

Estimación de compras por internet en los hogares españoles

Armas Méndez, Jennifer; Santana Jiménez, Gisela

2010



<http://www.centros.ulpgc.es/fcee/>
Universidad de Las Palmas de Gran Canaria.
Facultad de Economía, Empresa y Turismo

El conocimiento por parte de las nuevas sociedades avanzadas, adaptadas a los constantes y vertiginosos cambios en el ámbito de la tecnología, sobre si sus ciudadanos se adaptan al mismo ritmo a este fenómeno, suscita un gran interés a todos niveles. Inversiones en I+D y en educación orientadas a la mejora de la calidad de vida, combinando técnicas de desarrollo sostenible son objetivos punteros en las políticas gubernamentales y empresariales.

CONTENIDO:

1. *INTRODUCCIÓN*, pág. 3
 2. *ANTECEDENTES*, pág. 3
 3. *HIPÓTESIS DE LA ESTIMACIÓN: OBJETIVOS*, pág. 5
 4. *METODOLOGÍA: DATOS Y CONFECCIÓN DEL MODELO*, pág. 6
 - 4.1 *LA VARIABLE INGRESOS NETOS DEL HOGAR*, pág. 9
 - 4.2 *LA VARIABLE NIVEL DE ESTUDIOS*, pág. 10
 - 4.3 *LA VARIABLE SITUACIÓN LABORAL*, pág. 10
 - 4.4 *LA VARIABLE OCUPACIÓN*, pág. 11
 - 4.5 *RECODIFICACIÓN DEL RESTO DE VARIABLES*, pág. 12
 5. *ANÁLISIS DE RESULTADOS*, pág. 16
 - 5.1 *SIGNIFICATIVIDAD ESTADÍSTICA*, pág. 19
 - 5.2 *INTERPRETACIÓN DE LOS COEFICIENTES*, pág. 21
 - 5.2.1 *REGRESORES POISSON*, pág. 22
 - 5.2.2 *REGRESORES INFLATE*, pág. 23
 6. *MODELO ZIP vs MODELO ZINB*, pág. 24
 7. *CONCLUSIONES*, pág. 25
- BIBLIOGRAFÍA*, pág. 27
- ANEXO*, pág. 28

1. INTRODUCCIÓN

En este trabajo, deseamos abordar la cuestión de cómo se está esbozando en nuestras vidas, la compra de bienes y servicios a un click del ratón del ordenador o del teléfono móvil. Las Tecnologías de la Información y la Comunicación (TIC en adelante) son las culpables de que sea un hecho efectuar actividades de consumo sin que sea preciso nuestra presencia física. Es posible comprar cualquier cosa sin tenerla delante, con cierto grado de fiabilidad, y que nos lo envíen a una dirección determinada. Las TIC nos permiten establecer relaciones telemáticas con empresas, organizaciones y administraciones públicas sin restricción de horarios, de una forma cómoda, segura y rápida. La cuestión es, hasta qué punto los residentes en España son consecuentes con este hecho y si constituye una herramienta de sus pautas de consumo. A raíz de los datos aportados por una encuesta del INE en este ámbito, deseamos estimar de qué depende que una persona compre o no a través de internet, qué tipo de bienes y servicios son adquiridos y en definitiva, cuáles son las características o circunstancias que invitan al consumo virtual. Otra de las cuestiones que deseamos inferir son si las diferencias socioeconómicas entre individuos son determinantes en nuestro fenómeno de estudio, o bien si es posible establecer alguna diferenciación por razón de sexo, es decir, ¿compran más los hombres que las mujeres? ¿El nivel de estudios o la edad son determinantes en el uso de transacciones telemáticas?

Palabras clave: Compras por internet, Modelos Zip.

2. ANTECEDENTES

Las Tecnologías de la Información y la Comunicación contribuyen a aumentar nuestras capacidades, tanto físicas como mentales, y amplían el abanico de posibilidades a disposición de los ciudadanos para su desarrollo social. Conforman el pilar de una “*cultura tecnológica*” en la que las sociedades más avanzadas se encuentran inmersas.

Para hacerse una idea más o menos clara de lo que son las TIC, no debemos pensar únicamente en un concepto asociado a la informática y otras tecnologías agregadas, sino también a medios de comunicación de cualquier índole. Están sometidas a un ritmo vertiginoso en continuo proceso de cambio, a la par que el progreso científico y en un contexto de globalización económica y cultural. De este modo sobresalen como el soporte que ayuda al

alzamiento de nuevos valores y convertir los conocimientos adquiridos en casi prehistóricos. Este fenómeno implica profundos reajustes en nuestras estructuras sociales, culturales y económicas, recayendo en prácticamente todos los aspectos de la vida cotidiana. El efecto de las TIC en las acciones humanas, empresariales e institucionales es tal, que ha supuesto un gran cambio en el modo de comprender la realidad y confeccionar una filosofía de vida, siendo impensable desprendernos de ellas para actuar de un modo óptimo. Acceder al mercado laboral, la educación, la comunicación, la búsqueda y tratamiento de la información, entre otras, son actividades que cuentan con valiosos recursos que, por otra parte, no debemos obviar la exigencia de tener ciertos conocimientos previos para su uso.

Si tuviéramos que elegir, de entre los componentes que integran las TIC, aquel que ocupara el primer puesto en un podio, sin duda ese sería Internet. **Internet** abre las puertas a un mundo donde es posible realizar casi todo lo que hacemos en el mundo real y además otras funciones contribuyentes a mejorar nuestra calidad de vida.

No obstante, pese a que hoy en día podemos dividirnos en dos mundos, el físico y el ciberespacio, no tener acceso a internet puede producir serias desventajas con relación a los demás. Pensemos por ejemplo en una región no conectada a la red, perdería competitividad en el exterior y carecería de capacidad de adaptación a un nuevo modelo de desarrollo. También podríamos hablar de personas en paro sin conexión con la consiguiente debilidad que ello supone para su acceso al mercado de trabajo. De esto último se deduce que no disponer de internet, o disponer de la red pero no conectarse bien por desconocimiento, por desconfianza hacia la fiabilidad de procedimientos o cualquier otra causa, comportan lo que podemos denominar una brecha digital que asiste al hecho de aumentar las diferencias entre individuos.

“Una vez que toda la información está en la red, una vez que el conocimiento está en la red, el conocimiento codificado, pero no el conocimiento que se necesita para lo que se quiere hacer, de lo que se trata es de saber dónde está la información, cómo buscarla, cómo procesarla, cómo transformarla en conocimiento específico para lo que se quiere hacer. Esa capacidad de aprender, esa capacidad de saber qué hacer con lo que se aprende, esa capacidad es socialmente desigual y está ligada al origen social, al origen familiar, al nivel cultural, al nivel de educación.” (Castells, 2001)

3. HIPÓTESIS DE LA ESTIMACIÓN: OBJETIVOS

Previamente a iniciar el proceso de estimación, hicimos un análisis de qué hipótesis de partida esperábamos refutar con respecto a nuestra intención principal de estimar de qué dependen los productos y servicios que se compran por Internet. Muy a grosso modo, las hipótesis y, por tanto, la intención última de lo que queremos obtener de nuestro estudio, son las siguientes:

1. El nivel de estudios, los ingresos y la ocupación en actividades TIC presenta una relación directa positiva con nuestra variable endógena.
2. Los individuos que más uso hacen de Internet, probablemente compren también. Por tanto, cuantos más servicios (mail, etc.) y descargas se hagan por Internet (siendo esto indicador de un uso significativo de Internet), mayores serán las compras de productos y servicios que se efectúen a través de la misma.
3. Observar si la forma en la que los posibles compradores acceden a Internet incide en las compras o no.
4. Los problemas experimentados en Internet por parte de los usuarios, como el robo de identidad o información bancaria, actúa como freno a las compras por Internet (relación negativa).
5. Deseamos comprobar si la edad y el sexo son factores que funcionan como acicate a las compras por Internet.
6. Por último, verificar el efecto de la nacionalidad del individuo para el caso español, el uso de buscadores para encontrar información y algún efecto región dado por la geografía española.

En función de estas premisas previas y de lo explicado anteriormente, se hace necesario seleccionar las variables que van a intervenir en el proceso de estimación. En el apartado siguiente se explica cuál ha sido el procedimiento de selección de las variables relevantes y las razones que han motivado la elección del modelo.

4. METODOLOGÍA: DATOS Y CONFECCIÓN DEL MODELO

El punto de partida de nuestro estudio fue la “Encuesta sobre Equipamiento y Uso de Tecnologías de Información y Comunicación en los Hogares” del año 2010 (TIC-H). Esta encuesta se viene elaborando anualmente desde 2002 por el INE y recoge información sobre “los productos de tecnologías de información y comunicación de los hogares españoles de estos productos, de Internet y del comercio electrónico” (ine.es).

La Encuesta TIC-H 2010 se hizo a través de un muestreo trietápico estratificado¹. El tamaño de la muestra es de 25.000 viviendas y se dirige a las personas a partir de 10 años de edad residentes en la vivienda. De entre toda la información recolectada (equipamiento del hogar en tecnologías de la información y la comunicación y sobre el uso de ordenador, Internet y comercio electrónico) la estimación está centrada en lo referente al consumo de productos y servicios de distintas tipologías vía Internet por parte de los individuos con una edad igual o superior a 16 años recopilando así 19.384 observaciones. Ante la variedad y número de variables disponibles, hemos hecho un filtrado de las mismas (Ver Tabla 1 Anexo) hasta determinar, según lo que nos dicta más el sentido común que la teoría económica, cuáles son las variables más significativas. Así, entendimos que eran muy vinculantes la provincia de origen, sexo, edad, nacionalidad, nivel de educación, situación laboral, ocupación (nos interesa sobre todo si está relacionado con TICs), ingresos mensuales netos del hogar, posesión de aparatos de información y comunicación (televisión en la vivienda y cómo recibe los canales; ordenador en la vivienda y, si dispone de conexión a Internet, a través de qué vía), causas de no disponer de Internet, servicios a los que accede por Internet, productos y servicios comprados por Internet, descargas, problemas más frecuentes en Internet, etc. Cabe aducir que, salvo contadas excepciones, se trata de variables ficticias.

Pese a contar con el fichero de microdatos de la encuesta, en primera instancia tuvimos que modificar las variables para transformarlas en variables binarias (cero-uno). Originalmente, el INE codificó las respuestas afirmativas (en general a la posesión de, la

¹ Consiste en la división previa de la población de estudio en grupos o clases que se suponen homogéneos con respecto a alguna característica de las que se van a estudiar. A cada uno de estos estratos se le asignaría una cuota que determinaría el número de miembros del mismo que compondrán la muestra. Dentro de cada estrato se suele usar la técnica de muestreo sistemático, una de las técnicas de selección más usadas en la práctica. En este caso concreto, la primera etapa se corresponde con la estratificación por Comunidades Autónomas, la segunda se refiere a las viviendas familiares, y por último, la tercera etapa selecciona a 1 persona en cada vivienda de 16 o más y a todos los niños de 10 a 15 años de la vivienda.

compra de, etc.) con un 1, a las negativas con 6, y las respuestas en blanco y los “no sabe/no contesta” con el valor 9. Para nosotros, sin embargo, era más fácil trabajar con variables “cero-uno” que con 3 categorías². Por tanto, transformamos en Dummies dichas variables.

Por construcción del cuestionario, más de 9.000 missing values no indicaban realmente ausencia de respuestas. En la pregunta nº 31 se les preguntaba a los individuos si alguna vez habían usado internet; aquellos cuyas respuestas fuera NO pasaban directamente a contestar a partir de la pregunta nº 55. Por consiguiente, todas las preguntas intermedias carentes de datos las interpretamos como respuestas de tipo negativo.

Nuestra variable dependiente, el número de compras efectuadas por Internet, cuyos valores van desde 0 hasta 17, es una variable recuento determinada por un modelo Poisson. En una primera fase de construcción del modelo, se creó la variable COMPINT, de naturaleza binaria que toma el valor 1 si el entrevistado ha comprado algo en internet y 0 en caso contrario, de esa forma pudimos averiguar qué porcentaje de la muestra representa las compras virtuales. A continuación se adjunta una tabla donde se observa que el 82,35% de la muestra no realiza compras vía internet frente al 17,65% que sí lo hace. Por consiguiente, el modelo a estimar se denomina *Modelo Poisson inflado de ceros* (en adelante ZIP), donde la parte “inflada” está determinada por un modelo binario tipo logit.

Para realizar las estimaciones, contrastes y demás cálculos pertinentes, el software utilizado será el Stata/SE 11.1.

Tabla 1. Variable COMPINT 1

COMPINT	Freq.	Percent	Cum.
0	15,962	82.35	82.35
1	3,422	17.65	100.00
Total	19,384	100.00	

Otra de las modificaciones previas que introdujimos fue la compilación de las 17 variedades de “productos y servicios comprados en Internet” en una única variable, nuestra endógena del modelo, en la nueva variable “NCOMPINT” la cual representa el número de tipos de bienes y servicios adquiridos vía internet. Para ello, creamos la variable como la suma de

² Convertir las variables en binarias atendía más al hecho de que, en esencia, una respuesta en blanco o “NS/NC” son “missing values”, de ahí que, en aquellas variables donde apareciera el 9 con este significado lo convertimos en Missing. Por otro lado, cabe apuntar que, hemos tenido en cuenta variables ficticias para categorías múltiples en lo referente al nivel de educación o al país de nacionalidad, por ejemplo.

todos los elementos de la encuesta que empezaban por PROD, puesto que éstos representan los tipos de bienes y servicios comprados. Cada dato PROD toma el valor 1 si éste es comprado y cero en otro caso. De este modo la variable respuesta toma el valor 0 si no ha comprado nada, 1 si el individuo ha comprado una única cosa por internet, 2 si ha comprado dos, y así sucesivamente hasta 17. En la siguiente tabla, se observa la clasificación de bienes y servicios.

Tabla 2. Productos y servicios comprados por internet últimos 12 meses³

TIPO DE BB/SS	
PROD1	Alimentación
PROD2	bienes hogar
PROD3	Medicamentos
PROD4	películas, música
PROD5	libros, revistas periódicos
PROD6	material formativo online
PROD7	material deportivo, ropa
PROD8	software juegos de ordenador
PROD9	otro software de ordenador
PROD10	equipo informático
PROD11	equipo electrónico
PROD12	servs. de telecomunicaciones
PROD13	acciones, seguros
PROD14	alojamiento vacaciones
PROD15	otros servicios para viajes
PROD16	entradas para espectáculos
PROD17	otros productos o servicios

Cabe aducir que, unido al “inflado de ceros” ya comentado, la variable endógena presenta otra anomalía muestral que es la sobre-dispersión dado que, como vemos en el siguiente cuadro, la varianza es superior a la unidad.

³ La encuesta se realizó en Julio de 2010.

Tabla 3. Media y Desviación típica Variable Respuesta

Variable	Obs	Mean	Std. Dev.	Min	Max
NCOMPINT	19384	.5605654	1.544719	0	16

4.1 LA VARIABLE INGRESOS NETOS DEL HOGAR

A tenor de la posible relevancia en el nivel de ingresos sobre el consumo virtual, se hace necesario una mención especial a esta variable y cuál ha sido su tratamiento en la encuesta. Se consideran ingresos netos de cada hogar los siguientes (según la metodología del cuestionario):

“Ingresos mensuales netos: recibidos después de las retenciones a cuenta por impuestos, cotizaciones sociales y otros pagos asimilados. Incluidas todas las fuentes de ingreso (en el caso de existir más de una), considerando (para los ingresos del trabajo por cuenta ajena) la parte proporcional de las pagas extraordinarias y otros ingresos extraordinarios percibidos regularmente.”

“Se deberá contabilizar la suma de los ingresos regulares percibidos por todos los miembros del hogar en la actualidad, aporten ó no estos ingresos en su totalidad ó en parte para sufragar los gastos del hogar.”

En el cómputo del importe de estos ingresos mensuales se tendrá en cuenta:

- para los ingresos de trabajo por cuenta ajena, debe sumarse al importe mensual el prorrateo de los ingresos regulares que no tengan periodicidad mensual (pagas extraordinarias y otros ingresos extraordinarios que se perciban regularmente)
- para los ingresos de trabajo por cuenta propia, debe descontarse del importe de los ingresos mensuales, los gastos deducibles, retenciones ó pagos fraccionados.

CODIFICACIÓN DE LOS INGRESOS POR INTERVALOS

- Menos de 1.100 euros..... 1
- De 1.100 a 1.800 euros..... 2

De 1.800 a 2.700 euros.....	3
Más de 2.700 euros.....	4
NS/NR.....	5 ⁴

Tomamos como datos relevantes para esta variable las marcas de clase de cada intervalo con el fin de especificarla como continua. Para el último intervalo de renta, la marca de clase se ha propuesto como 3.500 €. Del total de los encuestados, 4.301 individuos han sido reacios a responder cuáles son los ingresos netos del hogar. Sin embargo, a pesar del sesgo que se produce en la muestra, hemos decidido incorporar la variable al modelo, dada su significación estadística.

4.2 LA VARIABLE NIVEL DE ESTUDIOS

El nivel máximo de estudios adquirido por los entrevistados está ordenado de menor a mayor formación según la siguiente clasificación:

Analfabetos	1
Educación Primaria.....	2
Primera etapa de la Educación Secundaria.....	3
Segunda etapa de la Educación Secundaria.....	4
Formación Profesional de Grado Superior.....	5
Educación Superior Universitaria.....	6
No se puede codificar.....	9

Unimos las categorías analfabetos y educación primaria en un grupo único y tomamos las observaciones que no se podían codificar como categoría de referencia para evitar que educación superior universitaria fuera omitida por multicolinealidad (problemas surgidos realizando pruebas). De esta forma, cada una de las categorías está codificada como una dummy, la cual toma valor 1 cuando se da el suceso y 0 en otro caso. Finalmente contemplamos en el modelo cinco categorías de formación.

4.3 LA VARIABLE SITUACIÓN LABORAL

Siguiendo el mismo criterio del apartado anterior, recodificamos la situación laboral de los encuestados en variables dummy. La clasificación construida en función del cuestionario es la siguiente:

⁴ Estos valores quedarán como missing values en el análisis porque no están codificados en ningún intervalo, con el consiguiente coste de perder 4.301 observaciones.

Trabajando por cuenta ajena.....	1
Trabajando por cuenta propia (se incluye ayuda familiar).....	2
Parado	3
Estudiante.....	4
Labores del hogar.....	5
Pensionista.....	6
Otra situación	7

A tenor de la similitud entre las dos primeras categorías, creamos la variable dicotómica TRABAJA relativa al hecho de estar ocupado (valor 1 si, 0 en otro caso) bajo las siguientes circunstancias:

“Ha de tenerse en cuenta que el informante se considerará trabajando cuando lo haya hecho al menos una hora en la semana de referencia ó cuando, no haciéndolo, tuviera un trabajo del que estuvo ausente temporalmente (vacaciones).”

La categoría “otra situación” ha sido la seleccionada para jugar el papel de referencia mientras que el resto se han convertido en variables binarias tipo cero-uno.

4.4 LA VARIABLE OCUPACIÓN

En esta parte del trabajo nos encontramos con un problema que sesgaba sumamente la muestra haciendo que el modelo perdiera observaciones. La ocupación ha sido codificada a través del cuestionario en dos partes. Primero como OCUPACION1, relativa a individuos con trabajos manuales y no manuales y, OCUPACION2, para trabajadores de las TIC's, cuyo valor 1 indica pertenecer a este grupo y 0 en caso contrario.

La razón por la cual las observaciones se perdían estriba en que solamente los individuos en situación de trabajar son los que contestan a esta parte, los cuales ascienden a 8.043, de éstos, únicamente el 3,23% son trabajadores TIC. En consecuencia, se consideró necesario (dada la posibilidad de su significación estadística) incluir ambos tipos de ocupación como dummies. La primera DOCUPACION1, que toma el valor 1 en caso de que el trabajo en cuestión sea manual y 0 si es no manual. La segunda DOCUPACION2, con valor 1 si el trabajador es TIC y 0 en caso contrario.

4.5 RECODIFICACIÓN DEL RESTO DE VARIABLES

Examinando el cuestionario nos dimos cuenta de que la variable NIVELEST sólo afectaba a los encuestados de nacionalidad española, por consiguiente, convertimos la variable NACIONALIDAD (categórica hasta el momento) en una dummy que toma el valor 1 para españoles y 0 para cualquier otro caso. No obstante, en posteriores pruebas resultó ser no significativa para el modelo ZIP.

Un hecho curioso fue el descubrir que disponer de televisión en la vivienda carece de significación estadística cuando a priori pensábamos que la publicidad emitida a través de este medio podría suponer un impulso para el consumo vía internet. A consecuencia de esto, consideramos apropiado no incorporar como variables del modelo la recepción de canales de cualquier tipo como por ejemplo antena parabólica o ADSL.

Por otro lado, aunque resulte obvio, cabe destacar que disponer de internet en la vivienda (ver tabla VIV_INTER en Anexo) sí es sumamente significativo aunque no lo es la forma de conexión a internet. De hecho, disponer o no de internet da resultados muy ajustados dentro de la muestra: un 49,16% no dispone de internet en casa mientras que el 50,84% sí (a excepción de 66 missing values). Nos pareció razonable efectuar pruebas sobre si los motivos de no disponer de internet facilitados por los encuestados y seleccionados a priori, eran relevantes para la estimación. Nuestra conclusión resultó ser que solamente la percepción de los individuos sobre si los equipos son caros o no (NOINT4) era significativa, no obstante, incluirla suponía omitir la variable VIV_INTER por colinealidad.

En los microdatos que manejamos, la variable "APINT" engloba la realización de apuestas o loterías por internet. Según el modelo, realizar esta actividad por internet, pese a que es realmente exiguo el porcentaje de individuos que afirman hacerlo (1,30%), consideramos incluirla en el modelo porque este servicio no estaba incorporado como tal dentro de las variables codificadas como "PROD". Asimismo, esperamos que tenga una relación positiva con nuestra variable endógena.

De todos los problemas que los internautas pueden experimentar en sus transacciones los más importantes a nuestro criterio son: virus con pérdida de información, intimidad, fishing y el uso fraudulento de tarjetas de crédito (PROBINT). Nos pareció importante incluir estas

variables en el modelo para determinar de qué forman condicionan el comportamiento de los encuestados con respecto a sus compras online.

Decidimos que el tratamiento más adecuado para la variable EDAD era elevándola al cuadrado con el fin de mejorar la estimación del modelo. Consideramos apropiado proceder de esta manera con EDAD porque resulta poco creíble establecer una relación lineal entre ella y la endógena. En este contexto, el cumplir más años no implica comprar más por internet.

Por último, con el fin de averiguar si existe un “efecto región” en función de la Comunidad Autónoma de residencia, y puesto que en la muestra está especificada solamente la provincia, contrastamos si existían diferencias significativas entre las mismas. Como resultado obtuvimos la siguiente tabla:

Tabla 4. Diferencias Significativas Provincias

CPRO	Summary of NCOMPINT		Freq.
	Mean	Std. Dev.	
1	.97058824	2.1566132	136
2	.59171598	1.9192453	169
3	.48514851	1.4479479	404
4	.56598985	1.5781461	394
5	.32258065	1.0677722	62
6	.41545894	1.3974562	414
7	.73308958	1.6458579	547
8	.84089101	1.8708666	1257
9	.71794872	1.7479753	156
10	.36032389	1.2541495	247
11	.26174497	1.1175963	596
12	.67320261	1.5969958	153
13	.52475248	1.7282785	202
14	.25858124	1.0770546	437
15	.58351893	1.5462498	449
16	.37647059	1.3090934	85
17	.71428571	1.8056229	147
18	.20649652	.82493384	431
19	.44927536	1.207041	69
20	.92592593	1.9104711	297
21	.38343558	1.2267161	326
22	.58598726	1.5277123	157
23	.25784753	1.1171267	446
24	.46025105	1.5869377	239
25	.37362637	1.1513118	91
26	.49191686	1.381275	433
27	.51748252	1.7314536	143
28	.92832765	1.9668429	1172
29	.4021164	1.2170153	567
30	.39490446	1.2578707	628
31	.73888183	1.6716629	1574
32	.29770992	1.2322254	131
33	.53001937	1.5355087	1549
34	.41891892	1.1468451	74
35	.35421687	1.0642751	415
36	.56441718	1.651244	326
37	.31724138	1.1468765	145
38	.28065395	.95827236	367
39	.6349892	1.6575377	463
40	.12068966	.53237588	58
41	.52289513	1.5244046	677
42	.86363636	2.173743	44
43	.66447368	1.6395124	152
44	.35	1.038404	100
45	.66129032	1.7742517	248
46	.48773449	1.4280154	693
47	.55701754	1.4755851	228
48	.6989011	1.7139936	455
49	.16	.77180028	75
50	.55802048	1.5516891	586
51	.65957447	1.5488684	94
52	.71052632	1.4770763	76
Total	.56056541	1.5447193	19384

Source	Analysis of Variance		
	SS	df	MS
Between groups	783.340195	51	15.3596117
within groups	45467.556	19332	2.35193234
Total	46250.8962	19383	2.38615778

Bartlett's test for equal variances: $\chi^2(51) = 1.6e+$

A la luz de los datos, podemos observar que efectivamente existen diferencias significativas entre las provincias. Destacan las provincias de Álava (1), Barcelona (8), Madrid (28) y Soria (42) con medias superiores al 0,80. Descartamos Soria por disponer únicamente de 44 observaciones. Debido a esto, construimos las variables relativas a las Comunidades Autónomas de País Vasco, Cataluña y Madrid.

Una vez recodificadas nuestras variables preseleccionadas, hicimos un nuevo filtrado de las mismas, pero esta vez en busca de aquellas que no aportaran mayor sesgo importante en la muestra en relación a una pérdida considerable de observaciones, puesto que ya habíamos asumido la pérdida de más de 4.000 observaciones por incorporar ING_HOG. Tras este último filtrado, obtuvimos una muestra final de 15.039 observaciones.

Elaborada la reconstrucción de los datos en base a las modificaciones efectuadas citadas anteriormente, adjuntamos una tabla con lo que a priori, hemos considerado que puede ser el modelo final:

Tabla 5. Variables Modelo 1: (ZIP) Número de bienes y servicios comprados vía internet

SEXO	Sexo de la persona informante
EDAD	Edad de la persona informante
EDAD2	Edad de la persona informante al cuadrado
DOCUPACION1	Ocupación principal: Manual/No manual
DOCUPACION2	Ocupación principal: Trabajador TIC / No TIC
ING_HOG	Ingresos mensuales netos del hogar (euros)
TRABAJA	Individuo ocupado SI=1, NO=0
ESTUDIANTE	Individuo estudiante SI=1, NO=0
PARADO	Sin ocupación SI=1, NO=0
PENSIONISTA	Jubilado SI=1, NO=0
LAB_HOGAR	Labores del hogar SI=1, NO=0
OTRASIT	Otra situación laboral SI=1, NO=0
VIV_INTER	Dispone de acceso a Internet en la vivienda SI=1, NO=0
SERV37.2	Servicio usado de Internet; viajes y alojamiento SI=1, NO=0
SERV37.5	Servicio usado de Internet; descargar software (excluido el de juegos) SI=1, NO=0
SERV37.6	Servicio usado de Internet; jugar ó descargar juegos, imágenes,... SI=1, NO=0
SERV37.7.1	Servicio usado de Internet; suscripción servicio de noticias SI=1, NO=0
SERV37.9	Servicio usado de Internet; temas de salud SI=1, NO=0
SERV38.1	Servicio usado de Internet; banca electrónica SI=1, NO=0
SERV38.2	Servicio usado de Internet; ventas de bienes y servicios SI=1, NO=0
SERV39.1	Servicio usado de Internet; información de educación y cursos SI=1, NO=0
ANALFABETOS	Sin estudios SI=1, NO=0
PRIMETEDSEC	Estudios primarios SI=1, NO=0
SECUNTEDSEC	Estudios secundarios SI=1, NO=0
FORPROGRADSUP	Formación Profesional de Grado Superior SI=1, NO=0
DESCARGA3	Descargas a través de página Web: programas de ordenador y actualizaciones
APINT	Realización de apuestas (o loterías) a través de Internet
PROBINT1	Problemas experimentados en Internet (ult. 12 meses): virus con pérdida de

	información SI=1, NO=0
PROBINT3	Problemas experimentados en Internet (ult. 12 meses): intimidad SI=1, NO=0
PROBINT4	Problemas experimentados en Internet (ult. 12 meses): phishing / pharming SI=1, NO=0
PROBINT5	Problemas experimentados en Internet (ult. 12 meses): uso fraudulento tarjetas SI=1, NO=0
TAREAIN1	Tarea relacionada con Internet: usar un buscador SI=1, NO=0

5. ANÁLISIS DE LOS RESULTADOS

Los regresores obtenidos en la endógena NCOMPINT del modelo se interpretan igual que en un modelo Poisson convencional, mientras que los relativos a la parte INFLATE no pueden ser tratados como en un modelo logit binomial donde la probabilidad de éxito está determinada por:

$$E(Y_i|X_i) = X_i' * \beta = \Pr(Y_i = 1) = P_i$$

Para el caso que nos ocupa, la probabilidad de éxito es que la variable endógena tome el valor 0:

$$E(Y_i|X_i) = X_i' * \beta = \Pr(Y_i = 0) = P_i = \frac{1}{(1 + e^{X_i\beta})}$$

Por tanto, si esperamos una relación positiva entre una variable explicativa y la variable respuesta, en la parte ínflate esperamos un coeficiente negativo.

El primer análisis que debemos realizar es el de los signos de los coeficientes. No todos los regresores obtenidos poseían el signo esperado.

Los problemas experimentados en internet por los usuarios, variables incluidas en la parte Poisson, arrojaron signos positivos, indicando entonces, que dichos problemas incrementaban el número esperado de compras de bienes y servicios por internet, manteniéndose el resto de variables constantes. Por esta causa se hizo necesario, a fin de lograr un buen modelo (cuanto menos, coherente), eliminarlas de la estimación final, aunque fueran estadísticamente

significativas. A continuación presentamos la tabla con el Modelo 1 donde se observa lo explicado anteriormente:

Tabla 6. Modelo 1

Zero-inflated Poisson regression	Number of obs	=	15039
	Nonzero obs	=	2698
	Zero obs	=	12341
Inflation model = logit	LR chi2(21)	=	2029.76
Log likelihood = -9118.308	Prob > chi2	=	0.0000

NCOMPINT	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
NCOMPINT						
SEXO	.0167574	.0254034	0.66	0.509	-.0330324	.0665471
EDAD	.0193292	.0066077	2.93	0.003	.0063783	.03228
EDAD2	-.0002323	.0000778	-2.99	0.003	-.0003849	-.0000798
ING_HOG	.0000853	.0000128	6.68	0.000	.0000603	.0001103
MADRID	.0843111	.0382674	2.20	0.028	.0093084	.1593137
PAISVASCO	.2070636	.0477882	4.33	0.000	.1134004	.3007267
CATALUÑA	.0388725	.0363465	1.07	0.285	-.0323652	.1101103
SERV37_2	.9234705	.0453988	20.34	0.000	.8344904	1.012451
SERV37_5	.2917234	.0270013	10.80	0.000	.2388017	.344645
SERV37_6	.0707142	.0259131	2.73	0.006	.0199256	.1215029
SERV37_9	.082945	.0263227	3.15	0.002	.0313534	.1345366
SERV38_1	.5282397	.0323272	16.34	0.000	.4648796	.5915997
SERV38_2	.1445832	.0290111	4.98	0.000	.0877224	.201444
SERV39_1	.2098605	.0264565	7.93	0.000	.1580068	.2617143
APINT	.2536368	.0423615	5.99	0.000	.1706098	.3366637
PROBINT1	.0566259	.0239345	2.37	0.018	.0097152	.1035367
PROBINT3	.150786	.0345793	4.36	0.000	.0830119	.2185601
PROBINT4	.0868451	.0495959	1.75	0.080	-.0103611	.1840512
PROBINT5	.1689797	.0717452	2.36	0.019	.0283617	.3095976
TAREAINT1	.473718	.069922	6.77	0.000	.3366733	.6107627
DOCUPACION2	.1158867	.0445463	2.60	0.009	.0285776	.2031957
_cons	-1.669692	.1558343	-10.71	0.000	-1.975122	-1.364262
inflate						
SEXO	-.381004	.0709176	-5.37	0.000	-.5199999	-.242008
EDAD	-.0179065	.01846	-0.97	0.332	-.0540875	.0182744
EDAD2	.0005741	.0002119	2.71	0.007	.0001588	.0009893
ING_HOG	-.0002757	.0000387	-7.12	0.000	-.0003516	-.0001998
PRIMETEDSEC	-.3933661	.1486512	-2.65	0.008	-.6847171	-.102015
SEGUNETEDSEC	-.8522877	.1456171	-5.85	0.000	-1.137692	-.5668834
FORPROGRAD~P	-.9885963	.1569533	-6.30	0.000	-1.296219	-.6809735
EDUCUNI	-1.256654	.150637	-8.34	0.000	-1.551897	-.9614106
TAREAINT1	-1.212133	.1130371	-10.72	0.000	-1.433681	-.990584
VIV_INTER	-.9082432	.0899228	-10.10	0.000	-1.084489	-.7319977
DOCUPACION1	.2700522	.0978487	2.76	0.006	.0782723	.461832
TRABAJA	.3321158	.2984143	1.11	0.266	-.2527655	.9169971
PARADO	.6094075	.3095324	1.97	0.049	.0027351	1.21608
ESTUDIANTE	.9889699	.3331321	2.97	0.003	.336043	1.641897
LABHOGAR	.8380531	.3332122	2.52	0.012	.1849692	1.491137
PENSIONISTA	.5122312	.3274189	1.56	0.118	-.1294981	1.153961
_cons	2.687622	.4986265	5.39	0.000	1.710332	3.664912

Una vez eliminadas las PROBINT, el modelo se presenta como sigue:

Tabla 7. MODELO 2: Eliminadas Var PROBINT

Zero-inflated Poisson regression	Number of obs	=	15039
	Nonzero obs	=	2698
	Zero obs	=	12341
Inflation model = logit	LR chi2(17)	=	1989.46
Log likelihood = -9138.457	Prob > chi2	=	0.0000

NCOMPINT	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
NCOMPINT						
SEXO	.0217992	.025364	0.86	0.390	-.0279135	.0715118
EDAD	.0195301	.0065928	2.96	0.003	.0066084	.0324518
EDAD2	-.000235	.0000776	-3.03	0.002	-.0003871	-.0000829
ING_HOG	.0000853	.0000127	6.69	0.000	.0000603	.0001103
MADRID	.0910784	.0382168	2.38	0.017	.016175	.1659819
PAISVASCO	.2071008	.0477979	4.33	0.000	.1134186	.300783
CATALUÑA	.053038	.0362227	1.46	0.143	-.0179573	.1240332
SERV37_2	.9278679	.0453916	20.44	0.000	.838902	1.016834
SERV37_5	.2996183	.0269565	11.11	0.000	.2467846	.352452
SERV37_6	.0719635	.0258132	2.79	0.005	.0213705	.1225565
SERV37_9	.0890339	.0262945	3.39	0.001	.0374977	.1405702
SERV38_1	.5242543	.0322828	16.24	0.000	.4609811	.5875275
SERV38_2	.1558931	.0289434	5.39	0.000	.099165	.2126211
SERV39_1	.2232836	.0263306	8.48	0.000	.1716766	.2748906
APINT	.2670902	.0422938	6.32	0.000	.1841959	.3499845
TAREAINT1	.4858275	.0698543	6.95	0.000	.3489156	.6227394
DOCUPACION2	.1170604	.0444747	2.63	0.008	.0298915	.2042293
_cons	-1.665566	.1552138	-10.73	0.000	-1.969779	-1.361353
inflate						
SEXO	-.3784491	.0708002	-5.35	0.000	-.5172149	-.2396832
EDAD	-.0178113	.0184277	-0.97	0.334	-.053929	.0183064
EDAD2	.0005738	.0002115	2.71	0.007	.0001592	.0009883
ING_HOG	-.0002755	.0000387	-7.12	0.000	-.0003513	-.0001997
PRIMETEDSEC	-.3937075	.148593	-2.65	0.008	-.6849445	-.1024705
SEGUNTEDSEC	-.853192	.145566	-5.86	0.000	-1.138496	-.567888
FORPROGRAD~P	-.9877604	.1568297	-6.30	0.000	-1.295141	-.68038
EDUCUNI	-1.255716	.1505565	-8.34	0.000	-1.550801	-.9606305
TAREAINT1	-1.210854	.1130157	-10.71	0.000	-1.432361	-.9893475
VIV_INTER	-.9133806	.0897414	-10.18	0.000	-1.089271	-.7374906
DOCUPACION1	.2698124	.0977351	2.76	0.006	.0782551	.4613697
TRABAJA	.3379955	.2985313	1.13	0.258	-.2471151	.9231061
PARADO	.6182108	.3096356	2.00	0.046	.0113362	1.225086
ESTUDIANTE	.999229	.3331969	3.00	0.003	.346175	1.652283
LABHOGAR	.8498628	.3332156	2.55	0.011	.1967723	1.502953
PENSIONISTA	.5145604	.3273791	1.57	0.116	-.127091	1.156212
_cons	2.682753	.4981054	5.39	0.000	1.706484	3.659022

Por su parte, la variable ingresos netos del hogar ha resultado significativa tanto en la parte Poisson como en la parte ínflate. Se espera que si aumentan los ingresos netos del hogar, la probabilidad de comprar más productos o servicios por internet aumente, entonces es de esperar un signo positivo para el regresor. Mientras que si los ingresos aumentan, también se espera que la probabilidad de que NO se realice ninguna compra, sea más pequeña, por lo cual se espera un signo negativo como regresor.

Observemos el coeficiente de la variable ING_HOG para la parte Poisson, éste es:

ING_HOG⁵ (0 .0000853): Tal y como hemos apuntado, es positivo y se interpreta como en cuanto aumentaría el número de compras esperadas (NCOMPINT) si aumentaran los ingresos de los individuos. Por ejemplo, aumentar en 300 € los ingresos netos mensuales del hogar, implica aumentar el nº de compras esperado en un 25,77%.

Ahora bien, observemos el coeficiente de la parte “inflada” para la misma variable:

ING_HOG (-0,0002755): Supongamos una observación con las siguientes características:

Mujer, 30 años, desempleada, con ingresos netos mensuales de 800 €, con disponibilidad de internet en la vivienda; tendría una probabilidad de no comprar por internet del 23,34%⁶. El efecto marginal de un aumento en los ingresos, manteniendo el resto de características constantes, está determinado por la siguiente ecuación, cuyo resultado viene acompañado de la disminución en la probabilidad de no comprar si aumentaran los ingresos en 300 €:

$$\frac{dP_i}{dX_{ik}} = P_i * (1 - P_i) * \beta_k = 0,2334 * (1 - 0,2334) * (-0,0002755) = 4,29E - 5 \Rightarrow_{(+300)} 20,97\%$$

5.1 SIGNIFICATIVIDAD ESTADÍSTICA

Una vez examinado los signos, pasamos a contrastar la significación individual y global de los estimadores. Los estadísticos t nos demuestran algunas variables no significativas.

En la parte Poisson, la variable SEXO no es significativa, sin embargo sí lo es en la parte “íflate”. Así mismo, no existe efecto Comunidad Autónoma para Cataluña en el número de compras vía internet. No obstante, el contraste significación global nos informa de que los regresores son globalmente significativos.

En la parte ínflate, las variables EDAD, TRABAJA y PENSIONISTA han resultado no significativas. Respecto a la variable TRABAJA, no nos parece razonable que no sea significativa. Suponíamos en un principio que el hecho de trabajar, influía positivamente en la decisión de comprar,

⁵ Este coeficiente no se interpretará nuevamente en el siguiente apartado porque aunque sufre una variación muy pequeña en el valor numérico, se mantiene el signo y, por tanto, su significado.

⁶ En el supuesto caso de que la variable EDAD fuera significativa.

debido a que un individuo que trabaja dispone de menos tiempo y oportunidad de realizar compras físicamente y, por tanto, optaría por realizar un consumo virtual. Realizamos el contraste para NCOMPINT si existen diferencias entre trabajar y no trabajar y el resultado es el siguiente:

Ho: No existen diferencias entre trabajar y no trabajar.

Tabla 8. Contraste diferencia entre trabajar y no trabajar

. oneway NCOMPINT TRABAJA, tab

TRABAJA	Summary of NCOMPINT			Freq.
	Mean	Std. Dev.		
0	.214355	.94008813		11341
1	1.048738	2.0242528		8043
Total	.56056541	1.5447193		19384

Source	Analysis of variance			F	Prob > F
	SS	df	MS		
Between groups	3276.09851	1	3276.09851	1477.55	0.0000
within groups	42974.7977	19382	2.217253		
Total	46250.8962	19383	2.38615778		

Bartlett's test for equal variances: $\chi^2(1) = 5.5e+03$ Prob> $\chi^2 = 0.000$

Dado el valor del estadístico: 1.477,55 concluimos rechazar la hipótesis nula y existen diferencias entre trabajar y no trabajar.

Después de este resultado, en búsqueda de una explicación, nos percatamos de que las variables DOCUPACION1 Y 2, relativas a los individuos ocupados, resultaron significativas, por tanto eliminamos del modelo la variable TRABAJA. Además estas variables nos arrojan información adicional sobre el tipo de actividad desempeñada por lo trabajadores, si éstos son empleados TIC's o no.

Una vez eliminadas las variables irrelevantes, a continuación presentamos nuestro modelo final:

Tabla 9. MODELO ZIP FINAL

Zero-inflated Poisson regression	Number of obs	=	15039
	Nonzero obs	=	2698
	Zero obs	=	12341
Inflation model = logit	LR chi2(15)	=	1996.43
Log likelihood = -9142.063	Prob > chi2	=	0.0000

NCOMPINT	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
NCOMPINT						
EDAD	.0228124	.0062998	3.62	0.000	.0104649	.0351598
EDAD2	-.0002708	.0000745	-3.64	0.000	-.0004168	-.0001248
ING_HOG	.0000859	.0000127	6.76	0.000	.0000061	.0001109
MADRID	.0826932	.0377829	2.19	0.029	.0086401	.1567463
PAISVASCO	.2000878	.0475064	4.21	0.000	.106977	.2931985
SERV37_2	.9274175	.045358	20.45	0.000	.8385174	1.016318
SERV37_5	.3057598	.0263826	11.59	0.000	.2540508	.3574688
SERV37_6	.0726675	.0257897	2.82	0.005	.0221205	.1232144
SERV37_9	.084435	.0260171	3.25	0.001	.0334425	.1354275
SERV38_1	.525612	.032282	16.28	0.000	.4623404	.5888837
SERV38_2	.1619041	.0286639	5.65	0.000	.1057239	.2180843
SERV39_1	.2230263	.0263204	8.47	0.000	.1714392	.2746133
APINT	.2720886	.0420138	6.48	0.000	.1897431	.3544341
TAREINT1	.4860937	.069827	6.96	0.000	.3492352	.6229521
DOCUPACION2	.1226113	.0441279	2.78	0.005	.0361221	.2091004
_cons	-1.720816	.1500303	-11.47	0.000	-2.01487	-1.426762
inflate						
SEXO	-.3890076	.0686096	-5.67	0.000	-.52348	-.2545353
EDAD2	.0004043	.0000351	11.52	0.000	.0003355	.0004731
ING_HOG	-.0002803	.0000385	-7.29	0.000	-.0003557	-.0002049
PRIMETEDSEC	-.4049851	.1483931	-2.73	0.006	-.6958302	-.11414
SEGUNTEDSEC	-.8748173	.1448061	-6.04	0.000	-1.158632	-.5910024
FORPROGRAD~P	-1.00997	.1559518	-6.48	0.000	-1.31563	-.7043103
EDUCUNI	-1.27977	.1494543	-8.56	0.000	-1.572695	-.9868455
TAREINT1	-1.209679	.1128884	-10.72	0.000	-1.430936	-.9884215
VIV_INTER	-.9245254	.0895519	-10.32	0.000	-1.100044	-.749007
DOCUPACION1	.2580736	.0967822	2.67	0.008	.068384	.4477633
PARADO	.2771348	.110435	2.51	0.012	.0606862	.4935835
ESTUDIANTE	.763816	.1431	5.34	0.000	.4833452	1.044287
LABHOGAR	.4871958	.1681023	2.90	0.004	.1577214	.8166702
_cons	2.64827	.1984518	13.34	0.000	2.259311	3.037228

5.2 INTERPRETACIÓN DE LOS COEFICIENTES

En este apartado vamos a definir el significado de cada coeficiente. Al tratarse en algunos casos de variables que forman parte de un mismo grupo analizaremos el efecto de una de ellas en concreto sobre la endógena. De tal manera que al analizar el efecto de una explicativa, se aplicará la condición *ceteris paribus* para el resto.

El efecto de la edad: para este caso disponemos de 2 variables. Por un lado, EDAD representando el hecho de cumplir más años implica aumentar el número de compras esperadas para ciertos grupos de edad. Y por otro lado, EDAD2 hace alusión al hecho de que, a partir de una cierta edad (40 años), el cumplir más años no implica comprar más, ver tabla 2

Anexo, se observan los efectos mencionados y de ahí la coherencia de los signos de sus regresores.

5.2.1 REGRESORES POISSON

Efecto región: habíamos visto que las regiones pertenecientes a las Comunidades Autónomas de Madrid y País Vasco era donde más se compraba. Por tanto, cabía esperar que existiera un efecto región positivo. Así, ser de Madrid frente a proceder de otra Comunidad Autónoma (*ceteris paribus*), incrementa el número de compras esperadas en un 8,2%. Cabe apuntar que el efecto región del País Vasco es mayor que el de Madrid, lo que implica que residir en el País Vasco implica un aumento en las compras esperadas de un 20%.

Efecto de los servicios usados en internet: para las variables relativas a los servicios usados por internet, esperábamos signos positivos porque esto influiría positivamente en el número de compras esperadas. Basaremos la interpretación en el coeficiente más alto relativo a la variable `SERV37_2`: buscar viajes y alojamiento aumenta el número esperado de compras por internet en un 92,74% con respecto a los demás servicios. Estos servicios provocan un efecto sumamente elevado sobre las compras esperadas dado que representan el 98% de la muestra sobre el total de compras de bienes y servicios efectuadas por internet (el número de individuos que compran por internet asciende a 3.422, de los que compran viajes y alojamiento son 3.355. (Ver tabla 3 Anexo PROD14, PROD15 y NCOMPINT.).

Realización de apuestas y loterías por internet (si nos permite la licencia, lo hemos denominado “efecto ludópata/consumista”): tal y como esperábamos, el coeficiente de esta variable tiene un signo positivo. No obstante, encontramos sorprendente que esta variable, no siendo importante en la muestra, sea significativa y produzca un aumento del 27% en el número de compras esperadas.

Tareas realizadas a través de internet: la búsqueda de información utilizando en la red aumenta el número de compras esperadas. De modo que usar buscadores por internet aumenta el número esperado de compras en 48,6%.

Efecto ocupación en TIC's: ser un trabajador TIC frente a tener una ocupación no TIC incrementa el número esperado de compras en un 12,26%.

Constante: si la hipótesis nula del contraste de significación global se acepta (todos los regresores son 0 excepto la constante. Hipótesis siempre rechazada⁷), esto implica que el número de compras disminuye en 1,72.

5.2.2 REGRESORES INFLATE

Efecto género: en este caso vemos que ser hombre implica comprar más. De hecho, ser hombre es un factor de protección ($e^{\beta} < 1$), y eso que supone menor proporción respecto a las mujeres. Los hombres representan un 44,62% de la muestra.

Efecto edad: tal y como esperábamos existe una correspondencia entre los coeficientes de la EDAD2 entre la parte Poisson y la parte Inflate. Así, en este caso, si escogiéramos a un hombre, de 50 años, con internet en casa y con una ocupación no manual, cumplir un año más implica un efecto marginal del 7,53% en detrimento del número esperado de compras.

Efecto educación: a medida que el nivel de formación aumenta, vemos que los coeficientes en valor absoluto también aumente, y es de esperar que aumenten las compras por internet (coherencia con el signo negativo). Para calcular el efecto real sobre la variable respuesta nos centraremos en la educación universitaria, siendo este análisis análogo para el resto. Así, poseer educación universitaria disminuye la probabilidad de no comprar.

Tareas realizadas a través de internet: vemos como en este caso disminuye la probabilidad de no comprar, correspondiéndose este resultado con lo obtenido en la parte Poisson donde las compras esperadas aumentan.

Internet en la vivienda: obviamente, este regresor cumple con todas nuestras expectativas y no nos ha defraudado, en el sentido de que la probabilidad de efectuar compras por internet (es decir, $\Pr(Y=1)$) aumenta.

Efecto ocupación en trabajos manuales: recordemos que para este caso tener la característica tomaba valor 1 y 0 en caso de trabajos no manuales. Por tanto, desempeñar una ocupación manual disminuye la probabilidad de no comprar.

Efecto de la situación laboral: los estadios contemplados son trabajar, estudiar, estar parado, dedicarse a sus labores y pensionista. Recordemos que la variable TRABAJA y PENSIONISTAS

⁷ Todos los contrastes de significación global dan como resultado que los coeficientes son globalmente significativos.

fueron eliminadas por no ser significativas, y dejamos como categoría de referencia “otra situación”. Analizaremos el resto de los coeficientes. Estar parado, o bien ser estudiante, o dedicarse a labores del hogar, frente a estar en otra situación, supone un factor de riesgo para nuestra variable endógena ($e^{\beta_k} > 1$)

Constante: si no se da ninguna de las características de la parte “inflat”, la probabilidad de no comprar supondría un 91,27%.

6. MODELO ZIP vs MODELO POISSON

En este apartado último la intención es verificar si para nuestra variable respuesta, que atiende a un modelo Poisson inflado de ceros, el modelo correcto hubiera sido un Poisson estándar.

Para ello, empleamos el contraste de Vuong:

Vuong test of zip vs. standard Poisson: $z = 20.62$ $Pr > z = 0.0000$

Ho: El modelo es Poisson.

Conclusión: Rechazamos la hipótesis nula, el modelo ZIP es el correcto.

7. CONCLUSIONES

Como discusión final solo queda concluir si se han alcanzado los objetivos propuestos a priori en función de las premisas planteadas como hipótesis.

Hemos conseguido demostrar que efectivamente los servicios usados en internet suponen un gran impulso para efectuar compras de bienes y servicios en la red, y puesto que, los problemas sufridos del tipo pérdida de información por virus, o uso fraudulento de tarjetas, por un porcentaje no muy elevado de la muestra no han resultado relevantes, consideramos que existe un alto grado de confianza por parte de los usuarios con respecto a sus transacciones telemáticas.

Dentro de todos los servicios buscados en la red, los más relevantes han sido sin duda aquellos destinados a la búsqueda de viajes y alojamiento. De este modo, podemos afirmar que las actividades de ocio preferidas por los españoles internautas son las que están relacionadas con las vacaciones, ocupando un papel secundario el resto de bienes y servicios. Esto puede explicar en gran parte un fenómeno que estamos viviendo en nuestro país, y en general, la continua desaparición de las agencias de viajes y la dura competencia en precios entre las empresas que ofertan paquetes vacacionales online.

En relación al sexo de los informantes, se ha constatado que ser hombre supone que el número de compras efectuadas de bienes y servicios es mayor que para las mujeres. Esto se debe, probablemente, a la adquisición por parte de los varones de productos tales como software, equipos informáticos y electrónicos entre otros.

Por otra parte, el simple hecho de obtener información a través de cualquier buscador, provoca un notable crecimiento del número de compras efectuadas. Pensemos, por ejemplo, en el bombardeo de publicidad de cualesquiera empresas que saltan a la pantalla del usuario, una vez que pincha en el vínculo que le interesa. Otra razón para afirmar que publicitarse en internet es efectivo.

Ha quedado claro que hogares con niveles de ingresos altos son más propensos a efectuar compras por internet que aquellos que perciben ingresos más bajos. Si dentro del

hogar, además, encontramos algún individuo cuya actividad laboral está dentro del grupo de las TIC's, esta propensión será aún mayor.

A pesar de la baja proporción de analfabetismo en España, hay un alto porcentaje de individuos con algún tipo de estudios situados en un grupo de edad que abarca de los 40 años en adelante que no realiza compras online, ni hacen uso de la red para ninguna otra transacción, siendo los individuos más jóvenes los que representan el grueso de los usuarios de internet.

La importancia del nivel de estudios ha quedado reflejada como un factor impulsor en el fenómeno estudiado. Aquellas personas con cierta formación, y a medida que esta va alcanzando niveles superiores, tendrán una mayor disposición a realizar sus compras en internet con mayor asiduidad. Obviamente, la consecución de estudios avanzados lleva implícito el manejo de herramientas informáticas y la adquisición de conocimientos relacionados con las mismas.

Teniendo en cuenta el papel espectacular que protagonizan las TIC's en las sociedades modernas, la importancia de la educación de los ciudadanos y las diferencias entre las provincias españolas, encontrado un "efecto región" en comunidades autónomas con mayor esfuerzo inversor en I+D, se hace necesario por parte de las Administraciones Públicas, nacionales y supranacionales, y del Sector Privado, una mayor coordinación entre regiones en su intensidad innovadora y educativa. Así mismo, no debemos olvidar el compromiso que deben adquirir los ciudadanos en aprovechar al máximo estos recursos.

BIBLIOGRAFÍA

LIBROS

Castells, Manuel: *"Internet y la Sociedad red"* (2001). Disponible en web:

[Castells, 2001](#)

Wooldridge, Jeffrey M.: *"Introducción a la econometría. Un enfoque moderno"* (2005). Ed.

Thompson

ARTÍCULOS

Stata Annotated Output Zero-inflated Poisson Regression. UCLA. Disponible en web:

[UCLA ZIP](#)

ANEXO

TABLAS VARIABLES FILTRADO

. tab ING_HOG

ING_HOG	Freq.	Percent	Cum.
550	6,029	39.97	39.97
1450	4,594	30.46	70.43
2250	2,632	17.45	87.88
3500	1,828	12.12	100.00
Total	15,083	100.00	

. tab SIT_LAB

SIT_LAB	Freq.	Percent	Cum.
1	6,426	33.15	33.15
2	1,617	8.34	41.49
3	1,956	10.09	51.58
4	986	5.09	56.67
5	2,444	12.61	69.28
6	5,646	29.13	98.41
7	309	1.59	100.00
Total	19,384	100.00	

. tab NIVELEST

NIVELEST	Freq.	Percent	Cum.
1	702	3.62	3.62
2	6,704	34.61	38.23
3	3,927	20.27	58.51
4	3,338	17.23	75.74
5	1,546	7.98	83.72
6	3,153	16.28	100.00
Total	19,370	100.00	

. tab DOCUPACION1

DOCUPACION1	Freq.	Percent	Cum.
0	16,313	84.16	84.16
1	3,071	15.84	100.00
Total	19,384	100.00	

. tab DOCUPACION2

DOCUPACION2	Freq.	Percent	Cum.
0	19,124	98.66	98.66
1	260	1.34	100.00
Total	19,384	100.00	

. tab NACIONALIDAD

NACIONALIDAD	Freq.	Percent	Cum.
0	1,109	5.72	5.72
1	18,275	94.28	100.00
Total	19,384	100.00	

. tab TELEV

TELEV	Freq.	Percent	Cum.
0	95	0.49	0.49
1	19,288	99.51	100.00
Total	19,383	100.00	

. tab RTV1

RTV1	Freq.	Percent	Cum.
0	16,324	85.07	85.07
1	2,865	14.93	100.00
Total	19,189	100.00	

. tab RTV2

RTV2	Freq.	Percent	Cum.
0	16,257	84.60	84.60
1	2,960	15.40	100.00
Total	19,217	100.00	

. tab RTV3

RTV3	Freq.	Percent	Cum.
0	2,976	15.48	15.48
1	16,247	84.52	100.00
Total	19,223	100.00	

. tab RTV4

RTV4	Freq.	Percent	Cum.
0	18,323	95.32	95.32
1	900	4.68	100.00
Total	19,223	100.00	

. tab ORD

ORD	Freq.	Percent	Cum.
0	7,755	40.01	40.01
1	11,627	59.99	100.00
Total	19,382	100.00	

. tab VIV_INTER

VIV_INTER	Freq.	Percent	Cum.
0	9,497	49.16	49.16
1	9,821	50.84	100.00
Total	19,318	100.00	

. tab FORINT1

FORINT1	Freq.	Percent	Cum.
0	2,621	26.96	26.96
1	7,102	73.04	100.00
Total	9,723	100.00	

. tab FORINT2

FORINT2	Freq.	Percent	Cum.
0	7,731	79.90	79.90
1	1,945	20.10	100.00
Total	9,676	100.00	

. tab FORINT4

FORINT4	Freq.	Percent	Cum.
1	225	2.32	2.32
6	9,460	97.68	100.00
Total	9,685	100.00	

. tab FORINT5

FORINT5	Freq.	Percent	Cum.
0	9,397	97.17	97.17
1	274	2.83	100.00
Total	9,671	100.00	

. tab NOINT1

NOINT1	Freq.	Percent	Cum.
0	8,377	88.64	88.64
1	1,074	11.36	100.00
Total	9,451	100.00	

. tab NOINT2

NOINT2	Freq.	Percent	Cum.
0	7,308	77.48	77.48
1	2,124	22.52	100.00
Total	9,432	100.00	

. tab NOINT3

NOINT3	Freq.	Percent	Cum.
0	3,653	38.64	38.64
1	5,801	61.36	100.00
Total	9,454	100.00	

. tab NOINT4

NOINT4	Freq.	Percent	Cum.
0	7,500	80.81	80.81
1	1,781	19.19	100.00
Total	9,281	100.00	

. tab NOINT7

NOINT7	Freq.	Percent	Cum.
0	9,205	98.31	98.31
1	158	1.69	100.00
Total	9,363	100.00	

. tab DISPMOVINT1

DISPMOVINT1	Freq.	Percent	Cum.
0	17,747	91.55	91.55
1	1,637	8.45	100.00
Total	19,384	100.00	

. tab DISPMOVINT2

DISPMOVINT2	Freq.	Percent	Cum.
0	18,943	97.72	97.72
1	441	2.28	100.00
Total	19,384	100.00	

. tab DISPMOVINT3

DISPMOVINT3	Freq.	Percent	Cum.
0	18,794	96.96	96.96
1	590	3.04	100.00
Total	19,384	100.00	

. tab SERV37_2

SERV37_2	Freq.	Percent	Cum.
0	13,566	69.99	69.99
1	5,818	30.01	100.00
Total	19,384	100.00	

. tab SERV37_5

SERV37_5	Freq.	Percent	Cum.
0	16,659	85.94	85.94
1	2,725	14.06	100.00
Total	19,384	100.00	

. tab SERV37_6

SERV37_6	Freq.	Percent	Cum.
0	15,689	80.94	80.94
1	3,695	19.06	100.00
Total	19,384	100.00	

. tab SERV37_7_1

SERV37_7_1	Freq.	Percent	Cum.
0	18,784	96.90	96.90
1	600	3.10	100.00
Total	19,384	100.00	

. tab SERV37_9

SERV37_9	Freq.	Percent	Cum.
0	14,525	74.93	74.93
1	4,859	25.07	100.00
Total	19,384	100.00	

. tab SERV38_1

SERV38_1	Freq.	Percent	Cum.
0	15,687	80.93	80.93
1	3,697	19.07	100.00
Total	19,384	100.00	

. tab SERV39_1

SERV39_1	Freq.	Percent	Cum.
0	14,721	75.94	75.94
1	4,663	24.06	100.00
Total	19,384	100.00	

. tab DESCARGA1

DESCARGA1	Freq.	Percent	Cum.
0	170	52.96	52.96
1	151	47.04	100.00
Total	321	100.00	

. tab DESCARGA2

DESCARGA2	Freq.	Percent	Cum.
0	593	70.93	70.93
1	243	29.07	100.00
Total	836	100.00	

. tab DESCARGA3

DESCARGA3	Freq.	Percent	Cum.
0	186	33.27	33.27
1	373	66.73	100.00
Total	559	100.00	

. tab SEXO

SEXO	Freq.	Percent	Cum.
0	10,734	55.38	55.38
1	8,650	44.62	100.00
Total	19,384	100.00	

. tab APINT

APINT	Freq.	Percent	Cum.
0	19,132	98.70	98.70
1	252	1.30	100.00
Total	19,384	100.00	

. tab PROBINT1

PROBINT1	Freq.	Percent	Cum.
0	16,506	85.15	85.15
1	2,878	14.85	100.00
Total	19,384	100.00	

. tab PROBINT3

PROBINT3	Freq.	Percent	Cum.
0	18,783	96.90	96.90
1	601	3.10	100.00
Total	19,384	100.00	

. tab PROBINT4

PROBINT4	Freq.	Percent	Cum.
0	19,077	98.42	98.42
1	307	1.58	100.00
Total	19,384	100.00	

. tab PROBINT5

PROBINT5	Freq.	Percent	Cum.
0	19,226	99.18	99.18
1	158	0.82	100.00
Total	19,384	100.00	

. tab TAREAINT1

TAREAINT1	Freq.	Percent	Cum.
0	11,877	61.27	61.27
1	7,507	38.73	100.00
Total	19,384	100.00	

Tabla 1 Anexo. Variables encuesta INE

CPRO	Identificador de la Provincia
SEXO	Sexo de la persona informante
EDAD	Edad de la persona informante
NACIONALIDAD	Nacionalidad de la persona entrevistada
NIVELEST	Estudios terminados, código
SIT_LAB	Situación laboral en la que se encuentra
OCUPACION2	Ocupación principal: Trabajador TIC / No TIC
ING_HOG	Ingresos mensuales netos del hogar
TELEV	Televisión en la vivienda
RTV1	Recibe canales a través de antena parabólica
RTV2	Recibe canales a través de cable
RTV3	Recibe canales de televisión digital terrestre (TDT)
RTV4	Recibe canales a través de ADSL o línea telefónica
ORD	Disponibilidad de ordenador
VIV_INTER	Dispone de acceso a Internet en la vivienda
FORINT1	Forma de conexión a Internet en la vivienda=ADSL
FORINT2	Forma de conexión a Internet en la vivienda=Cable
FORINT4	Forma de conexión a Internet en la vivienda=Otras conexiones banda ancha
FORINT5	Forma de conexión a Internet en la vivienda=Módem/RDSI
NOINT1	Motivo de NO disponer conexión a Int en vivienda=Acceden desde otros lugares
NOINT2	Motivo de NO disponer conexión a Int en vivienda=No quieren
NOINT3	Motivo de NO disponer conexión a Int en vivienda=No necesitan
NOINT4	Motivo de NO disponer conexión a Int en vivienda=costes equipo caros
NOINT7	Motivo de NO disponer conexión a Int en vivienda=seguridad, privacidad
DISPMOVINT1	Uso dispositivos móviles para acceso a Internet=Teléfono móvil banda ancha
DISPMOVINT2	Uso dispositivos móviles para acceso a Internet=Otro teléfono móvil
DISPMOVINT3	Uso dispositivos móviles para acceso a Internet=Ordenador de mano
SERV37.2	Servicio usado de Internet; viajes y alojamiento
SERV37.5	Servicio usado de Internet; descargar software (excluido el de juegos)
SERV37.6	Servicio usado de Internet; jugar ó descargar juegos, imágenes,...
SERV37.7.1	Servicio usado de Internet; suscripción servicio de noticias
SERV37.9	Servicio usado de Internet; temas de salud
SERV38.1	Servicio usado de Internet; banca electrónica
SERV38.2	Servicio usado de Internet; ventas de bienes y servicios
SERV39.1	Servicio usado de Internet; información de educación y cursos
PROD1	Productos/servicios comprados en Internet (últimos 12 meses), alimentación
PROD2	Productos/servicios comprados en Internet (últimos 12 meses), bienes hogar
PROD3	Productos/servicios comprados en Internet (últimos 12 meses), medicamentos
PROD4	Productos/servicios comprados en Internet (últimos 12 meses), películas, música
PROD5	Productos/servicios comprados en Internet (últimos 12 meses), libros, revistas periodicos
PROD6	Productos/servicios comprados en Internet (últimos 12 meses), material formativo online
PROD7	Productos/servicios comprados en Internet (últimos 12 meses), material deportivo, ropa
PROD8	Productos/servicios comprados en Internet (últimos 12 meses), software de juegos de ordenador
PROD9	Productos/servicios comprados en Internet (últimos 12 meses), otro software de ordenador
PROD10	Productos/servicios comprados en Internet (últimos 12 meses), equipo informatico
PROD11	Productos/servicios comprados en Internet (últimos 12 meses), equipo electrónico
PROD12	Productos/servicios comprados en Internet (últimos 12 meses), serv. de telecomunic.
PROD13	Productos/servicios comprados en Internet (últimos 12 meses), acciones, seguros
PROD14	Productos/servicios comprados en Internet (últimos 12 meses), alojamiento vacaciones
PROD15	Productos/servicios comprados en Internet (últimos 12 meses), otros servicios para viajes
PROD16	Productos/servicios comprados en Internet (últimos 12 meses), entradas para espectáculos

PROD17	Productos/servicios comprados en Internet (últimos 12 meses), otros productos o servicios
DESCARGA1	Descargas a través de página Web: películas, música
DESCARGA2	Descargas a través de página Web: libros, revistas, periódicos, material formativo online
DESCARGA3	Descargas a través de página Web: programas de ordenador y actualizaciones
APINT	Realización de apuestas (o loterías) a través de Internet
PROBINT1	Problemas experimentados en Internet (ult. 12 meses): virus con pérdida de información
PROBINT3	Problemas experimentados en Internet (ult. 12 meses): intimidad
PROBINT4	Problemas experimentados en Internet (ult. 12 meses): phishing / pharming
PROBINT5	Problemas experimentados en Internet (ult. 12 meses): uso fraudulento tarjetas
TAREAIN1	Tarea relacionada con Internet: usar un buscador

Tabla 2 Anexo 2. Edad-Variable respuesta

EDAD	Summary of NCOMPINT		Freq.
	Mean	Std. Dev.	
16	.27428571	.84020484	175
17	.34285714	.96319801	175
18	.49132948	1.3537996	173
19	.65662651	1.5597741	166
20	.88823529	1.6513872	170
21	.93167702	1.823884	161
22	.9625	1.7729182	160
23	1.1156463	1.9532797	147
24	1.357513	2.1217274	193
25	1.4583333	2.4294948	168
26	1.431694	2.2297886	183
27	1.3497537	2.2930885	203
28	1.4676617	2.2716071	201
29	1.087156	1.7378201	218
30	.97095436	1.7063566	241
31	1.0221402	1.8874509	271
32	1.2191235	2.1042321	251
33	1.1737805	2.115253	328
34	1.2098361	2.1093234	305
35	1.0253165	1.9597517	316
36	1.1437309	2.0546555	327
37	1.1923077	2.1808953	364
38	1.1815718	2.0869904	369
39	.95616438	2.0443561	365
40	1.0531915	2.0994832	376
41	.87960688	1.7540247	407
42	.8537859	1.8488797	383
43	.91056911	2.1277468	369
44	.87087087	1.8493303	333
45	.99722992	2.0956285	361
46	.78362573	1.9852745	342
47	.84639498	1.9372818	319
48	.70392749	1.5906782	331
49	.6966967	1.6936863	333
50	.58181818	1.6357433	330
51	.52698413	1.4915053	315
52	.6097561	1.6367798	328
53	.59558824	1.583488	272
54	.36462094	1.1674531	277
55	.47583643	1.3917306	269
56	.39372822	1.2125635	287
57	.34589041	1.1666398	292
58	.30314961	1.1756043	254
59	.28937729	1.3341377	273
60	.22468354	.81464272	316
61	.22153846	1.0570884	325
62	.19292605	.86588466	311
63	.16376307	.84311928	287
64	.16816817	.92257919	333
65	.16225166	.82938089	302
66	.10447761	.66217954	268
67	.1010453	.62598087	287
68	.04508197	.43746416	244
69	.04388715	.40198243	319
70	.06763285	.40056477	207
71	.01421801	.15364377	211
72	.00772201	.08770449	259
73	.00321543	.0567048	311
74	.04240283	.34462008	283
75	.02068966	.24870503	290
76	.025	.23051291	280
77	.00732601	.08543462	273
78	0	0	265
79	.01083032	.13415745	277
80	.01793722	.21147542	223
81	.00854701	.13074409	234
82	0	0	207
83	0	0	156
84	0	0	174
85	0	0	127
86	0	0	119
87	0	0	97
88	0	0	84
89	0	0	56
90	0	0	47
91	0	0	32
92	0	0	34
93	0	0	15
94	0	0	13
95	0	0	9
96	0	0	11
97	0	0	8
98	0	0	3
99	0	0	4
100	0	0	2
Total	.56056541	1.5447193	19384

Tabla 3 Anexo

. tab PROD14

PROD14	Freq.	Percent	Cum.
0	1,730	50.56	50.56
1	1,692	49.44	100.00
Total	3,422	100.00	

. tab PROD15

PROD15	Freq.	Percent	Cum.
0	1,759	51.40	51.40
1	1,663	48.60	100.00
Total	3,422	100.00	

. tab NCOMPINT

NCOMPINT	Freq.	Percent	Cum.
0	15,962	82.35	82.35
1	1,026	5.29	87.64
2	652	3.36	91.00
3	516	2.66	93.66
4	399	2.06	95.72
5	306	1.58	97.30
6	218	1.12	98.43
7	120	0.62	99.05
8	73	0.38	99.42
9	52	0.27	99.69
10	30	0.15	99.85
11	14	0.07	99.92
12	7	0.04	99.95
13	4	0.02	99.97
14	4	0.02	99.99
16	1	0.01	100.00
Total	19,384	100.00	

.