras que las variaciones del gusto aleatorias las hemos estudiado usando especificaciones del modelo logit mixto, ya que permiten que los parámetros varíen aleatoriamente entre los turistas.

Los resultados de los modelos estimados muestran que el atractivo de una casa rural para los turistas rurales de la Región de Murcia mejora – en diferentes grados – si es una casa de construcción tradicional que ha sido re-

modelada para ser usada para el turismo, localizada en el campo o montaña, con un amplio número de habitaciones, que dispone de la certificación de calidad “Q” y ofrece la posibilidad de alquilar caballos. Hay otros atributos que también ejercen un efecto positivo sobre el atractivo de las casas rurales, pero sólo para algunos segmentos de turistas. Por ejemplo, la oferta de una minigranja podría ser atractiva para aquellos turistas que viajan con niños, o la posibilidad de reservar una casa por habitaciones puede atraer a los turistas que hacen turismo rural con frecuencia. Por último, otras características de las casas pueden tener tanto un efecto positivo como negativo para su elección, como son el compartir el cuarto de baño, ser una casa de nueva construcción, estar situada en un área de “huerta” o si la casa se alquila por habitaciones o al completo.

La información obtenida con este segundo estudio puede ser valiosa para los planificadores y administradores de establecimientos rurales, ya que les puede ayudar a un desarrollo mayor de su actividad económica y en la planificación de estrategias de marketing y el diseño de campañas de publicidad. La comparación de los atributos de un alojamiento rural existente con aquellos preferidos por los turistas pueden mejorar el conocimiento sobre la casa y proporcionar la base para promociones específicas. Por ejemplo, los gerentes de una casa rural pueden decidir tomar las medidas necesarias para conseguir ganar la certificación de calidad “Q”, o simplemente ofrecer la posibilidad de montar a caballo (a través de establos locales) como medio para hacer su establecimiento más atractivo. También pueden proponer ofertas de alquilar la casa en su totalidad o por habitaciones para atraer a los turistas frecuentes, o la creación de una mini-granja como un imán para los turistas que viajan con niños. El conocimiento de las preferencias del consumidor para los atributos como el tipo de construcción, la situación o el tamaño, que no son fáciles de modificar, es también una herramienta importante para el desarrollo del turismo rural, especialmente en el caso de las decisiones de inversión en nuevas casas rurales.

En el Capítulo 4, tomando como base la teoría del proceso de elección constructiva del consumidor propuesta por Bettman et al. (1998) y teniendo presente la premisa de que no todos los turistas que acceden al medio rural tienen las mismas motivaciones o buscan las mismas satisfacciones (Frochot, 2005; Molera y Albaladejo, 2007; Kastenholz et al., 1999; entre otros), hemos probado que las motivaciones o intereses que los turistas buscan cuando se desplazan al entorno rural son decisivos en la elección de su alojamiento y en la determinación de los atributos relevantes de un alojamiento. Además, también hemos constatado que el efecto de las motivaciones difiere en fun-

ción de los atributos del alojamiento, mostrando preferencias heterogéneas de los turistas respecto de dichos atributos. Por último, hemos conseguido extender la hipótesis de Kim y Park (2017) de que el contexto de elección influye en la decisión de elegir un hotel a la de elección de un alojamiento en el entorno rural, al verificar que las motivaciones implican ciertas restricciones o limitaciones a la hora de elegir el alojamiento rural. Por tanto, los resultados testifican que el contexto de decisión es decisivo en la elección del alojamiento turístico, ya sea un hotel (Kim y Park, 2017) o un alojamiento rural (Albaladejo y Díaz, 2020).

Las motivaciones, a diferencia de las variables cognitivas, son variables latentes, esto es, no observables directamente, multidimensionales y muy difíciles de medir. Así, la consideración de que las motivaciones pudieran influir en la elección de alojamiento rural ha supuesto una dificultad añadida respecto a la utilización de únicamente variables cognitivas, y nos ha llevado a la utilización de los modelos híbridos de elección discreta (HDC). Esto se debe a la incapacidad que presentan los modelos de elección discreta tradicionales para trabajar con variables latentes. Los modelos HDC surgen de la combinación de los modelos de elección discreta, muy útiles en la determinación de los atributos observables que son relevantes en la elección de alojamiento rural, tal y como se ha visto en el capítulo 3, pero que no permiten la inclusión de variables latentes, y los modelos de ecuaciones estructurales, una técnica estadística multivariante utilizada para explicar las relaciones causales entre variables latentes que, a su vez, se miden a través de su relación con otras variables observadas directamente denominadas variables indicadores o manifiestas. De esta manera, el empleo de los modelos híbridos de elección discreta nos ha permitido la inclusión tanto de variables observables (atributos de los alojamientos) como latentes (motivaciones) en la formalización de la elección de un alojamiento rural. En concreto, hemos empleado un modelo HDC que incorpora el efecto de las variables latentes a través de un modelo de ecuaciones estructurales denominado modelo de múltiples indicadores y múltiples causas (MIMIC) (Ben-Akiva, Walker et al., 2002; Bolduc y Daziano, 2010; Bahamonde-Birke et al., 2010; Bahamonde-Birke et al., 2017; Márquez et al., 2020, entre otros), que vincula las variables latentes, por un lado, a los indicadores motivacionales a través de las ecuaciones de medida y, por otro, con las características socio-económicas de los turistas y sus condiciones de viaje por medio de las ecuaciones estructurales.

La estimación del modelo HDC ha sido secuencial o en dos etapas. En una primera etapa hemos estimado las variables latentes usando el modelo MIMIC propuesto, previamente validado, y en una segunda fase, hemos añadido los

valores esperados de las variables latentes, detectadas en la primera fase, al conjunto de variables explicativas observables del modelo de elección discreta para la estimación de los parámetros asociados a unas y otras variables de la forma estándar. Así, con la estimación del modelo MIMIC hemos detectado la existencia de cuatro variables latentes que identifican las motivaciones de los turistas rurales que visitan la Comarca del Noroeste, denominadas *Medio ambiente*, *Individualidad*, *Familia* y *Sociabilidad*. La variable latente *Medio Ambiente* está vinculada con características propiamente del entorno rural, como son el contacto con la naturaleza, ambiente tranquilo, belleza del entorno, ausencia de masificación y ambiente rústico. Así, niveles altos en estos indicadores se deben a que la principal motivación de estos turistas para ir al entorno rural es el *Medio ambiente*. Además, según el modelo MIMIC, un turista que viaja más de 5 veces al año, en grupos pequeños (5 personas máximo) y con familia e hijos, cuya residencia habitual está fuera de la Región de Murcia y carece de estudios universitarios se relaciona con niveles elevados en esta variable latente. Por otro lado, la variable latente *Individualidad*, que alude a la búsqueda de un alojamiento rural situado cerca de la vivienda habitual del individuo, con un precio razonable, que le permita descansar y que le proporcione una sensación de independencia y flexibilidad, se explica por el lugar de residencia, el nivel de formación alcanzado y el tamaño del grupo de viaje. En consecuencia, puntuaciones altas en la variable latente *Individualidad* puede ser causadas por vivir fuera de la Región de Murcia, viajar en grupos pequeños y tener educación primaria, como máximo nivel de formación alcanzado. La variable latente *Familia* está relacionada con los individuos que están interesados principalmente en pasar un buen rato con la familia y la posibilidad de viajar con niños. Esta variable está explicada principalmente por la frecuencia y los compañeros de viaje. Niveles altos en la variable latente *Familia* pueden ser causados por viajar más de 5 veces al año, con la familia e hijos, mientras que los niveles bajos se asocian con viajar con amigos, ser soltero y tener estudios superiores. Por último, niveles elevados en la variable latente *Sociabilidad*, asociada a personas que ven en elegir un alojamiento rural una oportunidad para conocer gente y estar en contacto con los residentes locales, están vinculados con estar ocupado, tener una edad comprendida entre 26 y 40 años, con un presupuesto diario de más 60 euros y con una frecuencia anual de viaje superior a 5 veces.

En una segunda fase, y continuando con nuestro objetivo de contrastar si la motivación latente que tiene un turista para visitar el entorno rural puede influir en la elección de alojamiento rural, hemos calculado, en primer lugar, los valores esperados de las variables latentes detectadas para cada uno de los individuos de la muestra y, posteriormente, hemos estimado tres modelos

de elección discreta con los datos obtenidos sobre los turistas que visitaron el Noroeste de la Región de Murcia: un MNL y dos especificaciones del modelo HDC. En el modelo MNL únicamente se consideraron como variables explicativas los atributos objetivos y tangibles de los alojamientos (tamaño, ubicación, precio, entre otros), mientras que en las dos especificaciones del modelo HDC, además de incorporar las variables anteriores, también incluyeron las cuatro motivaciones latentes detectadas. En la primera especificación su inserción fue en forma lineal, mientras que en la segunda se introdujeron linealmente y también como interacciones con los atributos de los alojamientos rurales. Los resultados obtenidos indican que los dos modelos HDC ofrecen mejoras significativas en la bondad del ajuste respecto al modelo MNL y evidencian que las motivaciones pueden afectar a la probabilidad de elección de alojamiento rural. En concreto, observamos que los turistas con una mayor motivación por socializarse, tanto con residentes locales como con otros turistas, tienen una mayor probabilidad de elegir un alojamiento rural en la Región de Murcia que aquellos turistas motivados por pasar tiempo con la familia o motivados por hacer un viaje corto, tranquilo e individual. La segunda especificación del modelo HDC nos ha permitido detectar preferencias heterogéneas entre los turistas en cuanto a los atributos de los alojamientos rurales. Así, por ejemplo, el hecho de que el alojamiento tenga una mini-granja influye positivamente en la elección del alojamiento para los turistas motivados por viajar con la familia y los niños, pero influye negativamente en los turistas que buscan hacer un viaje corto, tranquilo e individual.

Por último, los resultados de este tercer estudio proporcionan también información sobre los atributos de los alojamientos que son apreciados por los turistas, aquellos que no son valorados por ningún turista sea cual sea su motivación para ir al entorno rural, o los que sólo son atractivos para los turistas que tienen una motivación específica. No obstante, hay que señalar que, aunque modelos HDC proporcionan mejores ajustes y son más explicativos, su proceso de estimación es más complejo y requiere de mucha más información que del modelo MNL. Así que, de acuerdo con Palma et al. (2018) y Vij y Walker (2016), los modelos HDC deben usarse sólo para estudios en un contexto particular. Tal es el caso de nuestro estudio, que tenía como objetivo hacer un análisis de la elección del alojamiento rural, proponiendo que la motivación los objetivos de los turistas también influyen en sus elecciones.

En resumen, los resultados derivados de esta memoria muestran que:

1. Los alojamientos en un destino turístico se pueden clasificar en grupos homogéneos o tipologías de alojamientos y existen diferencias significativas entre los turistas que eligen cada una de las diferentes tipologías.

2. Los datos de preferencias declaradas, a pesar de que su obtención no es tan directa como los datos de preferencias reveladas, ya que hay que recurrir a técnicas estadísticas del diseño de experimentos, han resultado ser más idóneos para estimar los modelos de elección discreta.
3. Los atributos físicos y tangibles de los alojamientos rurales y las actividades y servicios que ofertan son determinantes para su elección.
4. Los modelos de elección discreta con datos de preferencias declaradas son viables, desde una perspectiva conceptual y práctica, para:
 - la detección de los factores determinantes en la elección de alojamiento rural y,
 - la evaluación de las preferencias de los turistas sobre las instalaciones y los servicios que ofrece cada establecimiento.
5. Los turistas tienen preferencia por ciertos atributos de un alojamiento, aunque la percepción y evaluación de estos atributos no es igual para todos los individuos. Existen segmentos de turistas con diferentes preferencias, pero el comportamiento diferente de los turistas también puede tener un carácter aleatorio.
6. La elección del alojamiento rural varía en función de la motivación que tiene el turista para ir al entorno rural (Medio Ambiente, Individualidad, Familia o Sociabilidad). El contexto de decisión también condiciona las preferencias en relación con los atributos del alojamiento.
7. El contexto de decisión es determinante en la elección de un alojamiento turístico rural. Este resultado extiende la propuesta de Kim y Park (2017) para la elección de hotel.
8. Los modelos híbridos de elección son una herramienta muy útil, aunque no sencilla, para analizar el impacto de factores decisivos, y no directamente observables, en la elección de alojamiento, tales como las motivaciones de los turistas.

Por consiguiente, las diferentes metodologías desarrolladas en esta memoria han ido dirigidas a la detección de factores observables o no observables que pueden ser relevantes en la elección de alojamiento rural y si estos factores son importantes para todos los individuos o sólo para un determinado grupo. Estas herramientas han resultado útiles para determinar las preferencias de los turistas por las instalaciones y servicios que ofrece un alojamiento rural.

Bibliografía

- Adamowicz, W., Louviere, J., y Williams, M. (1994). Combining revealed and stated preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management*, 26:271–294.
- Adamowicz, W., Swait, J., Boxall, P., Louviere, J., y Williams, M. (1997). Perceptions versus objective measures of environmental quality in combined revealed y stated preference models of environmental evaluation. *Journal of Environmental Economics and Management*, 32:65–84.
- Agresti, A. (2003). *Categorical data analysis*, volume 482. John Wiley and Sons.
- Aguiló, E. y Juaneda, C. (2000). Tourist expenditure for mass tourism markets. *Annals of Tourism Research*, Curso Avanzado de Turismo en Áreas Rurales: Estrategias y Promoción, (3):624 – 637.
- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Second International Symposium on Information Theory*, pages 267–281. Akademiai Kiado, Budapest, Hungary.
- Albaladejo, I. P. y Díaz Delfa, M. T. (2003). Un modelo de elección discreta en la determinación del perfil del turista rural: Una aplicación en Murcia. *Cuadernos de Turismo*, (11):7–20.
- Albaladejo, I. P. y Díaz Delfa, M. T. (2005). Rural tourism demand by type of accommodation. *Tourism Management*, 26:951–959.
- Albaladejo, I. P. y Díaz Delfa, M. T. (2007). *Tourism management in the 21st century*, chapter Stated preferences of tourists for rural accommodation choice, pages 179–203. Nova Science Publishers, New York.
- Albaladejo, I. P. y Díaz Delfa, M. T. (2009). Tourist preferences for rural house stays: Evidence from discrete choice modelling in Spain. *Tourism Management*, 30(6):805–811.

- Albaladejo, I.P. y Díaz Delfa, M. T. (2020). The effects of motivations to go to the country on rural accommodation choice: A hybrid discrete choice model. *Tourism Economics*, <https://doi.org/10.1177/1354816620912062>.
- Aldas Manzano, J. y Uriel Jimenez, E. (2017). *Análisis multivariante aplicado con R*. Ediciones Paraninfo, SA.
- Aldenderfer, M. S. y Blashfield, R. K. (1991). *Cluster analalysis*. Series: Quantitative Applications in the Social Sciences. A Sage University Paper, University of Iowa.
- Alfnes, F. (2004). Stated preferences for imported and hormone-treated beef: application of a mixed logit model. *European Review of Agricultural Economics*, 31:19–37.
- Apostolakis, A. y Jaffry, S. (2005a). A Choice Modeling Application for Greek Heritage Attractions. *Journal of travel research*, 43(3):309–318.
- Apostolakis, A. y Jaffry, S. (2005b). Stated preference for two Cretan heritage attractions. *Annals of Tourism Research*, 32(4):985–1005.
- Araña, J. E., León, C. J., Carballo, M. M., y Moreno Gil, S. (2016). Designing tourist information offices: The role of the human factor. *Journal of Travel Research*, 55(6):764–773.
- Arimond, G. y Elfessi, A. (2001). A clustering method for categorical data in tourism market segmentation research. *Journal of Travel Research*, 39:91–98.
- Ashok, K., Dillon, W. R., y Yuan, S. (2002). Extending discrete choice models to incorporate attitudinal and other latent variables. *Journal of marketing research*, 39(1):31–46.
- Asparouhov, T. y Muthén, B. (2010). Simple second order chi-square correction. *Mplus technical appendix*, pages 1–8.
- Bagozzi, R. P. y Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the academy of marketing science*, 16(1):74–94.
- Bahamonde-Birke, F., Kunert, U., Link, H., y Ortúzar, J. d. D. (2017). About attitudes and perceptions: Finding the proper way to consider latent variables in discrete choice models. *Transportation*, 44(3):475–493.

- Bahamonde-Birke, F., Raveau, S., Yáñez, F., y Ortúzar, J. d. D. (2010). The role of tangible attributes in hybrid discrete choice models. In *Association for European Transport and contributors*.
- Bahamonde-Birke, F. J. y Ortúzar, J. d. D. (2014). On the variability of hybrid discrete choice models. *Transportmetrica A: Transport Science*, 10(1):74–88.
- Barke, M. (2004). Rural tourism in Spain. *International Journal of Tourism Research*, 6:137–149.
- Bates, J. (1988). Econometric issues in stated preference analysis. *Journal of Transport Economics and Policy*, 22(1):59–69.
- Ben-Akiva, M. y Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel demand*. MIT Press, Cambridge, Massachusetts.
- Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D., Daly, A., de Palma, A., Gopinath, D., Karlström, A., y Munizaga, M. (2002a). Hybrid choice models: Progress and challenges. *Marketing Letters*, 13:163–175.
- Ben-Akiva, M. y Morikawa, T. (1991). Estimation of travel demand models from multiple data sources. In Koshi, M., editor, *Transportation and Traffic Theory*, pages 461–476, Amsterdam. Elsevier. Proceedings of the 11th ISTTT.
- Ben-Akiva, M., Walker, J., Bernardino, A., Gopinath, D., Morikawa, T., y Polydoropoulou, A. (2002b). Integration of choice and latent variable models. *Perpetual Motion: Travel Behaviour Research Opportunities and Application Challenges*.
- Ben-Akiva, M., Walker, J., Bernardino, A. T., Gopinath, D. A., Morikawa, T., y Polydoropoulou, A. (2002c). Integration of choice and latent variable models. *Perpetual motion: Travel behaviour research opportunities and application challenges*, pages 431–470.
- Bettman, J. R., Luce, M. F., y Payne, J. W. (1998). Constructive consumer choice processes. *Journal of consumer research*, 25(3):187–217.
- Bhat, C. (1999). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. Working Paper, Department of Civil Engineering, University of Texas, Austin.

- Bhat, C. R. (1995). A heteroscedastic extreme value model of intercity travel mode choice. *Transportation Research B*, 29(6):471–483.
- Björk, P. y Jansson, T. (2008). Travel decision-making: The role of habit. *Tourismos: An International Multidisciplinary Journal of Tourism*, 3(2):11–34.
- Bolduc, D. y Daziano, R. A. (2010). On estimation of hybrid choice models. In Hess, S. and Daly, A., editors, *Choice Modelling: the state-of-the-art and the state-of-practice*, pages 259–287. Emerald.
- Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley and Sons, Chichester.
- Bolt, D. M. (2005). Limited-and full-information estimation of item response theory models. In Maydeu-Olivares and McArdle, J. J., editors, *Contemporary psychometrics: A festschrift for Roderick P. McDonald*, pages 27–71. Lawrence Erlbaum Associates Publishers.
- Boxall, P. C., Adamowicz, W. L., Swait, J., Williams, M., y Louviere, J. J. (1996). A comparison of stated preference methods for environmental valuation. *Ecological Economics*, 18:243–253.
- Bradley, M. A. y Kroes, E. (1990). Forecasting issues in stated preference survey research. In *Selected readings in transport survey methodology. Edited proceedings of the 3rd international conference on survey methods in transportation*, Washington DC, EE.UU. 69th TBR Annual Meeting.
- Brey, R., Riera, P., y Mogas, J. (2007). Estimation of forest values using choice modeling: an application to Spanish forests. *Ecological Economics*, 64:305–312.
- Brownstone, D., Bunch, D. S., y Train, K. (2000). Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research Part B: Methodological*, 34(5):315–338.
- Brownstone, D. y Train, K. (1999). Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, 89:109–29.
- Cánoves, G., Herrera, L., y Villarino, M. (2005). Turismo rural en España: Paisajes y usuarios, nuevos usos y nuevas visiones. *Cuadernos de Turismo*, 15:63–76.
- Cánoves, G., Villarino, M., Priestley, G. K., y Blanco, A. (2004). Rural tourism in Spain: an analysis of recent evolution. *Geoforum*, 35:755–769.

- Carson, R., Louviere, J. J., Anderson, D. A., Arabie, P., Bunch, D. S., Hensher, D. A., Johnson, R. M., Kuhfeld, W. F., Steinberg, D., Swait, J., Timmermans, H., y Wiley, J. B. (1994). Experimental analysis of choice. *Marketing Letters*, 5(4):351–368.
- Castro López, C. R. y Tenorio Aguirre, A. J. (2010). Modelos de clases latentes aplicados a las encuestas de percepción ciudadana: estudio de caso. *Revista legislativa de estudios sociales y de opinión pública*, 3(6):187–200.
- Charrad, M., Ghazzali, N., Boiteau, V., y Niknafs, A. (2014). Nbclust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software, Articles*, 61(6):1–36.
- Chen, C.-M., Tsai, Y.-C., y Chiu, H.-H. (2017). The decision-making process of and the decisive factors in accommodation choice. *Current Issues in Tourism*, 20(2):111–119.
- Costa, P. y Manente, M. (1995). Venice and its visitors: A survey and a model of qualitative choice. *Journal of travel Research*, 43(2):118–130.
- Croissant, Y. (2020). Estimation of random utility models in R: The mlogit package. *Journal of Statistical Software, Articles*, 95(11):1–41.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3):297–334.
- Crouch, G. I. y Louviere, J. J. (2000). A review of choice modeling research in tourism, hospitality, and leisure. *Tourism Analysis*, 5(2-3):97–104.
- Crouch, G. I. y Louviere, J. J. (2004). The determinants of convention site selection: a logistic choice model from experimental data. *Journal of Travel Research*, 43(2):118–130.
- Daly, A., Hess, S., Patruni, B., Potoglou, D., y Rohr, C. (2012). Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour. *Transportation*, 39(2):267–297.
- Dávila, D., Rodríguez, S., y Rodríguez, A. (2002). La promoción turística espacial dentro de un destino vacacional homogéneo. una aplicación a Canarias. *Estudios de Economía Aplicada*, 20(1):45–60.
- Daziano, R. A. y Bolduc, D. (2013). Incorporating pro-environmental preferences towards green automobile technologies through a bayesian hybrid choice model. *Transportmetrica A: Transport Science*, 9(1):74–106.

- Dernoi, L. A. (1991). About rural and farm tourism. *Tourism Recreation Research*, 16(1):3–6.
- Domencich, T. A. y McFadden, D. (1975). *Urban travel demand: A behavioral analysis*. North-Holland Publishing Company Limited, Oxford, England.
- Epstein, S. (2003). Cognitive-experiential self-theory of personality. *Handbook of psychology*, pages 159–184.
- Espino, R. (2003). *Análisis y predicción de la demanda de transporte de pasajeros: Una aplicación al estudio de dos corredores de transporte en Gran Canaria*. PhD thesis, Departamento de Análisis Económico, Universidad de las Palmas de Gran Canaria.
- Espino, R., Martín, J. C., y Román, C. (2008). Analyzing the effect of preference heterogeneity on willingness to pay for improving service quality in an airline choice context. *Transportation Research Part E: Logistics and Transportation Review*, 44(4):593–606.
- Espino, R., Ortúzar, J., y Román, C. (2004). Diseño de preferencias declaradas para analizar la demanda de viajes. *Estudios de Economía Aplicada*, 22:725–758.
- Espino, R., Román, C., y Ortúzar, J. (2006). Analysing demand for suburban trips: A mixed rp/sp model with latent variables and interaction effects. *Transportation*, 33:241–261.
- Eugenio-Martin, J. L. (2003). Modelling determinants of tourism demand as a five-stage process: A discrete choice methodological approach. *Tourism and Hospitality Research*, 4(4):341–354.
- Eymann, A. y Ronning, G. (1997). Microeconomic models of tourists' destination choice. *Regional Science and Urban Economics*, 27(6):735–761.
- Eymann, A., Ronning, G., y Zimmermann, K. F. (1992). Discrete choice analysis of foreign travel demand. In *European integration in the world economy*, pages 641–674. Springer.
- Fleischer, A. y Felsenstein, D. (2000). Support for rural tourism. does it make a difference? *Annals of Tourism Research*, 27(4):1007–1024.
- Fornell, C. y Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, 18(1):39–50.

- Forster, M. R. (2000). Key concepts in model selection: Performance and generalizability. *Journal of mathematical psychology*, 44(1):205–231.
- Frochot, I. (2005). A benefit segmentation of tourists in rural areas: a Scottish perspective. *Tourism Management*, 26(3):335–346.
- Frochot, I. y Morrison, A. M. (2000). Benefit segmentation: A review of its applications to travel and tourism research. *Journal of Travel and Tourism Marketing*, 9(4):21–45.
- Fuentes, R. (1995). Análisis de las principales características de la demanda de turismo rural en España. *Estudios Turísticos*, 127:19–52.
- Gannon, A. (1994). Rural tourism as a factor in rural community economic development for economies in transition. *Journal of Sustainable Tourism*, 2:51–59.
- Gössling, S. y Lane, B. (2015). Rural tourism and the development of internet-based accommodation booking platforms: a study in the advantages, dangers and implications of innovation. *Journal of Sustainable Tourism*, 23(8-9):1386–1403.
- Green, P. E. (1984). Hybrid models for conjoint analysis: An expository review. *Journal of Marketing Research*, 21(2):155–169.
- Green, S. B., Akey, T. M., Fleming, K. K., Hershberger, S. L., y Marquis, J. G. (1997). Effect of the number of scale points on chi-square fit indices in confirmatory factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 4(2):108–120.
- Greenacre, M. J. (1984). *Theory and application of correspondence analysis*. Academic Press, London.
- Greene, W. y Hensher, D. (2003). A latent class model for discrete choice analysis: Contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37:681–698.
- Greene, W. H. (1999). *Análisis Econométrico*. Prentice Hall Iberia, Madrid.
- Greene, W. H. (2002). *Limdep version 8.0*. Econometric Software, Inc., Australia.
- Greene, W. H. (2012). *Nlogit version 5.0 Reference Guide*. Econometric Software, Inc., Australia.

- Hensher, D. A. y Bradley, M. (1993). Using stated response choice data to enrich revealed preference discrete choice models. *Marketing Letters*, 4(2):139–151.
- Hensher, D. A., Rose, J. R., y Greene, W. H. (2005). *Applied choice analysis*. Cambridge University Press, Cambridge.
- Hu, L.-t. H. y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1):1–55.
- Huber, J. y Zwerina, K. (1996). The importance of utility balance in efficient choice designs. *Journal of Marketing Research*, 33:307–17.
- Hummelbrunner, R. y Miglbauer, E. (1994). Tourism promotion and potential in peripheral areas: the Austrian case. *Journal of Sustainable Tourism*, 2:41–50.
- Huybers, T. (2003). Modelling short-break holiday destination choices. *Tourism Economics*, 9(4):389–405.
- Huybers, T. y Bennett, J. (2000). Impact of the environment on holiday destination choices of prospective UK tourists: implications for Tropical North Queensland. *Tourism Economics*, 6:21–46.
- Jang, S., Morrison, A. M., y O’Leary, J. (2002). Benefit segmentation of Japanese pleasure travelers to the USA and Canada: selecting target markets based on the profitability and risk of individual market segments. *Tourism Management*, 23:367–378.
- Jöreskog, K. G. (1994). Structural equation modeling with ordinal variables. *Lecture Notes-Monograph Series*, pages 297–310.
- Jöreskog, K. G. y Moustaki, I. (2001). Factor analysis of ordinal variables: A comparison of three approaches. *Multivariate Behavioral Research*, 36(3):347–387.
- Kastenholz, E., Davis, D., y Paul, G. (1999). Segmenting tourism in rural areas: the case of North and Central Portugal. *Journal of Travel Research*, 37(May):353–363.
- Keane, M. P. (1997). Modeling heterogeneity and state dependence in consumer choice behavior. *Journal of Business and Economic Statistics*, 15(3):310–327.

- Kelly, J., Haider, W., Williams, P. W., y Englund, K. (2007). Stated preferences of tourists for eco-efficient destination planning options. *Tourism management*, 28(2):377–390.
- Kim, D. y Park, B.-J. R. (2017). The moderating role of context in the effects of choice attributes on hotel choice: A discrete choice experiment. *Tourism Management*, 63:439 – 451.
- Kim, D. y Perdue, R. R. (2013). The effects of cognitive, affective, and sensory attributes on hotel choice. *International Journal of Hospitality Management*, 35:246–257.
- Komppula, R. (2014). The role of individual entrepreneurs in the development of competitiveness for a rural tourism destination: A case study. *Tourism Management*, 40:361–371.
- Kuhfeld, W. F., Tobias, R. D., y Garratt, M. (1994). Efficient experimental design with marketing research applications. *Journal of Marketing Research*, 31:545–557.
- Lancaster, K. (1966). A new approach to consumer theory. *Journal of Political Economy*, 74:132–157.
- Lane, B. (1994). What is rural tourism? *Journal of Sustainable Tourism*, 2(1-2):7–21.
- Lane, B. (2009). Rural tourism: An overview. *The Sage handbook of tourism studies*, pages 354–370.
- Lazarsfeld, P. F. (1950). *Measurement and Prediction*, chapter The Logical and Mathematical Foundations of Latent Structure Analysis, pages 362–412. John Wiley and Sons, New York.
- Lê, S., Josse, J., y Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, 25(1):1–18.
- Lebart, L., Morineau, A., y Warwick, K. M. (1984). *Multivariate descriptive statistical analysis: Correspondence analysis and related techniques for large matrices*. Wiley series in probability and mathematical statistics. Applied probability and statistics. Wiley, New York.
- Li, C.-H. (2016). The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychological methods*, 21(3):369–387.

- Lin, T. H. y Dayton, C. M. (1997). Model selection information criteria for non-nested latent class models. *Journal of Educational and Behavioral Statistics*, 22(3):249–264.
- Linzer, D. A. y Lewis, J. B. (2011). poLCA: An R package for polytomous variable latent class analysis. *Journal of statistical software*, 42(10):1–29.
- Liu, J. y Ma, Y. (2017). The perceptual differences among stakeholders in the tourism supply of Xi'an City, China. *Sustainability*, 9(2).
- Lockyer, T. (2005). The perceived importance of price as one hotel selection dimension. *Tourism Management*, 26(4):529–537.
- Louviere, J. J., Hensher, D. A., y Swait, J. D. (2000). *Stated Choice Methods: Analysis and Application*. Cambridge University Press, Cambridge.
- Louviere, J. J. y Woodworth, G. (1983). Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. *Journal of Marketing Research*, 20:350–367.
- Maldonado-Hinarejos, R., Sivakumar, A., y Polak, J. W. (2014). Exploring the role of individual attitudes and perceptions in predicting the demand for cycling: a hybrid choice modelling approach. *Transportation*, 41(6):1287–1304.
- Márquez, L., Cantillo, V., y Arellana, J. (2020). Assessing the influence of indicators' complexity on hybrid discrete choice model estimates. *Transportation*, 47(1):373–396.
- Márquez, L., Pico, R., y Cantillo, V. (2018). Understanding captive user behavior in the competition between BRT and motorcycle taxis. *Transport policy*, 61:1–9.
- Martín, J. C., Román, C., y Mendoza, C. (2018). Determinants for sun-and-beach self-catering accommodation selection: A stated preference approach. *Tourism Economics*, 24(3):319–336.
- Masiero, L., Nicolau, J. L., y Law, R. (2015). A demand-driven analysis of tourist accommodation price: A quantile regression of room bookings. *International Journal of Hospitality Management*, 50:1–8.
- Masiero, L., Pan, B., y Heo, C. (2016). Asymmetric preference in hotel room choice and implications on revenue management. *International Journal of Hospitality Management*, 26:18–27.

- McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In Zarembka, P., editor, *Frontiers in Econometrics*, pages 105–142. Academic Press, New York.
- McFadden, D. (1981). Econometric models of probabilistic choice. In Manski, C. y McFadden, D., editors, *Structural Analysis of Discrete Data with Econometric Applications*, pages 198–272. MIT Press, Cambridge, MA.
- McFadden, D. y Train, K. (2000). Mixed MNL models of discrete choice response. *Journal of Applied Econometrics*, 15:447–470.
- Mcintosh, A. J. y Siggs, A. (2005). An exploration of the experiential nature of boutique accommodation. *Journal of Travel Research*, 44(1):74–81.
- Merlo, E. M. y de Souza João, I. (2011). Consumers attribute analysis of economic hotels: An exploratory study. *African Journal of Business Management*, 5(21):8410–8416.
- Mmopelwa, G., Kgathi, D., y Molefhe, L. (2007). Tourists' perceptions and their willingness to pay for park fees: A case study of self-drive tourists and clients for mobile tour operators in Moremi Game Reserve, Botswana. *Tourism management*, 28(4):1044–1056.
- Molera, L. y Albaladejo, I. P. (2007). Profiling segments of tourists in rural areas of South-Eastern Spain. *Tourism Management*, 28:757–767.
- Montgomery, D. C. (1991). *Diseño y análisis de experimentos*. Grupo Editorial Iberoamericana, México.
- Morales-Urrutia, X., Morales-Urrutia, D., y Simbaña Taipe, L. (2018). Rural tourism in Spain: An overview by Autonomous Communities. *Advanced Science Letters*, 24:9287–9290.
- Morikawa, T. (1989). *Incorporating stated preference data in travel demand analysis*. PhD thesis, Department of Civil Engineering, MIT, Cambridge, MA.
- Morikawa, T., Ben-Akiva, M., y McFadden, D. (2002). Discrete choice models incorporating revealed preferences and psychometric data. In *Advances in Econometrics*. Emerald Group Publishing Limited.
- Morley, C. L. (1994). Experimental destination choice analysis. *Annals of tourism research*, 21(4):780–791.

- Motoaki, Y. y Daziano, R. A. (2015). A hybrid-choice latent-class model for the analysis of the effects of weather on cycling demand. *Transportation Research Part A: Policy and Practice*, 75:217–230.
- Muthén, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika*, 49(1):115–132.
- Muthén, B. (1993). *Testing structural equation models*, chapter Goodness of Fit with Categorical and Other Non-Normal Variables, pages 205–234. Sage Publications Sage CA, Newbury Park, CA.
- Muthén, B. O., du Toit, S. H. C., y Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Technical report, no publicado.
- Nunkoo, R., Ramkissoon, H., y Gursoy, D. (2013). Use of structural equation modeling in tourism research: Past, present, and future. *Journal of Travel Research*, 52(6):759–771.
- Olsen, G. D. y Swait, J. D. (1998). Nothing is important. Technical report, Faculty of Management, University of Calgary.
- Oppermann, M. (1996). Rural tourism in Southern Germany. *Annals of tourism research*, 23(1):86–102.
- Oppewal, H., Timmermans, H. J. P., y Louviere, J. J. (1997). Modelling the effects of shopping centre size and store variety on consumer choice behaviour. *Environmental and Planning A*, 29:1073–1090.
- Ortuzar, J. D. (2000). *Fundamentals of Stated Preference. Stated Preference Modelling Techniques*. PTRC Education and Research Service Ltd, London.
- Ortuzar, J. D. y Willumsen, L. (2001). *Modelling Transport. 3a edition*. John Wiley and Sons, Chichester.
- Palma, D., Ortúzar, J., Rizzi, L., y Casaubon, G. (2018). Modelling consumers' heterogeneous preferences: a case study with Chilean wine consumers. *Australian Journal of Grape and Wine Research*, 24(1):51–61.
- Papatheodorou, A. (2001). Why people travel to different places. *Annals of Tourism Research*, 28(1):164–179.

- Park, D.-B. y Yoon, Y.-S. (2009). Segmentation by motivation in rural tourism: A Korean case study. *Tourism Management*, 30:99–108.
- Perles Ribes, J., Ramón-Rodríguez, A., y Sevilla, M. (2011). Determinants of the competitive advantage of residential tourism destinations in Spain. *Tourism Economics*, 17:373–403.
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Raveau, S., Daziano, R., Yanez, M., Bolduc, D., y Ortuzar, J. d. D. (2010). Sequential and simultaneous estimation of hybrid discrete choice models: Some new findings. *Transportation Research Record Journal of the Transportation Research Board*, 2156:131–139.
- Revelle, W. (2021). *psych: Procedures for Psychological, Psychometric, and Personality Research*. Northwestern University, Evanston, Illinois. R package version 2.1.6.
- Román, C. y Martín, J. C. (2016). Hotel attributes: Asymmetries in guest payments and gains—a stated preference approach. *Tourism management*, 52:488–497.
- Rosalina, P., Dupre, K., y Wang, Y. (2021). Rural tourism: A systematic literature review on definitions and challenges. *Journal of Hospitality and Tourism Management*, 47:134–149.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. version 0.5–12 (beta). *Journal of statistical software*, 48(2):1–36.
- Ryan, C. y Glendon, I. (1998). Application of leisure motivation scale to tourism. *Annals of Tourism Research*, 25:169–184.
- Satorra, A. y Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. *Latent Variable Analysis in Developmental Research*, pages 399–419.
- Savalei, V. (2014). Understanding robust corrections in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 21(1):149–160.
- Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461 – 464.

- Seddighi, H. R. y Theocharous, A. L. (2002). A model of tourism destination choice: a theoretical and empirical analysis. *Tourism management*, 23(5):475–487.
- Sharma, S. (1996). *Applied multivariate techniques*. John Wiley and Sons Inc.
- Sharpley, J. y Sharpley, R. (1997). *Rural tourism, an introduction*. International Thomson Business Press, London.
- Sharpley, R. (2000). The influence of the accommodation sector on tourism development: lessons from Cyprus. *International journal of hospitality management*, 19(3):275–293.
- Solsona, J. (1999). El turismo rural en la Comunidad Valenciana: análisis y planificación. Aplicación al Alto Mijares. *Sociedad Castellonense de Cultura, serie estudios económicos. Castelló de la Plana*.
- Soto, J. J., Márquez, L., y Macea, L. F. (2018). Accounting for attitudes on parking choice: An integrated choice and latent variable approach. *Transportation Research Part A: Policy and Practice*, 111:65–77.
- Stringam, B. B., Gerdes Jr, J., y Vanleeuwen, D. M. (2010). Assessing the importance and relationships of ratings on user-generated traveler reviews. *Journal of Quality Assurance in Hospitality and Tourism*, 11(2):73–92.
- Swait, J. y Louviere, J. J. (1993). The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research*, 30:305–314.
- Swait, J., Louviere, J. J., y Williams, M. (1994). A sequential approach to exploiting the combined strengths of (SP) and (RP) data: Application to freight shipper choice. *Transportation*, 21:135–152.
- Tam, M. L., Lam, W. H., y Lo, H. P. (2010). Incorporating passenger perceived service quality in airport ground access mode choice model. *Transportmetrica*, 6(1):3–17.
- Thorhauge, M., Cherchi, E., Walker, J. L., y Rich, J. (2019). The role of intention as mediator between latent effects and behavior: application of a hybrid choice model to study departure time choices. *Transportation*, 46(4):1421–1445.
- Train, K. E. (2003). *Discrete choice methods with simulation*. Cambridge University Press, Cambridge.

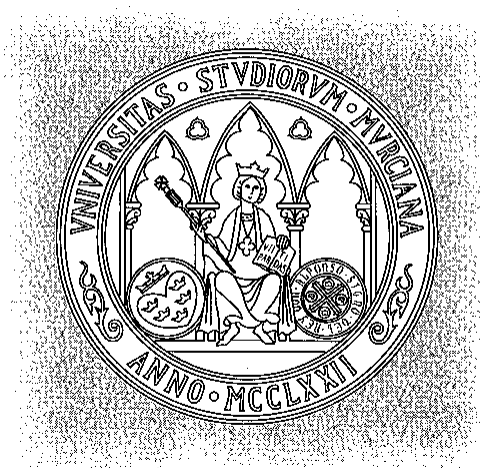
- Uysal, M. y Hagan, L. A. R. (1993). Motivation of pleasure travel and tourism. *Encyclopedia of hospitality and tourism*, 21(1):798–810.
- Valdés, L. (1996). *Introducción a la Economía en España*, chapter El turismo rural en España. E. Civitas, Madrid.
- Valdés, L. y Del Valle, E. (2003). El turismo rural en España. In *Curso Avanzado de Turismo en Áreas Rurales: Estrategias y Promoción*. Instituto Agronómico Mediterráneo de Zaragoza.
- Van Oel, C. J. y van den Berkhof, F. D. (2013). Consumer preferences in the design of airport passenger areas. *Journal of Environmental Psychology*, 36:280–290.
- Victorino, L., Verma, R., Plaschka, G., y Dev, C. (2005). Service innovation and customer choices in the hospitality industry. *Managing Service Quality*, 15:555–576.
- Vij, A. y Walker, J. L. (2016). How, when and why integrated choice and latent variable models are latently useful. *Transportation Research Part B: Methodological*, 90:192–217.
- Vredin-Johansson, M., Heldt, T., Johansson, P., y Vredin Johansson, M. (2006). The effects of attitudes and personality traits on mode choice. *Transportation Research Part A*, 40(6):507–525.
- Walker, J. y Ben-Akiva, M. (2002). Generalized random utility model. *Mathematical Social Sciences*, 43:303–343.
- Wen, J., Me, K., Yang, S., y Liu, F. (2020). Covid-19: potential effects on Chinese citizens' lifestyle and travel. *Tourism Review*, ahead-of-print.
- Yagüe, R. M. (2002). Rural tourism in Spain. *Annals of Tourism Research*, 29:1101–1110.
- Yáñez, M. F., Raveau, S., y Ortúzar, J. d. D. (2010). Inclusion of latent variables in mixed logit models: modelling and forecasting. *Transportation Research Part A: Policy and Practice*, 44(9):744–753.
- Ye, S., Xiao, H., y Zhou, L. (2019). Small accommodation business growth in rural areas: Effects on guest experience and financial performance. *International Journal of Hospitality Management*, 76:29–38.
- Yoo, C.-K., Yoon, D., y Park, E. (2018). Tourist motivation: An integral approach to destination choices. *Tourism Review*.

- Yu, C. Y. y Muthén, B. (2002). Evaluation of model fit indices for latent variable models with categorical and continuous outcomes. Technical report, Graduate School of Education and Information Studies, University of California, Los Angeles.
- Yuan, K.-H. y Bentler, P. M. (1998). Normal theory based test statistics in structural equation modelling. *British Journal of Mathematical and Statistical Psychology*, 51(2):289–309.

Anexos

Anexo I. Cuestionario

UNIVERSIDAD DE MURCIA



NOMBRE DE LA ENCUESTA:

LA DEMANDA DE TURISMO RURAL EN EL NOROESTE DE LA
REGIÓN DE MURCIA

UNIVERSIDAD:

Universidad de Murcia

FACULTAD:

Facultad de Economía y Empresa

PROFESORES RESPONSABLES:

Isabel P. Albaladejo Pina

M^a Teresa Díaz Delfa

M^a Victoria Caballero Pintado

Lourdes Molera Peris

N° cuestionario: _____

CASA: _____

P1. SEXO: Hombre Mujer

P2. ¿De qué población y provincia proviene? _____

P3. ¿Qué años tiene? _____

- | | |
|---|---|
| <input type="checkbox"/> Hasta 20 años. | <input type="checkbox"/> De 36 a 40 años. |
| <input type="checkbox"/> De 21 a 25 años. | <input type="checkbox"/> De 41 a 45 años. |
| <input type="checkbox"/> De 26 a 30 años. | <input type="checkbox"/> De 46 a 50 años. |
| <input type="checkbox"/> De 31 a 35 años. | <input type="checkbox"/> Más de 50 años. |

P4. ESTADO CIVIL:

- | | |
|---|----------------------------------|
| <input type="checkbox"/> Soltero. | <input type="checkbox"/> Casado. |
| <input type="checkbox"/> Separado/Divorciado. | <input type="checkbox"/> Viudo. |

P5. ESTUDIOS:

- Primarios
- Bachillerato o asimilado
- Diplomado Universitario o asimilado
- Licenciado, Ingeniero, Arquitecto, E. Superior Militar o asimilado

P6. SITUACIÓN LABORAL:

- Estudiante (sólo cuando está matriculado en algún centro de estudios).
- En paro.
- Trabaja en tareas del hogar.
- Jubilado.
- Ocupado.

P7. PROFESIÓN:

- | | |
|--|---|
| <input type="checkbox"/> Empresario con trabajadores a su cargo. | <input type="checkbox"/> Empresario sin trabajadores. |
| <input type="checkbox"/> Ejerce una profesión liberal. | <input type="checkbox"/> Funcionario. |
| <input type="checkbox"/> Trabajador por cuenta ajena. | |

P8. Excluyendo el individuo entrevistado, ¿cuántas personas conviven en el núcleo familiar (personas que conviven en un mismo domicilio)? _____

Señale la relación con el individuo entrevistado:

- | | | |
|---|---------------------------------|--|
| <input type="checkbox"/> Padre. | <input type="checkbox"/> Madre. | <input type="checkbox"/> Cónyuge o Pareja. |
| <input type="checkbox"/> Hermanos (indicar el número): _____ | | |
| <input type="checkbox"/> Hijos (indicar el número): _____ | | |
| <input type="checkbox"/> Otros (especificar quienes y cuántos) (Abuelos, suegros, amigos, etc.) _____ | | |

P9. Edades de los hijos:

- 1° hijo: _____ 2° hijo: _____ 3° hijo: _____ 4° hijo: _____ 5° hijo: _____

(Rellenar sólo si tiene cónyuge o pareja que convive en el mismo domicilio)

P10. ESTUDIOS DEL CONYUGE Ó PAREJA:

- Primarios
- Bachillerato o asimilado
- Diplomado Universitario o asimilado
- Licenciado, Ingeniero, Arquitecto, E. Superior Militar o asimilado

P11 SITUACIÓN LABORAL DEL CONYUGE Ó PAREJA:

- Estudiante.
- En paro.
- Trabaja en tareas del hogar.
- Jubilado.
- Ocupado.

P12. PROFESIÓN DEL CONYUGE Ó PAREJA:

- Empresario con trabajadores a su cargo.
- Ejerce una profesión liberal.
- Empresario sin trabajadores.
- Trabajador por cuenta ajena.
- Funcionario.

P13. INGRESOS MENSUALES DEL INDIVIDUO ENTREVISTADO:

- Hasta 601 euros (hasta 100.000 ptas.).
- De 601 a 1.202 euros (de 100.000 a 200.000 ptas.).
- De 1.202 a 1.803 euros (de 200.000 a 300.000 ptas.).
- De 1.803 a 2.404 euros (de 300.000 a 400.000 ptas.).
- Más de 2.404 euros (más de 400.000 ptas.)

P14. INGRESOS FAMILIARES MENSUALES:

- Hasta 1.202 euros (hasta 200.000 ptas.).
- De 1.202 a 2.404 euros (de 200.000 a 400.000 ptas.).
- De 2.404 a 3.606 euros (de 400.000 a 600.000 ptas.).
- Más de 3.606 euros (más de 600.000 ptas.)

P15. ¿Cómo ha alquilado el alojamiento (la casa o la cabaña o la habitación en la hospedería)?

- Directamente a los propietarios.
- Vía NORATUR.
- Otros _____

P16. ¿Cómo obtuvo referencia del lugar?

- A través de conocidos.
- Anuncios en prensa.
- Otros viajes.
- Internet.
- Otros _____

P17. ¿Cuántas noches se va a alojar en la casa? _____

P18. ¿Qué servicios tiene contratados?

- Sólo alojamiento.
- Algunas comidas.
- Senderismo.
- Montar a caballo.
- Alquiler de bicicletas.
- Bajada del río Segura.
- Hacer pan o dulces.
- Colaborar en una matanza.
- Otros _____

P19. ¿Qué actividades le gustaría realizar durante su estancia?

- Senderismo.
- Pasear por los alrededores.
- Montar a caballo.
- Montar en bicicleta.
- Pasear en quarts.
- Bajada del río Segura.
- Hacer pan o dulces.
- Salir a comer a restaurantes típicos.
- Visitar lugares de interés arquitectónico o cultural.
- Observación de la naturaleza.
- Otros _____

P20. Personas con las que realiza el viaje. Señale el número y la relación con el individuo entrevistado:

- Viaja sólo con amigos (indicar el número): _____
- Viaja sólo con la familia (especificar quienes y cuántos): _____
 - Cónyuge o Pareja.
 - Hijos (indicar el número): _____
 - Padre.
 - Madre.
 - Hermanos (indicar el número): _____
 - Otros (especificar quienes y cuántos): _____
- Viaja con amigos y familia (especificar quienes y cuántos): _____
 - Cónyuge o Pareja.
 - Hijos (indicar el número): _____
 - Padre.
 - Madre.
 - Hermanos (indicar el número): _____
 - Amigos (indicar el número): _____
 - Otros (especificar quienes y cuántos): _____

¿Cuántas son menores de 16 años? _____

P21. ¿Cuánto estima que se va a gastar sin contar gastos de alojamiento en este viaje?

- Hasta 30 euros (5.000 ptas.)
- De 30 a 60 euros (de 5.000 a 10.000 ptas.)
- De 60 a 90 euros (de 10.000 a 15.000 ptas.)
- Más de 90 euros (más de 15.000ptas.)

P22. Valore los motivos por los que hace turismo rural:

POSIBLES MOTIVOS	1	2	3	4	5
• Ambiente tranquilo.					
• Contacto con la naturaleza.					
• Belleza del entorno.					
• Ausencia de masificación.					
• Posibilidad de desarrollar actividades al aire libre.					
• Posibilidad de contemplar monumentos o arquitectura típica.					
• Agroactividades (amasar pan, hacer queso,...)					
• Buena comida típica o tradicional.					
• Poder convivir con los amigos.					
• Poder convivir con la familia.					
• Posibilidad de relacionarse con la gente del lugar.					
• Posibilidad de conocer nueva gente.					
• Posibilidad de viajar con niños.					
• Posibilidad de descansar.					
• Independencia y flexibilidad.					
• Distancia de viaje razonable.					
• Precio.					

(1 = muy poco valorado; 5 = muy valorado).

P23. ¿Ha hecho turismo rural anteriormente?

- No.
 Sí. ¿Cuántas veces al año aproximadamente? _____

P24. ¿Habitualmente realiza otro tipo de viajes además de los de turismo rural?

- No.
 Sí. ¿Cuántas veces al año aproximadamente? _____

INTRUCCIONES PARA OBTENER LAS PREFERENCIAS ESTABLECIDAS

Piense que está interesado en pasar un fin de semana en un alojamiento rural. En cada una de las siguientes páginas le ofrecemos la posibilidad de elegir entre tres opciones: dos alojamientos que tienen distintas características o «ninguno de ellos».

ALOJAMIENTO A

ALOJAMIENTO B

OTRO TIPO DE ALOJAMIENTO

Después de leer atentamente las descripciones de los dos alojamientos rurales (A y B), por favor elija el alojamiento que encuentre más deseable (que más se adecue a sus preferencias) o, en caso de no satisfacer sus preferencias, elija «otro tipo de alojamiento». Recuerde, que en este experimento no existen ni respuestas correctas ni respuestas incorrectas, únicamente sus preferencias personales.

Por favor, señale con un círculo sus elecciones en cada una de las páginas. Cada página describe una situación diferente, por eso debe leer con cuidado las descripciones de cada alojamiento antes de hacer su elección. Es importante **COMPLETAR EL EXPERIMENTO ENTERO**. Los cuestionarios incompletos no podrán ser utilizados en este estudio.

Esperamos que disfrute participando en este experimento. Por favor, complételo ya que sus preferencias son importantes. Gracias por participar y ayudarnos.

ESCENARIO 1

	ALOJAMIENTO A	ALOJAMIENTO B
Descripción	<p>Casa de nueva construcción con decoración rústica en la huerta rodeada de árboles</p> <p>4 habitaciones y 4 cuartos de baño.</p>	<p>Casa de nueva construcción con decoración rústica en núcleo urbano con aproximadamente 12.500 habitantes.</p> <p>1 habitación y 1 aseo con ducha.</p>
Instalaciones y servicios	Cocina equipada, calefacción, pistas polideportivas y pequeño corral con animales.	Cocina equipada, calefacción, piscina, pistas polideportivas, pequeño corral con animales, parque infantil y picadero de caballos.
Precio del fin de semana	<p>Únicamente es posible alquilar el alojamiento completo.</p> <p>Completo: 320 €</p>	80 €
Sello del Instituto para la Calidad Turística (ICTE)	SI	NO
Reserva	Únicamente por teléfono/fax	Por Internet o por teléfono/fax

¿Cuál de estos alojamientos de turismo rural elegiría usted para pasar un fin de semana? (Rodee con un círculo su elección)

ALOJAMIENTO A

ALOJAMIENTO B

OTRO TIPO DE ALOJAMIENTO

ESCENARIO 2

	ALOJAMIENTO A	ALOJAMIENTO B
Descripción	Casa antigua rehabilitada con decoración rústica en zona de campo y montaña. 1 habitación y 1 aseo con ducha.	Casa antigua rehabilitada con decoración rústica en núcleo urbano de 12.500 habitantes. 2 habitaciones y 1 cuarto de baño.
Instalaciones y servicios	Cocina equipada, calefacción, parque infantil y piscina.	Cocina equipada, calefacción, comidas por encargo, parque infantil y pequeño corral con animales.
Precio del fin de semana	120 € (19.966 ptas.)	Es posible alquilar el alojamiento completo o por habitaciones. Completo: 200 € (33.278 ptas.) Por habitaciones: 100 € (16.639 ptas.)
Sello del Instituto para la Calidad Turística (ICTE)	SI	SI
Reserva	Por Internet o por teléfono/fax	Por Internet o por teléfono/fax

¿Cuál de estos alojamientos de turismo rural elegiría usted para pasar un fin de semana? (Rodee con un círculo su elección)

ALOJAMIENTO A

ALOJAMIENTO B

OTRO TIPO DE ALOJAMIENTO

ESCENARIO 3

	ALOJAMIENTO A	ALOJAMIENTO B
Descripción	Casa antigua rehabilitada con decoración rústica en zona de campo y montaña. 3 habitaciones y 3 cuartos de baño.	Cabaña de madera en una pedanía o núcleo rural. 3 habitaciones y 3 cuartos de baño.
Instalaciones y servicios	Cocina equipada, calefacción, piscina, parque infantil y picadero de caballos.	Cocina equipada, calefacción, piscina, parque infantil, pistas polideportivas y picadero de caballos.
Precio del fin de semana	Únicamente es posible alquilar el alojamiento completo. Completo: 240 € (39.933 ptas.)	Únicamente es posible alquilar el alojamiento completo. Completo: 360 € (59.898 ptas.)
Sello del Instituto para la Calidad Turística (ICTE)	NO	NO
Reserva	Únicamente por teléfono/fax	Únicamente por teléfono/fax

¿Cuál de estos alojamientos de turismo rural elegiría usted para pasar un fin de semana? (Rodee con un círculo su elección)

ALOJAMIENTO A

ALOJAMIENTO B

OTRO TIPO DE ALOJAMIENTO

ESCENARIO 4

	ALOJAMIENTO A	ALOJAMIENTO B
Descripción	Casa de nueva construcción con decoración rústica en la huerta rodeada de árboles. 2 habitaciones y 1 cuarto de baño.	Antiguo molino rehabilitado a las afueras de una pedanía o núcleo rural. 4 habitaciones y 4 cuartos de baño.
Instalaciones y servicios	Cocina equipada, calefacción, pistas polideportivas, pequeño corral con animales y picadero de caballos.	Cocina equipada, calefacción, comidas por encargo y parque infantil.
Precio del fin de semana	Es posible alquilar el alojamiento completo o por habitaciones. Completo: 240 € (39.933 ptas.) Por habitaciones: 120 € (19.966 ptas.)	Únicamente es posible alquilar el alojamiento completo. Completo: 560 € (93.176 ptas.)
Sello del Instituto para la Calidad Turística (ICTE)	NO	SI
Reserva	Por Internet o por teléfono/fax	Únicamente por teléfono/fax

¿Cuál de estos alojamientos de turismo rural elegiría usted para pasar un fin de semana? (Rodee con un círculo su elección)

 ALOJAMIENTO A

 ALOJAMIENTO B

 OTRO TIPO DE ALOJAMIENTO

Anexo II. Programación en SAS para la obtención de las alternativas y los conjuntos de elección del experimento de preferencias declaradas


```

proc factex;
  factors  comidas      piscina      pistadep      minigranja
           pinfantil    caballos     calidad       ventaint
           aseos        reserva      comidas2     piscina2
           pistadep2    minigranja2 pinfantil2    caballos2
           calidad2     ventaint2   aseos2       reserva2;
  /**model res=4;*/
  model estimate=(comidas--reserva2);
  /**examine aliasing(2) confounding;*/
  size design=minimum;
  output out=turis1 comidas      cvals=('si' 'no')  piscina      cvals=('si'
'no')
                               pistadep  cvals=('si' 'no')  minigranja  cvals=('si'
'no')
                               pinfantil  cvals=('si' 'no')  caballos    cvals=('si'
'no')
                               calidad    cvals=('si' 'no')  ventaint    cvals=('si'
'no')
                               aseos      cvals=('compartido' 'privado')
                               reserva    cvals=('internet' 'telefono')
                               comidas2  cvals=('si' 'no')  piscina2    cvals=('si'
'no')
                               pistadep2  cvals=('si' 'no')  minigranja2 cvals=('si'
'no')
                               pinfantil2 cvals=('si' 'no')  caballos2   cvals=('si'
'no')
                               calidad2   cvals=('si' 'no')  ventaint2   cvals=('si'
'no')
                               aseos2     cvals=('compartido' 'privado')
                               reserva2   cvals=('internet' 'telefono')
  randomize;
run;

```

```

proc factex;
  factors construcc  situ  precio  tamano
           construcc2 situ2  precio2 tamano2 /nlev=4;
  model estimate=(construcc--tamano2);
  size design=minimum;
  output out=turis2 designrep=turis1
           construcc  cvals=('antigua' 'nueva' 'madera' 'molino')
           situ      cvals=('curbano' 'huerta' 'campo' 'curbpeq')
           precio    nvals=(80 100 120 140)
           tamano    nvals=(1 2 3 4)
           construcc2 cvals=('antigua' 'nueva' 'madera' 'molino')
           situ2     cvals=('curbano' 'huerta' 'campo' 'curbpeq')
           precio2   nvals=(80 100 120 140)
           tamano2   nvals=(1 2 3 4)
  randomize;
run;

```

```

proc optex data=turis2 coding=orth;
  class  comidas      piscina      pistadep      minigranja
         pinfantil    caballos     calidad       ventaint
         aseos        reserva      construcc     situ

```

```
        precio      tamano      comidas2   piscina2
        pistadep2   minigranja2 pinfantil2 caballos2
        calidad2    ventaint2    aseos2     reserva2
        construccion situ2      precio2    tamano2;
model comidas--tamano2;
blocks structure=(16) 4;
examine number=1 i var;
generate method=M_Federov n=64 iter=200 keep=10;
output out=turis3 blockname=bloque;
run;
```


Anexo III. Resultados de la estimación del modelo LAC

Tabla 8: Modelo estimado con cuatro grupos determinados mediante el análisis cluster (modelo LAC)

Atributos	Denominación	Modalidades	Grupo 1	Grupo 2	Grupo 3	Grupo 4
ACS	Constante		0	-0,3681	1,8128**	-1,6519
EDAD	Edad del turista	Hasta 25 años	0	-0,3586	-1,8604***	-1,7638
		De 26 a 40 años	0	-0,9215	-1,6817***	-4,1715***
ESTSUP	Educación superior	Sí	0	-0,9430***	-0,6671**	0,4470
ALQ	Forma de alquiler del alojamiento	Directamente al propietario	0	0,4586	-1,0661***	-1,0669
REFLUG	Referencia del lugar	A través de conocidos	0	1,0643***	1,8103***	0,5736
GASTO	Gasto diario por persona y día	Más de 60 euros	0	0,4987	-0,4284	1,4669
DEPOR	Actividades deportivas	Sí	0	0,7095**	-0,1130	-0,7036
CULTURA	Actividades culturales	Sí	0	0,1810	1,0688***	2,5675***
			307			
<i>Observaciones</i>						
<i>Número de parámetros</i>			27			
<i>Función Log-verosimilitud</i>			-302,77			
<i>Pseudo-R² de McFadden</i>			0,1709			
<i>AIC</i>			659,53			
<i>BIC</i>			760,16			
<i>% casos bien clasificados</i>			56,35 %			

Significatividad: '***': $p < 0,001$; '**': $p < 0,01$; '*': $p < 0,05$; '+': $p < 0,1$

Tabla 9: Modelo LCA con tres grupos

Atributos	Denominación	Modalidades	Grupo 1	Grupo 2	Grupo 3
ACS	Constante		0	0,8125	2,9148***
GEN	Género del turista	Hombre	0	-0,3743	-0,5435*
EDAD	Edad del turista	Hasta 25 años De 26 a 40 años	0	0,4725	-1,5365*
SLAB	Situación laboral	Estudiante	0	-1,2989**	-1,5271**
TIPFAM	Tipología familiar	Soltero que vive con sus padres	0	1,1185**	0,9140*
ALQ	Forma de alquiler del alojamiento	Directamente propietario	0	0,3457	-1,1083***
REFLUG	Referencia del lugar	A través de conocidos	0	1,1225***	1,8497***
PERS	Compañeros de viaje	Únicamente con amigos	0	-2,6422***	-1,7722***
GASTO	Gastos diarios por persona y día	Familia y amigos	0	-1,9584***	-1,6206***
DEPOR	Actividades deportivas	Más de 60 euros	0	0,5800*	-0,3153
CULTURA	Actividades culturales	Sí	0	0,4269	-0,3194
		Sí	0	0,4504	1,2066***
		<i>Observaciones</i>	299		
		<i>Número de parámetros</i>	26		
		<i>Función Log-verosimilitud</i>	-268,83		
		<i>Pseudo-R² de McFadden</i>	0,1807		
		<i>AIC</i>	589,66		
		<i>BIC</i>	685,87		
		<i>% casos bien clasificados</i>	57,86 %		

+Significativo al 10%. *Significativo al 5%. **Significativo al 1%. ***Significativo al 0.5 %

Anexo IV. Comparativa de los modelos logit LACL y LAC

Tabla 10: Características socio-económicas de los turistas significativas y signo, según modelo

Atributos	Denominación	Modalidad	Modelo	Significativa	Signo
GEN	Género del encuestado	Hombre	LACL4	Clase 3	-
			LACL3	Clase 3	-
			LCL4		
			LCL3	Grupo 3	-
PROC	Provincia de origen	Murcia	LACL4	Clase 3	-
			LACL3	Clase 3	-
			LCL4		
			LCL3		
EDAD	Edad del turista	Hasta 25 años	LACL4	Clase 3	-
			LACL3	Clase 3	-
			LCL4	Grupo 3	-
			LCL3	Grupo 3	-
		De 26 a 40 años	LACL4	Clases 3 y 4	-
			LACL3	Clase 3	-
			LCL4	Grupos 3 y 4	-
			LCL3	Grupo 3	-
ESTSUP	Educación superior	Sí	LACL4 LACL3 LCL4 LCL3	Grupos 2 y 3	-
SLAB	Situación laboral	Ocupado	LACL3 LCL3		
		Estudiante	LACL3 LCL3	Clase 2 Grupos 2 y 3	- -
TIPFAM	Tipología de unidad familiar	Soltero que vive con sus padres	LACL4	Clases 2 y 3	+
			LACL3	Clases 2 y 3	+
			LCL4		
		LCL3	Grupos 2 y 3	+	
INGFAM	Ingresos familiares mensuales	Más de 2405 euros	LACL4	Clase 4	+
			LACL3		
			LCL4		
			LCL3		

[*]LACL4=Modelo logit con alternativas de elección las 4 clases latentes obtenidas en el Análisis de clases latentes; LACL3=Modelo logit con alternativas de elección las tres primeras clases latentes obtenidas en el Análisis de clases latentes; LCL4=Modelo logit con alternativas de elección los 4 grupos obtenidos en el Análisis de cluster con datos categóricos; LCL3=Modelo logit con alternativas de elección los tres primeros grupos obtenidos en el Análisis de cluster con datos categóricos

Tabla 11: Condiciones de viaje del turista significativas y signo, según modelo

Atributos	Denominación	Modalidad	Modelo	Significativa	Signo
ALQ	Forma de alquiler del alojamiento	Directamente del propietario	LACL4	Clase 3	-
			LACL3	Clase 3	-
			LCL4	Grupo 3	-
			LCL3	Grupo 3	-
REFLUG	Referencia del lugar	A través de conocidos	LACL4	Clases 2 y 3	+
			LACL3	Clases 2 y 3	+
			LCL4	Grupos 2 y 3	+
			LCL3	Grupos 2 y 3	+
PERS	Compañeros de viaje	Únicamente amigos	LACL4	Clases 2, 3 y 4	-
			LACL3	Clases 2 y 3	-
			LCL4		
		Familia y amigos	LCL3	Grupos 2 y 3	-
			LACL4	Clases 2, 3 y 4	-
			LACL3	Clases 2 y 3	-
GASTO	Gastos diarios por persona y día (excluyendo alojamiento)	Más de 60 euros	LCL4	Grupos 2 y 3	-
			LACL4	Clase 3	-
			LACL3	Clase 3	-
			LCL3	Grupos 2	+
TRURAL	Experiencia previa en turismo rural	Sí	LACL3 LCL3		
DEPOR	Actividades deportivas	Sí	LACL4	Clase 2	+
			LACL3		
			LCL4	Grupos 2	+
			LCL3		
AREA	Actividades asociadas con el entorno	Sí	LACL3 LCL3	Clase 2	+
COMIDA	Actividades gastronómicas	Sí	LACL4		
			LACL3	Clase 2	-
			LCL4		
			LCL3		
CULTURA	Actividades culturales	Sí	LACL4	Clases 3 y 4	+
			LACL3	Clase 3	+
			LCL4	Grupos 3 y 4	+
			LCL3	Grupo 3	+

[*]LACL4=Modelo logit con alternativas de elección las 4 clases latentes obtenidas en el Análisis de clases latentes; LACL3=Modelo logit con alternativas de elección las tres primeras clases latentes obtenidas en el Análisis de clases latentes; LCL4=Modelo logit con alternativas de elección los 4 grupos obtenidos en el Análisis de cluster con datos categóricos; LCL3=Modelo logit con alternativas de elección los tres primeros grupos obtenidos en el Análisis de cluster con datos categóricos