

Medición y Determinantes de la Brecha Tecnológica en España¹

Leonel Cerno

Departamento de Economía
Universidad Europea de Madrid
C/ Tajo s/n
28670 – Villaviciosa de Odón - Madrid
e-mail: leonel.cerno@uem.es
Tel.: (34) 91 211 56 45

Teodosio Pérez Amaral

Departamento de Fundamentos del Análisis Económico II
Universidad Complutense de Madrid
28223 - Madrid
e-mail: teodosio@ccee.ucm.es
Tel.: (34) 91 394 23 80

En el siguiente artículo abordamos el problema de medir las diferencias de digitalización que existen entre las Comunidades Autónomas de España. Partiendo de esta idea proponemos un índice sintético que permite cuantificar dicha diferencia, utilizándolo de dos maneras: la primera como medida del gasto en equipamiento en la especificación de un modelo para el uso de Internet en cuya estimación las variables más significativas resultan ser el gasto en equipamiento, la edad y el nivel de estudios; y la segunda para explicar los determinantes del nivel de digitalización de las CCAA. Para encontrar una especificación del segundo modelo utilizamos el algoritmo RETINA de Pérez-Amaral, Gallo y White (2003) en donde vemos como se mejora la predicción fuera de la muestra a partir de la inclusión de interacciones y pendientes específicas entre los regresores del modelo original.

Palabras clave: gasto en equipamiento, digitalización, brecha tecnológica, Banda Ancha, modelo de Poisson, elasticidades, conglomerados, RETINA.

JEL-Code: C2, C21, C25

¹ La presente investigación está financiada por el Proyecto Cicyt nº SEJ2004 – 06948.

1. Introducción

Brecha tecnológica es el acceso desigual de las personas a la tecnología digital. Dicha desigualdad podrá ser considerada en distintas dimensiones: geográfica, económica, étnica, por edades, etc.

Actualmente la discusión de la existencia de una brecha tecnológica entre países o entre regiones dentro de cada país se lleva a cabo tanto a nivel descriptivo como a nivel cuantitativo. Preguntas acerca de los motivos de dichas diferencias vienen siendo formuladas hace relativamente poco tiempo, y en cuanto a su medición la mayoría de contribuciones existentes no pasan de ser visiones meramente descriptivas y sin adentrarse en los procesos de medición de los determinantes que lo provocan. A tales enfoques responderían estudios como los que elaboran determinadas organizaciones tanto públicas como privadas (Randstad, Red.es, Fundación AUNA, Banco Mundial u OCDE entre otros). También existen una serie de contribuciones previas a la elaboración de ciertas medidas (Corrocher y Ordanini, 2002; Martin, 1995; Mansell y When, 1998; Ricci, 2000). Estos trabajos intentan cuantificar la difusión del fenómeno digital utilizando metodología cuantitativa.

El presente estudio mide la brecha digital en España y analiza sus determinantes. Comenzamos en la siguiente sección con un análisis descriptivo. En la sección 3 elaboramos un índice de brecha tecnológica para España desde dos puntos de vista: primero a partir del gasto en tecnología y luego desde el punto de vista del grado de digitalización. El primero (índice de gasto en tecnología) es para utilizarlo como regresor en un modelo para el uso de Internet por CCAA. El segundo (digitalización) para utilizarlo como variable endógena en un modelo para sus determinantes. En la sección 4 y 5 ese modelo econométrico de comportamiento para medir el efecto de los factores que influyen en el nivel de digitalización, a nivel regional para toda España. Por las características que presenta la variable endógena, utilizamos primero un modelo de regresión de Poisson para estimar las elasticidades que nos ayudarán a separar la muestra en grupos homogéneos para tratar la heterogeneidad y obtener otro modelo para el grado de digitalización que realice la mejor predicción por fuera de la muestra utilizando el algoritmo RETINA (Relevant Transformation of the Inputs Network Approach) propuesto recientemente por Pérez-Amaral, Gallo y White (2003).

2. Análisis Descriptivo

La brecha digital puede ser medida tanto a nivel macro (continentes, países, etc.) como a nivel micro (regiones, sectores, consumidores, etc.). Generalmente el uso de Internet se considera un buen indicador del uso de la tecnología digital. Informes recientes² revelan que existen actualmente 812 millones de usuarios de Internet en todo el mundo. En el Gráfico 1 podemos observar cómo están distribuidos:

² Disponibles en <http://n-economia.com/frindex.html>

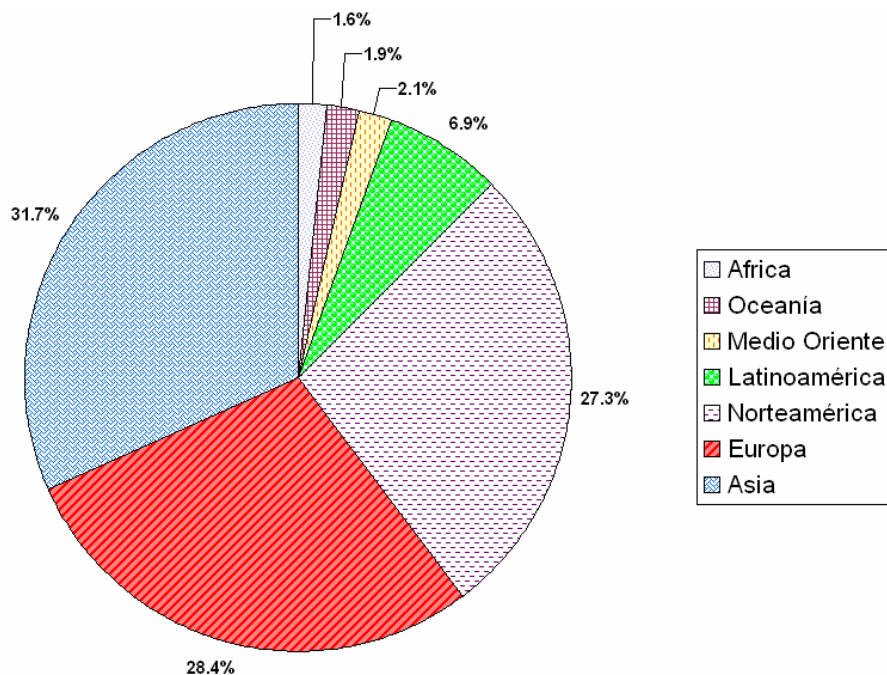


Gráfico 1: Distribución de Usuarios Mundiales de Internet

Fuente: Nielsen/NetRatings (2004) y elaboración propia

Asia concentra el 31,7% de la población on-line mundial, con más de 257 millones de usuarios. Le siguen Europa, con 230 millones de usuarios (28,4%) y Norteamérica con más de 222 millones de usuarios (27,3%). Luego, muy por debajo en cantidad le siguen Latinoamérica con casi 56 millones de usuarios (6,9%), Medio Oriente con 17 millones (2,1%), Oceanía³ con más de 15 millones (1,9%) y por último África con el menor número de usuarios (1,6%) equivalente a casi 13 millones de personas.

Centrándonos en el continente europeo podemos apreciar que la distribución de los 230 millones de usuarios es bastante desigual. Existen países sumamente avanzados en cuanto al uso de Internet y otras TIC y otros que apenas están surgiendo, entre los que se encontraría España. En el Gráfico 2 vemos que países como Suecia u Holanda están a la cabeza en la UE con un 74,6% y 66,5% de población usuaria de Internet. España se encuentra entre los países con brecha más pronunciada (sólo el 34,2% de la población en 2004 declara utilizar Internet), aunque es un porcentaje superior al de países como Grecia (15,3%), u otros recién ingresados en la Unión (Hungría 23,7%, Polonia 23,5% o Lituania 20,2%).

³ En términos relativos este continente estaría a la altura de Asia, Europa o Norteamérica si consideramos que entre Australia y Nueva Zelanda en conjunto no llegan a 26 millones de habitantes.

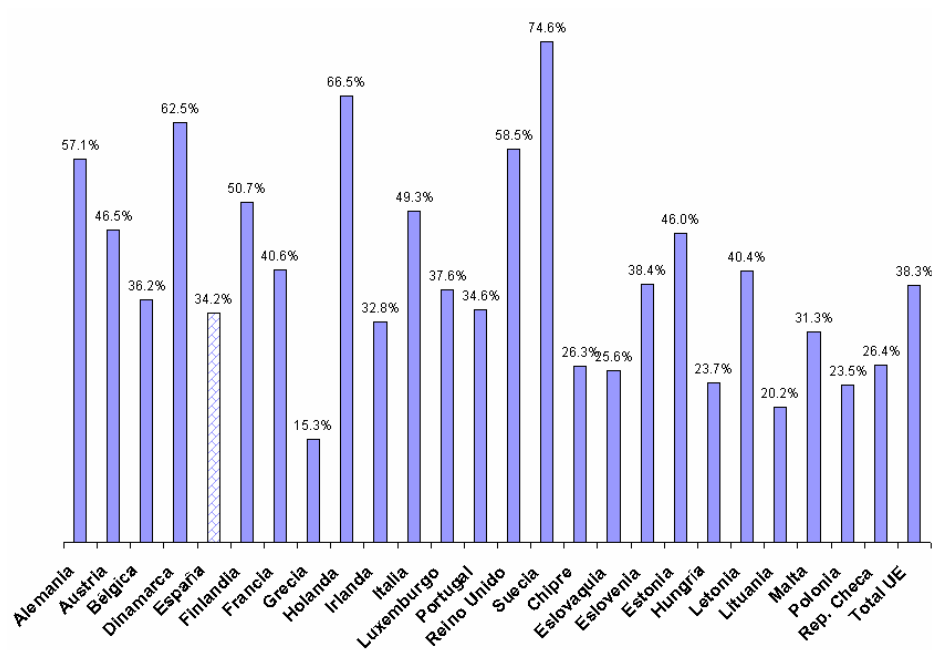


Gráfico 2: Distribución de Usuarios de Internet en la UE

Fuente: Nielsen/NetRatings (2004) y elaboración propia

España es un país en donde existe una diferencia considerable en cuanto al uso de las nuevas tecnologías (TIC). En un informe reciente (Red.es, diciembre 2004) en donde se comparan datos de uso de TIC entre regiones del país y con el resto de Europa, se destacan datos tales como que sólo el 23,5% de las empresas españolas utiliza el comercio electrónico, que no todas las empresas acceden a Internet (el 76,1% tiene acceso a Internet por Banda Ancha), o que la aportación al PIB del mercado de las TIC en España no supera el 1,68% en 2004. Sin embargo dicha brecha tecnológica no se ve tan claramente al realizar la comparación por regiones. Cataluña, Madrid y el País Vasco tienen niveles de convergencia tecnológica con las regiones más desarrolladas de Europa (más del 89% de sus empresas compran y venden a través de la Red), mientras que en Galicia y Extremadura no se alcanza al 21 % de sus empresas.

En cuanto al acceso a Internet desde los hogares, las diferencias entre regiones también son bastante evidentes como observamos en el Gráfico 3 tomado de Cerno y Pérez-Amaral (2005):

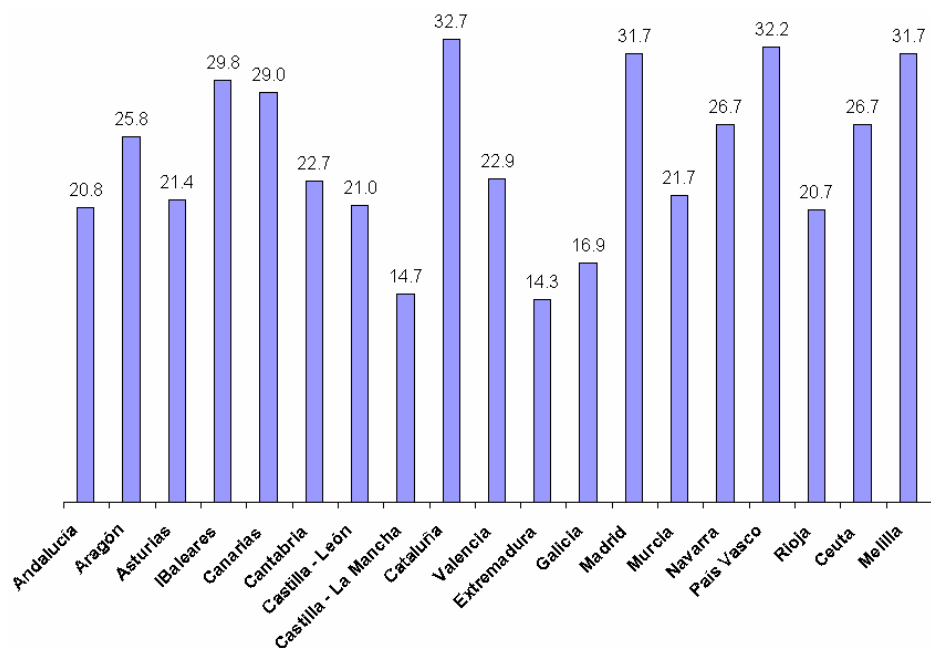


Gráfico 3: Distribución del Acceso a Internet en los Hogares Españoles
Fuente: INE

Si bien en los hogares españoles se experimentó una tasa de crecimiento media anual de 36,4% durante el período 2000-2004 (Red.es, 2004), en 2003 el 25,2% de los hogares ya cuenta con acceso a Internet (Cerno y Pérez-Amaral, 2005), dicho crecimiento fue dándose de manera desigual. En el Gráfico 3, vemos que Madrid, Cataluña y País Vasco son las regiones con mayor acceso a Internet.

A continuación ampliamos el análisis al resto de las tecnologías digitales. Esto pondrá en evidencia las diferencias existentes a nivel geográfico. A su vez, revela la necesidad de estudiar dichas diferencias, cuantificarlas a través de un índice y medir el efecto de los posibles determinantes.

3. Índice Sintético de Brecha Tecnológica Geográfica para España

Un concepto de *digitalización* es el que la define como la tendencia de las tecnologías de la información y la comunicación a converger, es decir, a interconectarse. Esta estandarización es la que va a permitir que una vez que los entes (individuos o empresas) estén conectados, sus sistemas de información se entiendan y puedan interactuar entre sí⁴.

Al iniciar la presente investigación nos hemos encontrado con que el concepto de brecha tecnológica es poco homogéneo dentro de la literatura. Las contribuciones realizadas hasta ahora en su mayoría usan diferentes medidas. Nos planteamos entonces cómo medir el grado de digitalización en las zonas geográficas españolas y cómo dicha medición puede ser utilizada como herramienta para orientar políticas.

Existen algunos puntos importantes a tener en cuenta relacionados con la agregación de elementos y con las ponderaciones para construir un índice sintético. El proceso podría

⁴ www.dmconsulting.es/digitalizacion.htm

seguir el criterio de ponderar de manera objetiva o basarse en ponderaciones subjetivas de acuerdo a criterios de importancia considerados de manera predeterminada. Para identificar los niveles de agregación y la provisión de las ponderaciones tenemos que pensar que la brecha tecnológica se está refiriendo a una *diferencia* entre las distintas unidades (en nuestro caso Comunidades Autónomas de España), entonces la medición a realizar deberá reflejar esta diferencia. Esto nos lleva al planteamiento de qué elementos se considerarán y qué ponderaciones se incluirán. El detalle de la construcción del índice se encuentra en el apéndice al final.

A continuación se presentan los valores del índice por CCAA. Los puntos de vista planteados en los pesos de importancia se refieren al gasto en equipamiento y tecnología (ITIC), y al nivel de digitalización (IDD). Este último cuantifica la predisposición del individuo hacia las nuevas tecnologías. Lo hace a partir de la asignación de pesos de importancia a los distintos grupos de tecnologías considerados en la construcción del índice dándole el mayor peso a Internet (para mayor detalle ver apéndice):

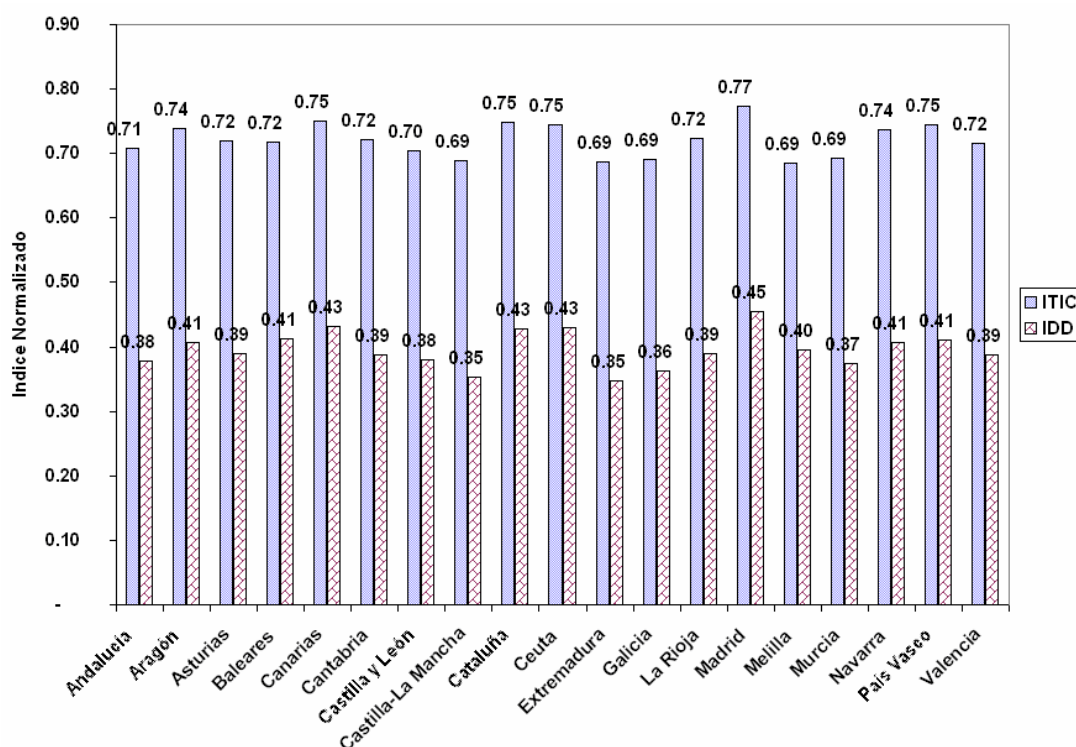


Gráfico 4: Índices Normalizados de Gasto en Tecnología (ITIC) y Nivel de Digitalización (IDD)

Vemos que existen diferencias a pesar de que el rango de variación aparentemente no es importante tanto en uno como en otro enfoque del índice. Sabiendo que valor 1 es el valor máximo, se observa que Madrid es la CCAA con mayor consumo de tecnología y con mayor nivel de digitalización, aunque en este último caso con salvedades, ya que vemos que ninguna CCAA llega siquiera al 50% del nivel considerado. Esto estaría relacionado con lo que vimos más arriba al comparar España con otras naciones europeas en donde se veían diferencias importantes en cuanto al uso de Internet, teniendo en cuenta de que es el factor más importante en la construcción del índice IDD.

Centrándonos en las CCAA, vemos que Melilla gasta menos en tecnología que Extremadura, pero sin embargo la supera bastante en el ranking de nivel de

digitalización. Otras CCAA como Cataluña o el País Vasco o como Extremadura o Castilla-La Mancha se mantienen en niveles altos y bajos en nivel de gasto en tecnología y nivel de digitalización respectivamente.

3.1. Comparación de Índices y Descubrimiento de la Brecha entre CCAA

El paso siguiente es observar y medir la brecha digital entre CCAA, definiendo esta como la diferencia existente entre las mismas en cuanto a gasto en tecnología y nivel de digitalización. Después de calcular niveles de gasto y digitalización para cada CCAA, utilizamos la media del índice como valor de referencia, obtenida a partir de los valores normalizados observados anteriormente. A partir de ahí medimos la distancia que cada Comunidad tiene respecto de esa media y entre ellas mismas. Esto tendrá las siguientes implicaciones:

- Un valor de brecha tecnológica “total” puede ser medida a partir de alguna medida de dispersión relativa tal como el coeficiente de variación.
- La brecha tecnológica por CCAA puede ser medida a partir de las distancias de su valor normalizado del índice con el resto de los valores o con la media total. Esta distancia representa la “similitud” o “diferencia” para cada CCAA, ya sea en cuanto a gasto como también a nivel de digitalización.
- Aunque por el tipo de observaciones que utilizamos no lo haremos en este artículo, sin embargo la evolución de la brecha tecnológica es susceptible a ser medida a través del tiempo, tanto a nivel total como a nivel de CCAA.

A continuación presentamos dos gráficos en donde observamos las diferencias existentes entre la media del índice y los valores por CCAA (valores tipificados):

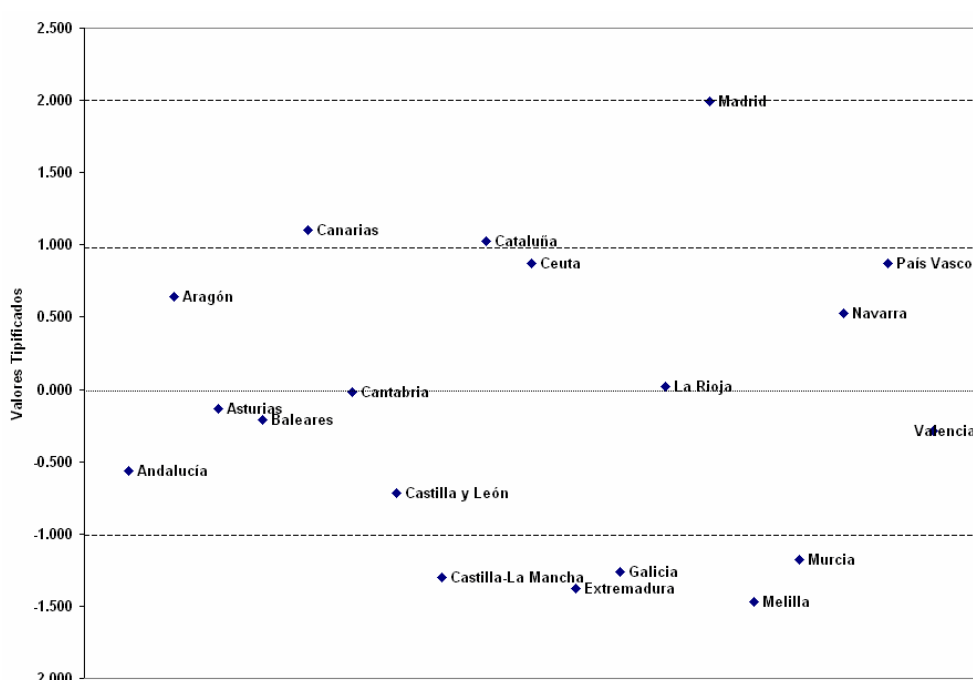


Gráfico 5: Comparación entre CCAA del Gasto en Equipamiento

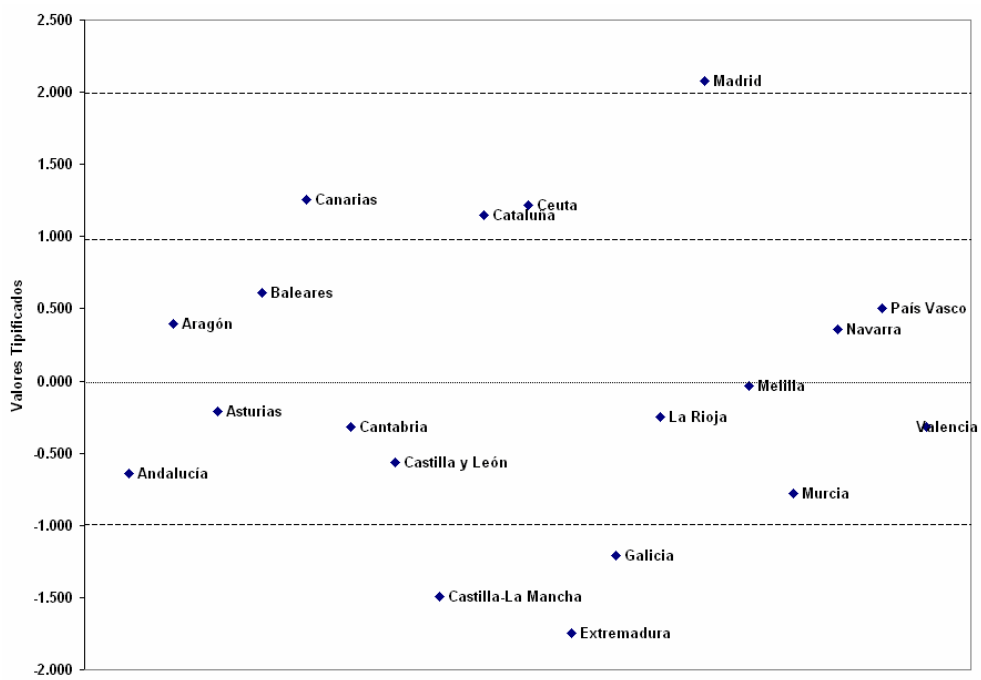


Gráfico 6: Comparación entre CCAA del Nivel de Digitalización

En la siguiente tabla están representados los valores de las diferencias observadas en ambos gráficos:

Tabla 1. Diferencia entre Valores del Índice con Respecto a la Media del Total

CCAA	Var. ITIC	Var. IDD
Andalucía	-11,2%	-1,8%
Aragón	13,6%	1,1%
Asturias	-3,9%	-0,6%
Baleares	-5,9%	1,7%
Canarias	17,5%	3,5%
Cantabria	3,8%	-0,9%
Castilla-León	-12,9%	-1,6%
Castilla-La Mancha	-17,8%	-4,2%
Cataluña	16,9%	3,2%
Ceuta	15,6%	3,4%
Extremadura	-18,3%	-4,9%
Galicia	-17,5%	-3,4%
La Rioja	5,0%	-0,7%
Madrid	23,1%	5,8%
Melilla	-18,8%	-0,1%
Murcia	-16,9%	-2,2%
Navarra	12,4%	1,0%
País Vasco	15,6%	1,4%
Valencia	-7,4%	-0,9%
Media	0,721	0,396
Desv. Típica	0,026	0,028
Coef. De Variación	0,0358	0,0707

De lo visto podemos decir:

- La media del gasto en equipamiento (0,721) tiene un valor mayor que la del nivel de digitalización (0,396). Es decir que en general se tendrá mayor peso para la renta destinada a equipamiento que para la afinidad hacia la tecnología digital.
- La brecha tecnológica en España en cuanto a porcentaje del ingreso destinado a equipamiento (ITIC) es menor que la existente en cuanto a nivel de digitalización y afinidad con las nuevas tecnologías (0,0358 es menor que 0,0707). Esto se debe a que un incremento en el nivel de renta per capita como el que experimentaron las regiones de España luego de su ingreso en la UE se tradujo, como es lógico, en un mayor gasto en equipamiento. Sin embargo el nivel de digitalización tiene otros determinantes además de la renta y tal como se estudia en detalle en las secciones precedentes de este estudio.
- Madrid es la CCAA que más gasta en equipamiento y que mayor nivel de digitalización tiene respecto al resto. La diferencia existente entre Madrid y el valor medio supera las dos desviaciones típicas positivas en ambos casos según podemos observar en los gráficos 5 y 6. La diferencia respecto a la media en ITIC es positiva de 23,1%. Respecto a IDD es también positiva, aunque bastante menor (5,8%). Es decir que Madrid es la CCAA más favorecida desde ambos puntos de vista del índice.
- Luego de Madrid, las CCAA que más favorecidas en cuanto a renta destinada a equipamiento son Cataluña y Canarias, con diferencias positivas respecto a la media total de 16,9% y 17,5% respectivamente. Le siguen cercanamente el País Vasco y Ceuta, ambos con 15,6%. Los valores de las diferencias en cuanto a nivel de digitalización serán también positivos de 3,5% para Canarias y de 3,2% para Cataluña. Para Ceuta la diferencia es de 3,4%. El País Vasco está más cercano a la media con una diferencia positiva de sólo 1,4%.
- Como contrapartida, las CCAA más desfavorecidas en cuanto a gasto en equipamiento y nivel de digitalización son Castilla-La Mancha, Galicia, Extremadura. Las diferencias tanto para el gasto en equipamiento como en el nivel de digitalización serán por debajo de la media en -17,8% y -4,2% para Castilla-La Mancha, -18,3% y -4,9% para Extremadura y en -17,5% y -3,4% para Galicia.
- El caso de Melilla o Murcia es especialmente interesante: los niveles de gasto en equipamiento de estas CCAA está entre los más bajos (-18,8% y -16,9%), pero su nivel de digitalización es más cercano a la media. (-0,1% y -2,2% respectivamente).
- En cuanto al gasto en equipamiento del resto de las CCAA, se encuentran dentro de la banda de más-menos una desviación típica, algunas muy cercanas a la media (Asturias, Baleares, Cantabria y La Rioja), otras más o menos favorecidas (dentro de las que se encuentran por encima de la media las que más destacan son Aragón y Navarra, y de las por debajo a Andalucía y Castilla y León).

- El cuanto al nivel de digitalización del resto de las CCAA, la dispersión es mayor. Exceptuando a Melilla con un valor muy cercano a la media, las Comunidades se diferencian por encima o por debajo de manera más destacada tal como observamos en el Gráfico 6.

A continuación comparamos en un gráfico bidimensional ambos índices. Al realizar esta comparación profundizamos más detalladamente en la idea de las diferencias tecnológicas existentes en España debido a que no todas las CCAA tienen ambos valores igualmente altos o bajos.

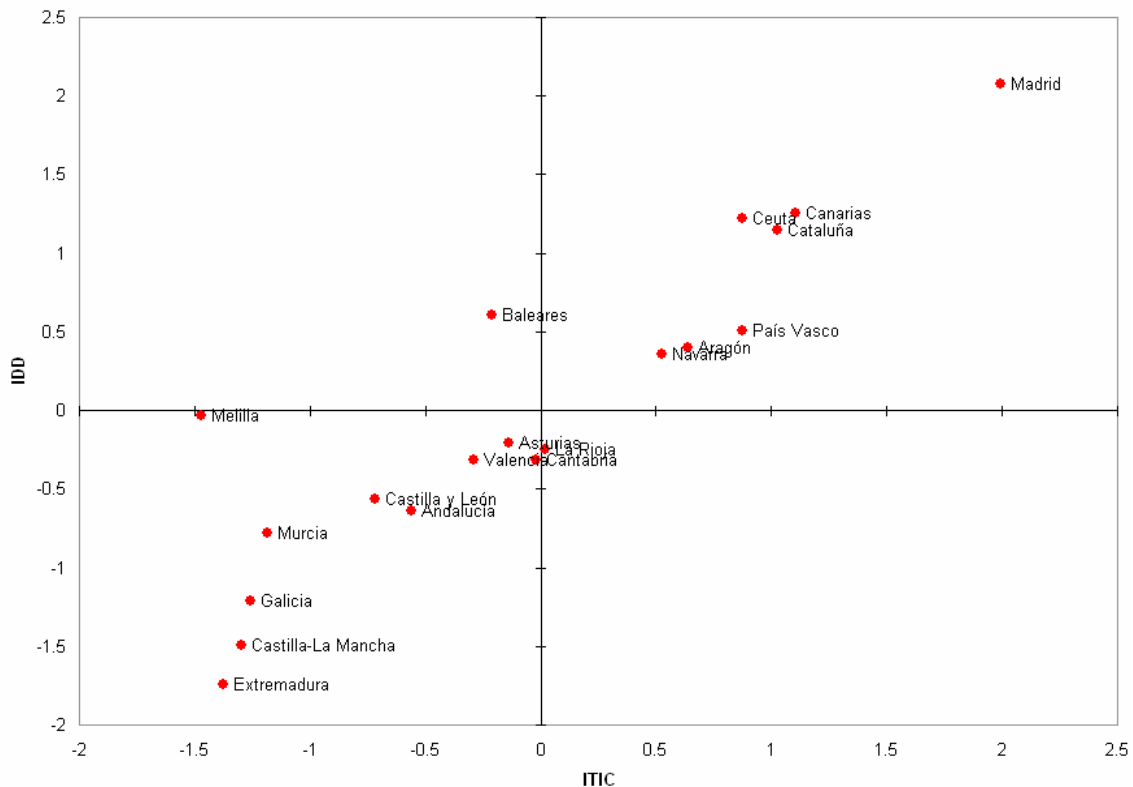


Gráfico 7: Diferencias entre CCAA

Las CCAA que se encuentran en el cuadrante derecho superior serán, en una visión general y a distintos niveles, las más favorecidas en tecnología, mientras que las que se encuentran en el inferior izquierdo serán lo contrario, también en general y a distintos niveles. Los casos “atípicos” de Melilla y Baleares se pueden explicar en que en cuanto a gasto en tecnología han llegado a una especie de punto de saturación que hace que sea especialmente bajo, mientras que en nivel de digitalización están en o por encima de la media.

También es posible realizar un análisis similar a partir de las matrices de distancias de ambos enfoques del índice entre las distintas CCAA. Para ello se deberían aplicar técnicas de análisis multivariante tales como el escalamiento multidimensional o los biplots, pero nuestro objetivo ahora es explicar los determinantes que hacen que existan estas diferencias. Para ello, en la sección siguiente calcularemos los impactos de los determinantes del uso de Internet haciendo el análisis por CCAA.

4. La Brecha Tecnológica como un Determinante del Uso de Internet

En el intento de modelizar la brecha tecnológica o para tomar decisiones de política, es necesario comenzar por las incidencias en los costes y los beneficios asociados con el uso de Internet.

De todos los servicios considerados, Internet es la que posee mayor potencial para reducir las distancias entre el grado de digitalización de distintas regiones. Esto es así debido a que el acceso y uso de Internet traen aparejados el establecimiento de una red para que sea posible tal conexión, además del uso de distintas forma de conexión ya sea a través de ordenadores, teléfonos móviles, etc. A partir de esto, comenzamos por modelizar los determinantes del uso de Internet, separando la muestra por CCAA.

La literatura empírica que analiza los determinantes de la demanda de acceso y uso de Internet es relativamente extensa, por ejemplo en Madden, Savage y Simpson (1996), Cassel (1999), Rappoport, Taylor y Kridel (2002) o Cerno y Pérez-Amaral (2005) entre otros, se destacan las implicaciones de Internet para el comercio, la educación y las relaciones personales.

Para elaborar un modelo empírico que mida los determinantes de la brecha tecnológica a nivel regional en España, usamos la base de datos de la encuesta del INE TIC-2003 realizada entre los meses de febrero y marzo de 2003 sobre acceso y uso de Nuevas Tecnologías en España. La información la hemos organizado por CCAA, 18 en total, en donde tomamos a Ceuta y Melilla como una sola de ellas.

Con la información del número de lugares en los que el individuo utiliza Internet construimos la variable endógena (USE_i) que toma valores en el intervalo discreto $[0, 4]$. Podemos observar que lo que más abunda son los valores nulos:

Tabla 2: Estadísticos del Número de Lugares de Uso de Internet por los Individuos

Frecuencias Absolutas					
0	1	2	3	4	Total
13.705	3.093	1.776	329	45	18.948

Media	Mediana	Desviación Típica	Coefficiente de Variación	Simetría	Curtosis
0,412	0,000	0,750	181,9	1,821	2,716

Fuente: Elaboración propia.

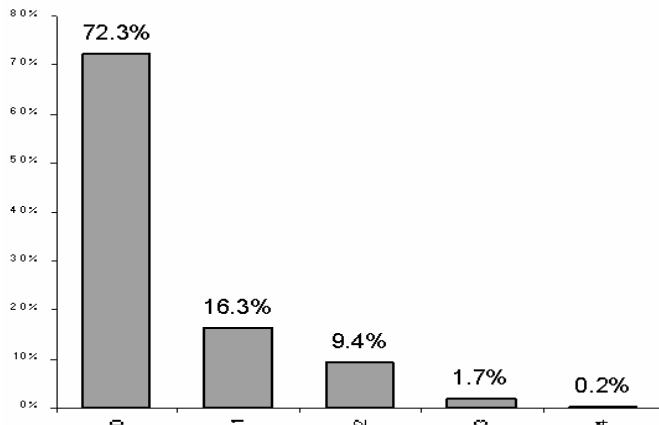


Gráfico 8: Frecuencias Relativas de la Variable USE_i

A primera vista vemos que más de la mitad de la población se declara no usuaria de Internet (72,3%). El 27,7% restante sí utiliza el servicio dentro del intervalo de 1 a 4 lugares distintos (Cerno y Pérez-Amaral, 2005).

Por el análisis previo que hicimos al intentar medir la brecha tecnológica, y por estar utilizando datos de corte transversal, es posible de que exista heterogeneidad. De no tenerla en cuenta en el modelo, nos podría llevar a resultados contradictorios y estimaciones sesgadas e inconsistentes. A partir de esta idea entonces, el análisis lo haremos a nivel de CCAA considerando que equivalen a submuestras homogéneas.

A continuación especificamos y estimamos un modelo para el uso de Internet y luego evaluamos las elasticidades del uso de Internet con respecto a las variables explicativas más significativas.

4.1. Especificación

Teniendo en cuenta que la variable USE_i es una variable de recuento (de naturaleza discreta), los valores pequeños que toma y la abundancia de ceros que tiene, a continuación especificamos un modelo que tenga en cuenta todas esas características. El más utilizado para datos de este tipo es el modelo de regresión de Poisson⁵. El modelo supondrá que cada y_i es una realización de una variable aleatoria con distribución de Poisson de parámetro λ_i y que el mismo esté relacionado con un conjunto de regresores x_i . La ecuación básica puede escribirse en términos de probabilidad tal como sigue:

$$\Pr(y = y_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} \quad (0.1)$$

$$y_i = [0, 4]$$

⁵ La distribución de Poisson es muy utilizada cuando el análisis estadístico paramétrico se realiza para variables aleatorias referidas a recuentos de sucesos, es decir al número de veces que un evento ocurre. Como la variable que pretendemos explicar se refiere al *número* de sitios de uso de Internet, lo especificamos con una distribución de probabilidad de número de ocurrencias. Para más detalles véase Davidson y MacKinnon, 2003.

Se puede demostrar que las probabilidades totales suman 1 con $y = 0, 1, 2, 3, 4$ y que la media y la varianza serán las mismas e iguales a λ que deberá tomar siempre valores no negativos (Davidson y MacKinnon, 2003). El parámetro que representa la media condicionada a los regresores de la variable endógena con distribución de Poisson es $\lambda_i = E(y_i|x_j) = V(y_i|x_j) = e^{x_j\beta}$, siendo de esta forma lo más habitual la formulación logarítmica lineal, que tiene la ventaja de ser lineal en los parámetros:

$$\ln \lambda_i = x_j' \beta = \beta_0 + \beta_1 x_{1j} + \dots + \beta_k x_{kj} \quad (0.2)$$

Los coeficientes β son interpretados como semielasticidades (excepto que los regresores sean en logaritmos), ya que $\beta_i = \partial \ln E(y_i|x_j) / \partial x_{jk}$ con $j = 1, 2, \dots, k$. Las variables explicativas utilizadas serán de cuatro tipos:

Variable referida al gasto en tecnología: Su efecto está cuantificado por la variable *ITIC*, y que representa una *proxy* de gasto en tecnología del individuo. Su elaboración está explicada en detalle en el apéndice y su utilidad en el apartado anterior. En la sección anterior a esta variable la propusimos como un índice para reflejar el grado de participación en el mercado de equipamiento digital a nivel individual y la posición en el mismo. En efecto, los valores de este indicador serán mayores cuando mayor sea el gasto en tecnología.

Variables referidas a la demanda de servicios complementarios: Estas se refieren a tener acceso a Internet en la vivienda y al equipamiento utilizado para conectarse. El acceso a Internet en la vivienda está recogido por la variable ficticia *HOUSINT* y el equipamiento para conectarse está expresado en la variable *EQ*.

Principales variables sociodemográficas: Estas son las referidas al sexo (*MALE*), la edad (*AGE*), el hábitat (*HABITAT*) y el nivel de estudios (*STUDYLEVEL* y *BESTUDYING*).

Variable endógena: Es la cantidad de lugares en donde la persona se conecta. La variable es *USE_i* y varía entre 0 y 4.

Para más detalles ver la tabla de Definición de Variables del apéndice.

4.2. Estimación y Resultados

El modelo para el uso de Internet para explicar sus determinantes para cada CCAA que nos planteamos es aquel que sigue la siguiente relación:

$$USO_i = f(\text{gasto en tecnología}_i, \text{equipamiento}_i, \text{atributos sociodemográficos}_i) + error_i$$

El gasto en tecnología se refiere a una de las formas planteadas de índice de brecha tecnológica y en esta especificación econométrica representa al gasto en bienes complementarios. Nuestro análisis por CCAA será intentando captar las diferencias

entre las mismas considerando la heterogeneidad. El modelo a estimar por máxima verosimilitud es:

$$USE_{ik} = \beta_{0k} + \beta_{1k}ITIC_{ik} + \beta_{2k}HOUSINT_{ik} + \beta_{3k}EQ_{ik} + \beta_{4k}MALE_{ik} + \beta_{5k}AGE_{ik} + \beta_{6k}HABITAT_{ik} + \beta_{7k}STUDYLEVEL_{ik} + \beta_{8k}BESTUDYING_{ik} \quad (0.3)$$

donde $USO_i = \ln \lambda_i$, el subíndice i representa al tamaño muestral de la CCAA, y el subíndice $k = 1, 2, \dots, 18$ a la CCAA considerada.

Los resultados detallados de cada regresión se pueden revisar en el apéndice, en donde vemos que las 18 ecuaciones tienen un valor de bondad de ajuste muy satisfactorio, lo que aleja cualquier sospecha de sobredispersión. A continuación presentamos en una tabla las elasticidades del uso de Internet con respecto al gasto en equipamiento, la edad y el nivel de estudios (entre paréntesis los estadísticos t):

Tabla 3: Elasticidades del Uso de Internet por CA

COMUNIDAD AUTÓNOMA	Tamaño de la Muestra	Elasticidades		
		Gasto	Edad	Nivel de Estudios
Andalucía	2.147	1,91 (5,07)	-1,80 (11,44)	1,21 (13,41)
Aragón	996	2,77 (4,28)	-2,02 (8,71)	1,34 (9,61)
Asturias	893	2,38 (3,83)	-1,80 (7,62)	1,10 (7,36)
Baleares	579	1,63 (2,44)	-1,63 (6,13)	1,11 (6,75)
Canarias	787	1,93 (3,38)	-1,65 (8,25)	1,01 (7,92)
Cantabria	544	1,30 (1,70)	-1,75 (5,15)	1,43 (7,00)
Castilla León	1.304	1,35 (2,82)	-1,88 (8,70)	1,10 (9,02)
Castilla La Mancha	1.012	2,42 (3,86)	-2,13 (7,53)	1,02 (7,71)
Cataluña	1.833	1,83 (4,53)	-1,58 (11,29)	0,81 (9,37)
Valencia	1.377	2,42 (4,76)	-1,76 (8,97)	1,17 (10,31)
Extremadura	831	1,55 (2,36)	-2,17 (6,79)	0,98 (6,34)
Galicia	1.171	2,13 (4,09)	-2,26 (9,45)	1,29 (10,02)
Madrid	1.567	2,11 (4,89)	-1,42 (10,99)	1,24 (11,86)
Murcia	773	2,97 (4,60)	-2,12 (7,38)	0,90 (5,79)
Navarra	1.202	1,49 (3,06)	-1,73 (9,30)	1,28 (10,32)
País Vasco	1.181	1,89 (2,78)	-1,50 (6,84)	1,05 (8,48)
La Rioja	549	1,44 (1,82)	-2,12 (6,94)	1,30 (7,02)
Ceuta-Melilla	203	2,16 (2,03)	-1,41 (3,41)	0,94 (2,40)

En el cuadro de arriba Podemos observar como varían las elasticidades por CCAA y vemos que dependerá de según qué variable estemos considerando. A continuación observamos tres gráficos comparativos de los valores de las elasticidades, ordenadas estas de menor a mayor, en donde podemos ver que las similitudes o diferencias en elasticidades entre CCAA dependerá de la variable a considerar:

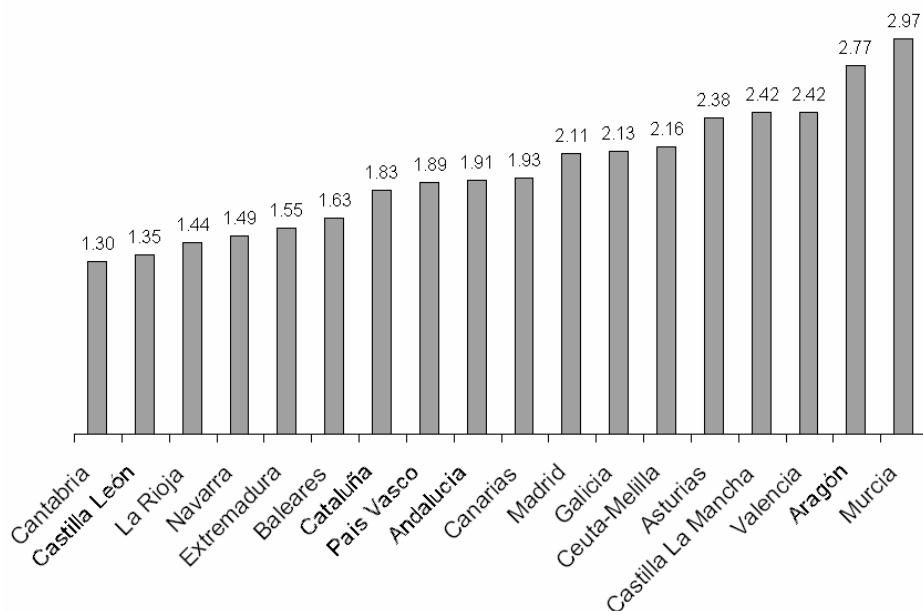


Gráfico 9: Elasticidades del uso de Internet con respecto al gasto en equipamiento

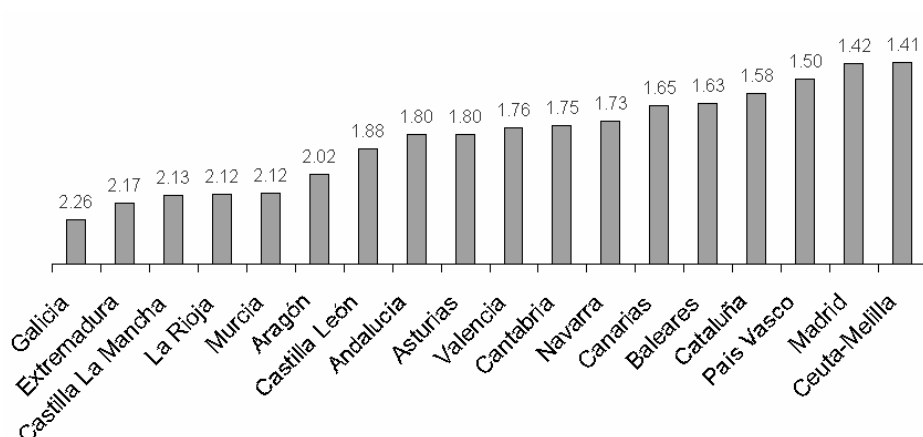


Gráfico 10: Elasticidades del uso de Internet con respecto a la edad (valores absolutos)

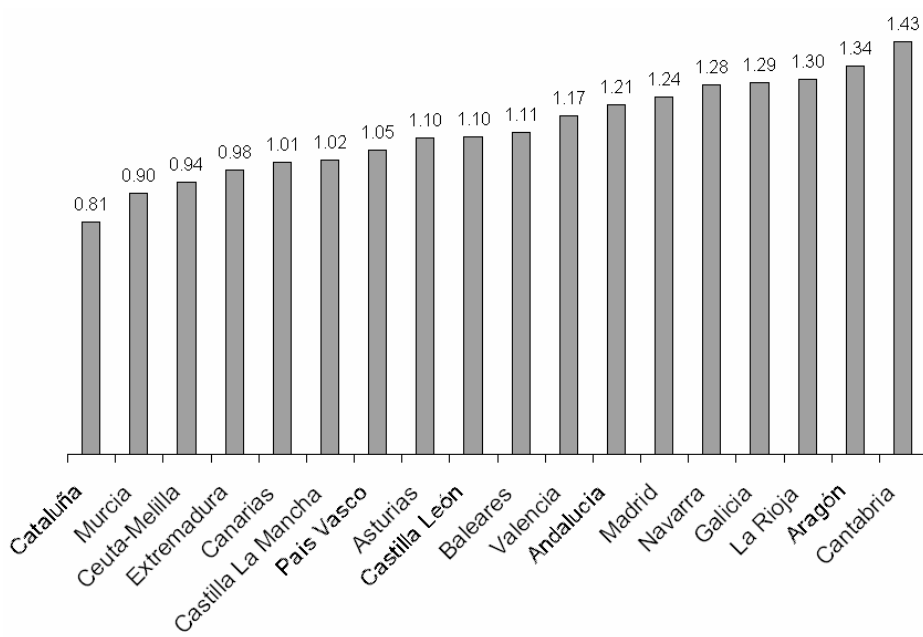


Gráfico 11: Elasticidades del uso de Internet con respecto al nivel de educación

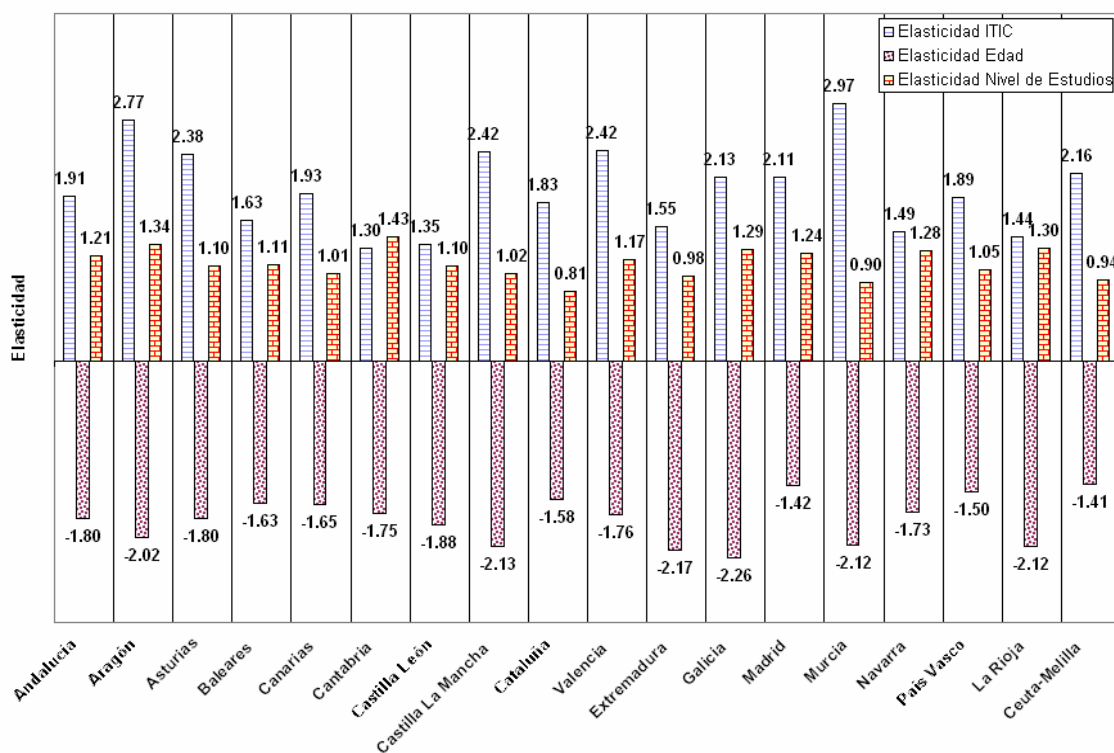


Gráfico 12: Elasticidades del uso de Internet (conjunto)

Aquí observamos que Murcia es la CCAA con mayor elasticidad de uso de Internet con respecto al gasto en tecnología con un valor de 2,97, esto quiere decir que si en Murcia aumenta el gasto individual en equipamiento un 1%, el uso de Internet se incrementará un 2,97%. Para el caso de Comunidades como Madrid o Cataluña, en donde anteriormente se veía que el gasto en equipamiento era de los primeros del ranking, dicho impacto es de 2,11% y 1,83% respectivamente. Pero observamos que en general el uso de Internet tendrá elasticidades mayores que uno respecto al gasto en equipamiento.

Por otro lado, observamos que en todos los casos la edad es una variable que tiene un impacto inverso en el uso de Internet. Las elasticidades serán todas negativas, encontrándose los mayores impactos en las CCAA más atrasadas en cuando a nivel de digitalización como hemos visto que serían Galicia, Extremadura o Castilla-La Mancha con valores de menos 2,26%, 2,17% y 2,13% respectivamente. Ceuta-Melilla, Madrid y País Vasco serán las CCAA cuyas elasticidades de la edad estarían en el otro extremo, con valores de -1,41, -1,42 y -1,50 y respectivamente.

En cuanto al nivel de educación, es interesante observar que Cataluña, una de las CCAA con mayor nivel de digitalización, tenga el menor valor de elasticidad del uso de Internet con respecto al grado de educación adquirido (0,81). Esto quiere decir que de aumentar 1% el nivel de educación adquirido a nivel agregado, el incremento en el uso de Internet sería menos que proporcional, de 0,81%. Algo parecido sucedería con Ceuta-Melilla o Murcia, ambas regiones con valores por debajo de la unidad (0,94 y 0,90 respectivamente). Entre estas se encuentra Extremadura, la CCAA con menor grado de digitalización y con un valor de 0,98. Las regiones con elasticidades mayores son aquellas de nivel de digitalización medio como Cantabria, Aragón y La Rioja, con elasticidades de 1,43, 1,34 y 1,30 respectivamente.

4.3. Heterogeneidad

Modelizar los determinantes del nivel de digitalización que provoca la brecha tecnológica en España requerirá tener en cuenta la heterogeneidad que hemos observado entre las distintas CCAA. En el apartado anterior hemos realizado regresiones por CCAA y hemos calculado elasticidades. Siguiendo este criterio, a continuación agruparemos a las CCAA según sus elasticidades, intentando de esta manera identificar en dichos grupos comportamientos similares en algún sentido⁶. Para lograr este objetivo utilizamos el método de Ward de agrupamiento jerárquico por variables, también denominado Análisis de Conglomerados. Con este método partimos desde las medias de los subgrupos definiendo una medida global de heterogeneidad que es la suma de las distancias euclídeas al cuadrado entre cada elemento y la media de su grupo:

$$W = \sum_{g=1}^G \sum_{i=1}^{n_g} (x_{ig} - \bar{x}_g)(x_{ig} - \bar{x}_g)^T \quad (0.4)$$

El criterio comienza suponiendo que cada dato forma un grupo, por tanto $W = 0$. Luego se unen los elementos más próximos de manera tal de que se produzca el mínimo incremento en W . Entonces en la siguiente etapa tendremos $n-2$ grupos de un elemento y uno de dos elementos. El algoritmo seguirá hasta unir todos los grupos en uno solo. En la representación gráfica (dendrograma) se determina el corte a un nivel de distancia dado y así obtenemos una clasificación del número de grupos existentes y los elementos que lo forman. Este tipo de análisis es muy utilizado y muy útil para explorar datos y detectar subgrupos homogéneos dentro del total de la muestra, determinando previamente el grado de similitud que se prefiere entre los elementos de cada grupo. En nuestro caso la evidencia es bien clara al observar el próximo Gráfico 12, en donde

⁶Agrupar por elasticidades es uno de los tantos criterios existentes para agrupar por CCAA. Por ejemplo, también se podría agrupar por efecto marginal del regresor o por ratio de incidencia, aunque los resultados terminarían siendo prácticamente los mismos.

vemos que se forman tres grupos de CCAA teniendo en cuenta el grado de homogeneidad (el resto de los resultados en el apéndice):

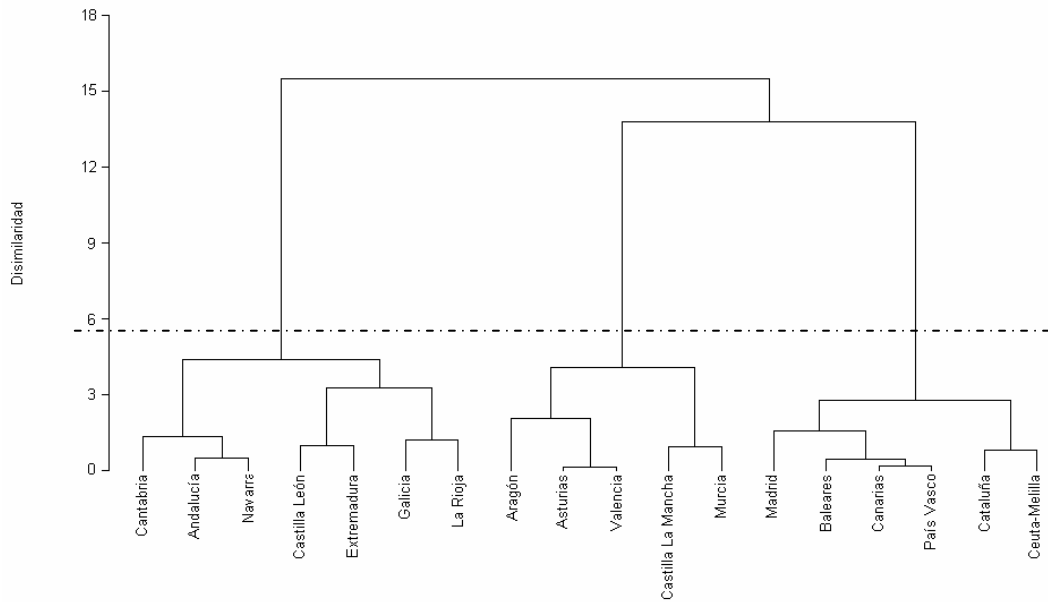


Gráfico 13: Árbol jerárquico de las CCAA a partir de las tres elasticidades calculadas

En este gráfico vemos que la partición se la puede realizar a un nivel bajo de agregación, observando la existencia clara de tres grupos dentro de los cuales existe una cierta similitud. A partir de aquí los grupos formados serán los siguientes:

Tabla 4: Conglomerados de CA

GRUPO 1: Menos Sensibles	Cantabria	Extremadura
	Andalucía	Galicia
	Navarra	La Rioja
	Castilla y León	
GRUPO 2: Sensibilidad Media	Aragón	Castilla – La Mancha
	Asturias	Murcia
	Valencia	
GRUPO 3: Más Sensibles	Madrid	País Vasco
	Baleares	Cataluña
	Canarias	Ceuta – Melilla

Entonces vemos la existencia de tres grupos diferenciados por su mayor o menor predisposición a las nuevas tecnologías, sea por el consumo, el uso o la misma afinidad. En la sección siguiente los consideraremos individualmente a cada uno en el modelo de los determinantes de la brecha tecnológica.

5. Determinantes del Grado de Digitalización en España

En el apartado anterior hemos detectado tres grupos caracterizados por su mayor o menor elasticidad de uso de Internet respecto a tres determinantes del mismo. A continuación analizaremos los determinantes de la brecha tecnológica en España utilizando un modelo de regresión lineal para explicar el índice de brecha tecnológica,

con este último tomando la forma que mide el grado de digitalización (IDD_i). Siguiendo los mismos pasos planteados en Marinucci y Pérez-Amaral (2005) y en Pérez-Amaral, Gallo y White (2005), comenzaremos con la especificación de un modelo lineal de referencia denominado modelo *Básico*. Se trata de un modelo en el que no se utiliza ninguna estrategia particular de selección y que es el punto de partida para encontrar la mejor especificación para la predicción fuera de la muestra. Luego utilizamos el algoritmo RETINA⁷ de Pérez-Amaral, Gallo y White (2003) para encontrar la mejor especificación que denominaremos modelo *Predictivo*. Este permitirá la inclusión de no linealidades en los regresores controlando siempre el nivel de multicolinealidad y su estimación solo la utilizaremos para realizar predicciones.

Suponemos que la forma funcional del modelo explicativo de la brecha tecnológica en España es desconocida a priori, pero RETINA nos ayudará a encontrar la especificación parsimoniosa del modelo que mejor prediga la variable endógena fuera de la muestra.

Suponemos que la especificación del modelo básico es la siguiente:

$$\ln IDD_i = \beta_0 + \beta_1 \ln EQ_i + \beta_2 \ln STUDYLEVEL_i + \beta_3 \ln AGE_i + \beta_4 IMILLS_i + \beta_5 MALE_i + \beta_6 HOUSINT_i + \sum_{h=1}^7 \gamma_h SIZE_{ih} + \sum_{m=1}^5 \phi_m WORK_{im} + u_i \quad (0.5)$$

Recordemos que IDD se refería al índice que mide el nivel de digitalización. Es un indicador que se construye a partir de establecer pesos de importancia a una canasta de tecnologías, en donde el peso más importante lo tiene el servicio de Internet (detalle de su construcción en el apéndice). Esta variable la hemos analizado en detalle a nivel de CCAA en el apartado 2. Aquí la analizamos a nivel individual, en donde $i = 1, \dots, n$, se refiere a los individuos (en total 18.948), $h = 1, \dots, 7$, se refiere a los tamaños de hábitat (detallado en la lista de variables en el apéndice) y $m = 1, \dots, 5$ se refiere a las distintas posiciones del individuo en el mercado laboral tales como si trabaja por cuenta propia o ajena, si se encuentra en el paro o si pertenece a la población que no busca trabajo como los estudiantes, los pensionistas o realiza labores del hogar (ver apéndice) y $u_i \square iid N(0, \sigma^2)$.

5.1. Análisis de los Resultados

A continuación presentamos los resultados más destacados. Buscaremos el mejor modelo predictivo fuera de la muestra. Para ello se seguirá el criterio del modelo con el mejor valor en el Error Cuadrático Medio de Previsión fuera de la Muestra (CMSPE).

Al modelo básico lo hemos planteado a partir de los siguientes dos puntos de vista:

- I. *Global*: Utilizando la muestra completa para las 18.948 observaciones. Los resultados de esta estimación serían aplicables a todos los individuos. Para referirnos a este modelo utilizaremos las siglas de la frase Modelo Básico Incondicional (MBI).

⁷ Relevant Transformation of the Inputs Network Approach.

- II. *Usuarios de Internet*: Utilizando solamente los usuarios de Internet. En total 5.240 observaciones. Los resultados serían sólo aplicables a los usuarios de Internet. Para este modelo utilizaremos las siglas (MBC).

Las estimaciones de ambos modelos arrojan los siguientes resultados:

Tabla 5: Comparación de los resultados de los Modelos Básicos

Endógena <i>IDD</i>	MBI	MBC
Constante	-0,66 (24,38)	-1,00 (28,66)
log(<i>EQ</i>)	0,17 (7,28)	0,22 (11,45)
log(<i>STUDYLEVEL</i>)	0,17 (30,53)	0,04 (4,80)
log(<i>AGE</i>)	-0,11 (15,23)	0,01 (0,75)*
<i>SEX_H</i>	0,01 (2,28)	0,02 (3,94)
<i>HOUSINT</i>	0,61 (34,25)	0,47 (31,20)
<i>SIZE1</i>	-0,01 (1,28)*	0,01 (0,73)*
<i>SIZE2</i>	----	----
<i>SIZE3</i>	-0,02 (2,38)	-0,01 (0,82)*
<i>SIZE4</i>	-0,04 (4,82)	-0,02 (2,08)
<i>SIZE5</i>	-0,05 (4,79)	-0,03 (2,75)
<i>SIZE6</i>	-0,09 (10,53)	-0,06 (4,98)
<i>SIZE7</i>	-0,01 (1,39)*	-0,01 (0,73)*
<i>CLUSTER1</i>	0,02 (3,07)	-0,004 (0,63)*
<i>CLUSTER2</i>	----	----
<i>CLUSTER3</i>	-0,01 (0,96)*	-0,02 (3,52)
<i>WORK0</i>	0,08 (8,29)	0,02 (1,44)*
<i>WORK1</i>	0,04 (4,73)	0,003 (0,36)*
<i>WORK2</i>	0,08 (6,54)	0,06 (5,28)
<i>WORK3</i>	0,02 (2,16)	0,01 (0,60)*
<i>WORK4</i>	-0,07 (6,79)	-0,02 (1,07)*
<i>WORK5</i>	----	----
<i>IMILLS</i>	----	0,0003 (0,16)*
Observaciones	18.673	5.204
\bar{R}^2	0,693	0,752
RCMSPE	0,262	0,181
AIC	-2,677	-3,427

Los resultados de las estimaciones en general nos muestran que:

- I. El nivel de digitalización individual de la sociedad española es sensible a la cantidad de equipamiento informático, al nivel de estudios alcanzados y a la edad. Vemos que las tres variables serán significativas y con los signos esperados en las estimaciones de MBI con muestra completa. En el caso del modelo condicional MBC vemos que el equipamiento y el nivel de estudios serán significativos mientras que la edad no lo será. Es decir que si restringimos la muestra a los usuarios de Internet, la edad no será importante en la explicación de las variaciones del nivel de digitalización. A su vez, esto confirma una vez más lo visto más arriba de que no sólo basta con acceder a la

tecnología, sino que además hay determinadas características adicionales tales como las que acabamos de mencionar (equipamiento, nivel de estudios y edad), que influyen en el nivel de digitalización.

- II. Si nos ceñimos sólo a los usuarios de Internet (MBC) observamos que se mejora ligeramente la bondad del ajuste (0,752 es mayor que 0,693). También hay un mejoramiento en los valores de RCMSPE y el AIC. Para el caso del valor de RCMSPE vemos que desciende sustancialmente (0,181 es bastante menor que 0,262).
- III. También vemos que el acceso a Internet en el hogar es significativo y con coeficientes con los signos esperados, igual que el sexo masculino. En el caso de variables que tienen en cuenta el tamaño del hábitat (*SIZE*) vemos son sólo significativas en el MBI. En el caso del MBC serán significativos sólo los tamaños intermedios de hábitat (*SIZE4*, *SIZE5* y *SIZE6*).

En el apartado anterior hemos visto que después de aplicar el análisis de conglomerados jerárquicos, las CCAA se separan en tres grupos que luego utilizamos para segmentar la muestra. Siguiendo la misma idea planteada en Marinucci *et al.* (2005) para recoger la heterogeneidad existen fuertes evidencias de utilizar estos grupos en el modelo final debido a que las diferencias entre los mismos son capturadas en las diferencias que hay entre los términos constantes. Esto viene a completar el análisis de los grupos iniciado en el apartado anterior.

Existen más evidencias de la heterogeneidad de los datos cuando vemos que algunas variables ficticias son estadísticamente significativas en algunos grupos y en otros no, como ocurre con la variable del hábitat (*SIZE_i*) o algunos tipos de situación en el mercado laboral (*WORK_i*). Aunque esto no estaría evidenciando diferencias de comportamiento entre individuos de diferentes CCAA, sino más bien por diferencias de rentas y otros factores demográficos.

5.2. Elección de un Modelo Predictivo para el Nivel de Digitalización

En esta sección presentamos los modelos predictivos (MPI para la muestra completa y MPC para la muestra condicional) seleccionados por RETINA. Compararemos los principales estadísticos para quedarnos con el mejor modelo y luego realizar predicciones. Llamaremos MPI y MPC a los modelos recomendados por RETINA para cada modelo básico (es decir MPI es el recomendado por RETINA para mejorar la predicción en MBI, y MPC lo es para MBC).

De los modelos recomendados por RETINA, el MPC será el mejor para predecir, debido a que tiene la mejor predicción fuera de la muestra (0,181 es menor que 0,257). La particularidad es que al comparar los RCMSPE entre MBC y MPC este no varía (será 0,181 para ambos modelos), aunque se gana sobre todo en parsimonia explicando las variaciones de la brecha tecnológica para los usuarios de Internet con bastante menos parámetros (antes con 20, ahora con 9). Una cuestión a tener en cuenta aquí es que los resultados de las predicciones realizadas con MPC sólo son aplicables a los usuarios de Internet, en cambio las predicciones realizadas con MPI son aplicables a toda la muestra. A continuación presentamos los principales resultados (para ver las tablas de regresión completa ir al apéndice):

Tabla 6: Comparación de los resultados de los Modelos Predictivos

	MPI	MPC
Nº de Parámetros	31	9
Transformaciones (W)	3	2
Constantes específicas	2	1
Pendientes específicas	15	5
\bar{R}^2	0,703	0,750
RCMSPE(1000)	0,257	0,181
AIC	-2,715	-3,422

Es evidente que tener en cuenta la heterogeneidad y la inclusión de no linealidades en las especificaciones mejorará tanto el ajuste como la habilidad predictiva de los modelos predictivos si los comparamos con los modelos básicos. En el caso del MPI vemos que RETINA no admite los regresores continuos originales, aunque sí transformaciones de estos. Así tenemos regresores tales como $\ln(STUDYLEVEL) * \ln(EDAD)$, como $\ln(EDAD)^2$ y como $\ln(EQ)^2$, todos estos significativos. Además dentro de las pendientes específicas que se han incluido están las referidas a los tres grupos de CCAA detectados en el cluster (*CLUSTER1*, *CLUSTER2* y *CLUSTER3*), que no podíamos tener en cuenta en el MPI por provocar la presencia de multicolinealidad exacta. Para el caso del MPC, vemos que tampoco se admiten los regresores continuos originales y sí transformaciones de estos tales como $\ln(STUDYLEVEL)^2$ y $\ln(AGE)^2$. Como ya lo comentamos arriba, una particularidad de este modelo es que tiene el mismo valor de RCMSPE que MPC, pero es mucho más parsimonioso. Es decir, está prediciendo igual, pero antes lo hacía con 20 regresores mientras que ahora lo hace sólo con 9.

Para validar ambos modelos de predicción utilizamos el contraste de especificación de Ramsey⁸ (RESET), y observamos que sólo para el caso del MPC no se rechaza la hipótesis de inexistencia de variables omitidas, mientras que en el MPI sí:

Tabla 15: Contraste RESET de los Modelos RETINA

MPI	MPC
$F(3,18612) = 253,58$	$F(3,5184) = 0,92$
$\text{Pr} > F = 0,000$	$\text{Pr} > F = 0,4318$

Por tener la mejor predicción fuera de la muestra, y por no rechazarse la hipótesis nula del contraste RESET, a partir de ahora nos centraremos en el MPC. La diferencia esencial entre el modelo básico MBC y el MPC se encuentra en que con el primero

⁸ El contraste RESET no solo es útil para contrastar la especificación del modelo. También se utiliza para detectar la presencia de no linealidades. La hipótesis nula es $H_0 : y_t = x_t^T \beta + u_t$ con $u_t \square iidN(0, \sigma^2)$, y la hipótesis alternativa $H_1 : E(y_t | x_t)$ es una función no lineal de los elementos de x_t . Para mayor detalle ver Davidson y McKinnon (2003).

hicimos la interpretación económica de los determinantes del nivel de digitalización de la sociedad española, mientras que con este último sólo realizaremos predicciones⁹.

5.3. *Análisis de Sensibilidad de la Brecha Tecnológica en los Usuarios de Internet*

Más arriba hemos visto que lo que más influirá en la existencia de una brecha tecnológica a nivel de CCAA y a nivel individual en España serán, además de factores tecnológicos tales como la adquisición y uso de equipamiento o el acceso a Internet desde el hogar, también lo serán factores demográficos tales como el nivel de estudios, la edad y el género sexual. Otra característica que hemos destacado además es que estos factores impactarán de manera diferente en cada CCAA.

En los últimos años en España se ha avanzado en el grado de sofisticación on line de la Administración tanto de las CCAA como del Estado central, aunque no todas las CCAA han avanzado por igual. Los presupuestos dedicados a las TIC por parte de las autonomías suponen ya el 73% del total, pero la inversión por habitante oscila entre los 87 euros en el País Vasco y los 10 euros que dedica la Región de Murcia (Fundación AUNA, 2005). También hay grandes diferencias en el número de personas que se han incorporado a Internet en el último año, en donde la comunidad con mayor crecimiento fue Cataluña (con un 22,65% del total), seguidas por Castilla y León, Andalucía y el País Vasco. Como contrapartida en el último puesto se encuentra Castilla-La Mancha, que incluso ha perdido usuarios.

Después de seleccionar el mejor modelo para la predicción, estudiamos los impactos de los distintos factores de la brecha tecnológica. El análisis lo haremos diferenciando por CCAA. Utilizaremos el MPC debido a que como hemos visto arriba, es el mejor de los dos que hemos estimado desde el punto de vista estadístico. Los resultados que obtendremos sólo serán aplicables a la población usuaria de Internet.

A continuación en el Gráfico 9 presentamos los valores pronosticados del índice de digitalización para los usuarios de Internet por CCAA:

⁹ Existe un gran paralelismo entre RETINA y el enfoque ARIMA para el estudio de series temporales, no solo porque los modelos estimados con el enfoque ARIMA se los utiliza únicamente para realizar predicciones y no para interpretarlos económicamente. Además tenemos que en ambos enfoques el orden inicial de los potenciales regresores se realiza a partir de correlaciones entre la variable dependiente y cada regresor candidato en el caso de RETINA y el valor retardado en el caso de ARIMA, la identificación está basada en funciones de correlación simples y parciales, y la elección del modelo para predecir se realiza a partir del RCMSPE.

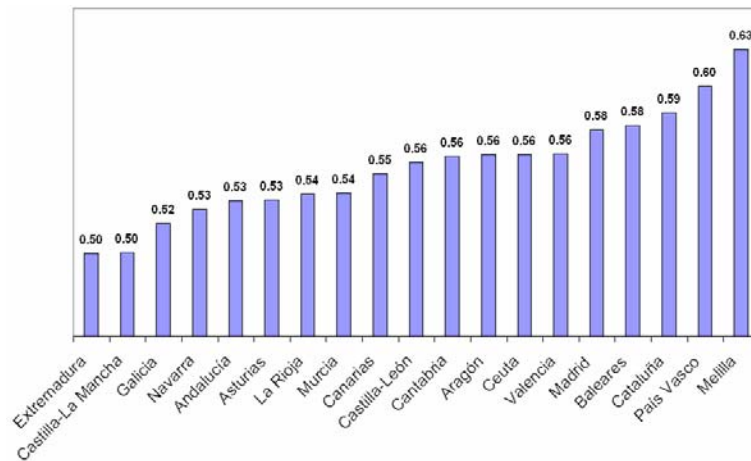


Gráfico 9: Valor del Pronóstico de Digitalización por CCAA

Considerando que cuando presentábamos IDD más arriba en el Gráfico 4 los valores se refieren a la muestra completa, vemos que las principales similitudes de los valores pronosticados con los valores observados se dan en las Comunidades de menor nivel de digitalización (Extremadura, Castilla-La Mancha, Galicia), mientras que esto no se mantiene con las Comunidades de mayor nivel como es el caso de Madrid y Melilla (ocuparán distintos lugares en el ranking de nivel de digitalización observado comparando con el nivel de digitalización predicho). Esto también se debe a que el modelo RETINA que estamos utilizando y los pronósticos que hemos obtenido son siempre sobre la muestra condicional a los usuarios de Internet.

A continuación se observa la Tabla 10 en donde se calcula la sensibilidad del nivel de digitalización respecto a los principales determinantes considerados para explicar el grado de digitalización. Estos son el equipamiento, el nivel de estudios, la edad y el acceso a Internet desde el hogar. Para ello hemos fijado cada una de las variables individualmente a su valor promedio¹⁰ y observando la variación provocada en el valor pronóstico.

¹⁰ Para el caso de la variable que considera el acceso a Internet desde el hogar, al tratarse de una variable ficticia, se la tomará siempre igual a 1. Es decir que simplemente su coeficiente estimado se suma al término constante.

Tabla 10: Sensibilidad del Nivel de Digitalización Respecto a sus Determinantes

Comunidad Autónoma	Valor Pronóstico	Equipamiento		Nivel de Estudios		Edad		Acceso a Internet	
		Valor	Δ %	Valor	Δ %	Valor	Δ %	Valor	Δ %
Andalucía	0,53	0,51	- 4,34	0,54	0,73	0,53	0,17	0,89	67,34
Aragón	0,56	0,54	- 4,48	0,56	0,20	0,56	0,10	0,87	54,39
Asturias	0,53	0,51	- 4,01	0,53	0,29	0,53	0,06	0,88	65,37
Baleares	0,58	0,55	- 4,88	0,59	0,83	0,58	0,19	0,87	51,15
Canarias	0,55	0,53	- 4,15	0,56	0,88	0,55	0,17	0,90	64,74
Cantabria	0,56	0,53	- 4,90	0,56	0,37	0,56	0,18	0,86	54,40
Castilla – León	0,56	0,53	- 4,62	0,57	0,76	0,56	0,10	0,87	56,83
Castilla La Mancha	0,50	0,48	- 3,33	0,50	0,50	0,50	0,25	0,88	75,71
Cataluña	0,59	0,56	- 5,04	0,60	0,73	0,59	0,10	0,88	50,18
Valencia	0,56	0,53	- 4,68	0,56	0,47	0,56	0,05	0,58	2,63
Extremadura	0,50	0,48	- 3,36	0,51	0,54	0,50	0,25	0,87	73,81
Galicia	0,52	0,50	- 3,91	0,52	0,40	0,52	0,06	0,89	71,98
Madrid	0,58	0,54	- 5,46	0,59	0,72	0,58	0,08	0,61	6,60
Murcia	0,54	0,51	- 4,19	0,55	0,76	0,54	0,10	0,58	7,81
Navarra	0,53	0,51	- 3,88	0,53	0,32	0,53	0,29	0,86	63,69
País Vasco	0,60	0,58	- 3,82	0,61	0,66	0,60	0,03	0,82	35,92
La Rioja	0,54	0,52	- 3,82	0,55	0,68	0,54	0,08	0,87	62,06
Ceuta	0,56	0,54	- 4,28	0,56	0,23	0,56	-0,01	0,91	62,09
Melilla	0,63	0,59	- 4,91	0,64	0,62	0,63	0,04	0,88	40,15

En cuanto a la sensibilidad del grado de digitalización respecto al equipamiento, cabe notar que si este fuera el mismo para todos los usuarios de Internet, el nivel de digitalización disminuiría en un porcentaje de 4,49% para España (es decir que aumentaría la brecha tecnológica). Para calcular este porcentaje de variación, igualamos el nivel de equipamiento a su media dentro de MPC, y vemos la variación de las predicciones. O sea que igualar a su media todas las observaciones del nivel de equipamiento provoca un descenso en el nivel de digitalización debido a que es una variable exclusivamente referida a la cantidad de ordenadores y otros equipos informáticos que posee el individuo. Encontramos aquí en esta muestra una relación inversa entre este equipamiento con la adquisición y uso del resto de las TIC. También vemos que el mayor impacto es provocado en las comunidades de mayor nivel de digitalización tales como Madrid (-5,46%) y Cataluña (-5,04), y el menor en las comunidades de baja digitalización tales como Extremadura (-3,36%) y Castilla-La Mancha (-3,33%). Esto podría deberse a que los usuarios de Internet de regiones con menor tendencia a lo digital sean mucho más propensos al uso del resto de las TIC que los usuarios de regiones de mayor tendencia.

En cuanto a las variaciones en el pronóstico provocadas por el nivel de estudios, se observa que la situación es bastante distinta. Los resultados de la Tabla 10 muestran que para todos los usuarios de Internet el nivel de digitalización se incrementaría sobre todo en Canarias y en Baleares (0,88% y 0,83% respectivamente). La brecha digital en regiones con menos nivel de digitalización es menos sensible a este atributo. Es el caso de Castilla-La Mancha, Extremadura o Galicia, con 0,50%, 0,54% y 0,40%. Las CCAA en donde la fijación del nivel de estudios impactará menos es en Aragón, Ceuta y Asturias, con 0,20%, 0,23 y 0,29% respectivamente. Es decir que la educación alcanzada por los usuarios de Internet impacta de distinta manera según de qué comunidad se trate.

Con respecto a la edad del individuo, se ve ya tanto en la literatura como en este trabajo que la edad es un atributo que influye en la afinidad por las nuevas tecnologías. También se observa de que existe una relación inversa entre estos dos fenómenos. En la Tabla 10 se comprueba que si bien la influencia de la edad es importante a nivel global, a nivel de usuario de Internet dicha influencia será mínima tal y como se veía en el modelo MBC del apartado anterior. Se observa que las variaciones serán mínimas, en donde las más sensibles con respecto al efecto de la edad serán la de Navarra (0,29%), Extremadura y Castilla-La Mancha (ambas con un 0,25%). La curiosidad aquí es que Navarra es una comunidad con un nivel de digitalización más cercano al de las comunidades con mayor nivel, sin embargo aquí se observa que su nivel de digitalización varía más con la edad que incluso comunidades bastante menos “digitales” como las ya mencionadas. En el otro extremo se encuentra Ceuta (con valor negativo de 0,01%), el País Vasco (0,03%) y Melilla (0,04%).

En cuanto a la influencia del acceso a Internet en el hogar se ven importantes diferencias en la sensibilidad de la brecha digital. Los resultados que se detallan en la Tabla 10 dan una idea de que el diverso nivel de digitalización que causa una brecha tecnológica en la mayoría de las comunidades será sensible principalmente a este atributo (el acceso a Internet desde el hogar). Sin embargo las diferencias aquí se hacen notar claramente en comunidades como Valencia o Madrid, en donde la sensibilidad de la brecha al acceso a Internet es de 2,63% y 6,60% respectivamente, en contraste con comunidades como Extremadura, Castilla - La Mancha y Galicia, con 75,71%, 73,81% y 71,98% respectivamente. Curiosamente también el grado de digitalización de Cataluña es bastante sensible al acceso a Internet ya que tiene un valor de 50,18%.

Entonces se puede concluir que:

- Las políticas de reducción de la brecha digital en CCAA con mayor nivel de digitalización tales como Madrid, País Vasco o Cataluña no deberían ser las mismas que en otras comunidades con un menor nivel como es el caso de Extremadura, Galicia o Castilla-La Mancha, ya que hemos visto que la sensibilidad del nivel de digitalización a sus determinantes en las distintas CCAA no vendrán a ser similares. Esto sugiere un estudio detallado a la hora de implementar una política de reducción de la brecha digital teniendo en cuenta otros factores tales como el nivel de estudios o la edad media de la población, el uso cotidiano de algún tipo específico de equipamiento tales como ordenadores, etc, además de sólo promover el acceso y uso de las TIC.
- Vemos que el grado de digitalización será más sensible al equipamiento en comunidades poco digitalizadas. Esto sugiere que una política de fomento de adquisición y uso de equipamientos podría ser adecuada. La liberalización de los mercados de telecomunicaciones ha estimulado nuevas demandas de equipamiento en los hogares y una política de reducción de la brecha digital podría ser apuntando a este aspecto.
- El acceso a Internet en el hogar es el atributo más importante de todos los considerados debido al impacto que provoca en el nivel de digitalización. Una política de reducción de la brecha digital entre CCAA podría considerar favorecer el acceso a Internet en todas sus formas (ADSL o cable módem). Sin embargo su impacto es bastante menor en regiones como Madrid,

Valencia o Murcia (estas dos últimas regiones con un incremento importante en su renta per capita desde el ingreso de España en la UE). Todo lo contrario ocurre con otras regiones, sobre todo en las más “atrasadas” como Castilla-La Mancha o Extremadura.

- El nivel de estudios que afectará a la digitalización de manera dispar. El nivel de digitalización será más sensible a este factor en Canarias y Baleares y menos en Aragón y Ceuta.

5.4. *Resumen de los Resultados*

Los resultados en general sugieren:

- I. La edad, el equipamiento y el nivel de estudios son buenos predictores de la brecha tecnológica, y su importancia para el modelo será grande tanto de manera lineal como interactuando entre sí o con otras variables.
- II. En todos los casos RETINA mejora la capacidad predictiva al incluir no linealidades y pendientes específicas.
- III. Los impactos de los factores que provocan la brecha tecnológica entre las CCAA de España no serán de la misma magnitud según se hable de un grupo con buena predisposición a la digitalización o de un grupo con menor predisposición. En el análisis de sensibilidad hemos visto cómo los atributos que influyen en la digitalización lo harán de manera distinta según se trate de una CCAA con menor o mayor nivel de digitalización. Esto debería ser tenido en cuenta en la formulación de políticas de reducción de la brecha digital tales como determinadas campañas de precios o descuentos.

6. Conclusiones

La importancia de la brecha digital entre países y regiones contrasta con el escaso bagaje teórico y empírico con el que cuenta su estudio. En este capítulo se centró el análisis en la medición de la brecha tecnológica dentro de las Comunidades Autónomas de España. El planteamiento del problema de cómo medir adecuadamente la dicha brecha se comienza intentando responder las siguientes preguntas:

- a. ¿Qué es la Brecha Digital?
- b. ¿Cómo se la puede medir?
- c. ¿Cuáles son sus causas más relevantes?
- d. ¿Cómo se la puede disminuir?

A partir de estas preguntas se propuso un indicador del nivel de digitalización y otro del gasto en tecnología. A partir de las dos formas que toma el indicador, se comparan los resultados para las CCAA considerando que un valor de brecha tecnológica “total” puede ser medida a partir de una medida tal como el coeficiente de variación, por ser esta una medida de dispersión relativa. Una vez calculado este coeficiente para toda la muestra, la brecha por CCAA se mide a partir de las distancias del valor normalizado del índice con la media total.

Una vez obtenido el cálculo del índice propuesto para cada CCAA y demostradas las diferencias existentes en los niveles de digitalización por CCAA, y siempre teniendo por finalidad explicar los determinantes de la existencia de estas diferencias, el siguiente paso será reagrupar las CCAA para recoger la heterogeneidad existente en la muestra. Para ello se estima un modelo econométrico de Poisson para el uso de Internet estimado con datos individuales para cada CCAA. Los resultados sugieren que los determinantes que provocan mayor variación porcentual son el gasto en equipamiento, la edad y el nivel de educación, aunque no con el mismo impacto para todas las Comunidades. A partir de aquí se formuló una pregunta adicional:

- e. ¿Las diferencias entre las elasticidades se deben a diferencias en el comportamiento entre CCAA, o las brechas observadas se pueden explicar por diferencias en el nivel de renta y otros factores demográficos?

Tener en cuenta las diferencias entre las elasticidades es importante debido a que equivale a decir que existen distintos impactos porcentuales en la variable a explicar por parte de los regresores. Entonces, para clasificar a las CCAA de acuerdo a su similitud se utiliza un análisis de conglomerados y en donde se distinguen tres grupos bien diferenciados: el de menor predisposición a la digitalización (Grupo 1), el de predisposición media (Grupo 2) y el de predisposición mayor (Grupo 3).

Una vez obtenidos dichos grupos, a continuación se los tiene en cuenta en la modelización del nivel de digitalización en donde se especifica un modelo básico que se mejora sustancialmente en habilidad predictiva por fuera de la muestra a partir de la inclusión de interacciones y no linealidades sugeridas por el algoritmo RETINA. A partir de él estimamos las diferencias que existen en los impactos de los atributos para los distintos grupos. Vemos que no serán de la misma magnitud según se trate de uno o de otro.

Referencias

1. Cassel C (1999) "Demand for and Use of Additional Lines by Residential Customers". In: Loomis, Taylor (eds.) The Future of the Telecommunications Industry: Forecasting and Demand Analysis. Kluwer Academic Publisher, Boston
2. Cerno, L. and Pérez-Amaral, T. (2005) "Demand of Internet Access and Use in Spain" mimeo presented in the ITS Conference on Regional Economic Development.
3. Corrocher, N. and Ordanini, A. (2002) "Measuring the digital divide: a framework for the analysis of cross-country differences" Journal of Information Technology, 17, 9-19.
4. Davidson R, MacKinnon J (2003) "Econometric Theory and Methods". Oxford University Press
5. Fundación AUNA (2005) "Informe España 2005" en http://www.fundacionauna.com/areas/25_publicaciones/

6. Lebart, L., Morineau, A. and Warwick, K. (1984) “Multivariate Descriptive Statistical Analysis”, Wiley.
7. Madden G, Savage S and Simpson M (1996) “Information Inequality and Broadband Network Access: An Analysis of Australian Household Survey Data”. Industry and Corporate Change, Oxford University Press, pp 1049–1056
8. Mansell, R. and When, U. (1998) “Knowledge Societies Information Technology for Suitable Development” Oxford University Press, Oxford and New York.
9. Marinucci, M. (2005) “RETINA Winpack for Real Data: A quick guide for Automatic Model Selection” Universidad Complutense de Madrid, Somosaguas, Madrid.
10. Marinucci, M. and Pérez-Amaral, T. (2005) “Econometric Modeling of Business Telecommunications Demand Using RETINA and Finite Mixtures” mimeo presented in the ITS Conference on Regional Economic Development.
11. Martin, W. (1995) “The Global Information Society” Bowker Saur, London.
12. OECD Report (2001) “Understanding the Digital Divide” in www.oecd.org/dataoecd/38/57/
13. Pelletiere, D. and Rodrigo, G. (2000) “An Empirical Investigation of the Digital Divide in the United States” mimeo presented at the Western Economics Association annual conference.
14. Peña, D. (2002) “Análisis Multivariante de Datos” Mc Graw Hill Eds., Madrid.
15. Pérez-Amaral, T., Gallo, G. and White, H. (2003) “A Flexible Tool for Model Building: The Relevant Transformation of the Inputs Network Approach (RETINA)” Oxford Bulletin of Economics and Statistics 65, supplement 1, 821-838.
16. Pérez-Amaral, T., Gallo, G. and White, H. (2005) “A Comparison of Complementary Automatic Modeling Methods: RETINA and PcGets” Econometric Theory, 21, 262-277.
17. Randstad Report (2004) “Quality at Work in the EU-15: The Knowledge Society” in www.eurofound.eu.int/ewco/reports/
18. Rappoport P, Taylor L and Kridel D (2002) “The Demand of Broadband: Access, Content, and the Value of Time”. In: Crandall RW, Alleman JH (eds.) Broadband: Should We Regulate High-Speed Internet Access? AEI-Brookings Joint Centre for Regulatory Studies, Washington, D.C.
19. Red.es – Ministerio de Industria, Turismo y Comercio (2004) “Las TIC en los Hogares Españoles” in <http://observatorio.red.es/estudios/consumo/index.html>

20. Ricci, A. (2000) "Measuring information society dynamics of European data on usage of information and communication technologies in Europe since 1995" *Telematics and Informatics*, 17, 141-67.
21. Sneath, P. and Sokal, R. (1973) "Numerical Taxonomy" W.H. Freeman, San Francisco.
22. Taylor LD (1994) "Telecommunications Demand in Theory and Practice". Kluwer Academic Publishers.
23. The World Bank (2005) "Financing Information and Communication Infrastructure Needs in the Developing World: Public and Private Roles" Draft for Discussion in <http://lnweb18.worldbank.org/ict/resources.nsf/>

Apéndice

Especificación de un Índice de Brecha Tecnológica para las CA de España

Siguiendo la idea original planteada en Corrocher y Ordanini (2002) en donde se desarrolla una medida para el nivel de digitalización dentro de un conjunto de países a través de indicadores agrupados por un factor de digitalización, la dispersión del valor final calculada para cada zona geográfica constituye una medida de brecha tecnológica dentro del sistema. Nosotros pretendemos elaborar una medida que además nos sirva para determinar las diferencias por zona geográfica (Comunidad Autónoma en nuestro caso), y a su vez pueda ser utilizada al nivel de desagregación tal que alcance al mismo individuo.

Como punto de partida consideramos todas las observaciones de la muestra (18.948 individuos entrevistados en 2003 en España) y la siguiente especificación:

$$I_{TIC_i} = \left\{ \sum_{j=1}^n \left[q_{ij} D_{ij} \prod_{k=1}^k (1 + q_{ijk} D_{ijk}) \right] \right\}$$

$i=1, \dots, 18.948$

en donde:

I_{TIC_i} : Índice TIC que cuantifica el grado de digitalización para cada individuo de la muestra.

D_{ij} : Variable ficticia intergrupales (=1 si el agente i posee algún elemento entre los distintos conjuntos de tecnología j).

q_{ij} : Peso de importancia del conjunto de tecnología j .

D_{ijk} : Variable ficticia intragrupal (=1 si el agente i dentro del conjunto de tecnología j posee algún elemento k).

$q_{.ij}$: Peso de importancia del elemento dentro del conjunto j .

Los grupos de elementos que se han considerado son los siguientes planteados en la Tabla A:

Tabla A: Equipamientos de Tecnología Considerados

Grupos j	Subgrupos k
<i>Televisión</i>	Antena convencional Antena parabólica Cable (incluye satélite) Televisión digital
<i>Ordenador</i>	Ordenador de sobremesa Portátil Otros tipo(incluye agenda electrónica)
<i>Telefonía</i>	Teléfono fijo Teléfono móvil
<i>Internet</i>	Conexión por banda ancha
<i>Otros Equipamientos</i>	Cadena musical Radio Vídeo (incluye videoconsola) DVD Fax Automóvil con sistema de navegación

Ponderaciones

En una segunda parte y una vez que ya hemos identificado y decidido los factores de digitalización a tener en cuenta, es ver qué nivel o grado de importancia se le dará a cada uno dentro del índice sintético. A partir de esta cuestión se nos presenta una primera posibilidad: la utilización de un peso *objetivo*, es decir una ponderación dentro de cada subgrupo dentro de cada grupo, y de cada grupo dentro del total. En este caso estaríamos haciendo referencia al *gasto* que se supone realiza el individuo que está orientado a las nuevas tecnologías.

Dentro de esta primera posibilidad, las ponderaciones se realizan entonces de acuerdo a la proporción de cada uno de los elementos de los subgrupos k dentro del total de cada grupo j . Es decir:

$$q_{ijk} = \frac{\text{Cantidad de elementos } k}{\text{Cantidad total de elementos en } j} \quad q_{ij} = \frac{\text{Cantidad de elementos } j}{\text{Cantidad total de elementos}}$$

$i = 1, 2, \dots, 18.948$ individuos.

$j = 1, 2, \dots, 5$ grupos de equipamientos.

Utilizando estas ponderaciones planteadas aquí, entonces lo que obtenemos es una variable *proxy* del gasto individual en tecnología digital, debido a que correspondería a la parte del ingreso individual a adquirir y mantener los elementos que el individuo dice poseer.

En la segunda posibilidad se varían los pesos de importancia anteriormente considerados. Por ejemplo, si consideramos pesos objetivos, el tener televisión o línea

de teléfono fijo en el hogar será más importante que tener conexión de Banda Ancha, debido a que es más probable que el individuo posea más de lo primero que de lo segundo¹¹. A partir de esta idea hemos rediseñado las ponderaciones de acuerdo a la siguiente tabla:

Tabla B: Pesos de cada Equipamiento en el Índice

Grupos <i>j</i>	Subgrupos <i>k</i>
<i>Televisión</i> (15%)	Antena convencional (30%) Antena parabólica (15%) Cable (20%) Televisión digital (35%)
<i>Ordenador</i> (15%)	PC (35%) Portátil (45%) Otros tipos (20%)
<i>Telefonía</i> (15%)	Teléfono fijo (50%) Teléfono móvil (50%)
<i>Internet</i> (35%)	Conexión por banda ancha (100%)
<i>Otros Equipamientos</i> (10%)	Cadena musical (20%) Radio (15%) Vídeo (15%) DVD (15%) Fax (15%) Automóvil con sistema de navegación (20%)

Al variar las ponderaciones de esta manera, es posible obtener un tipo de *Índice Sintético del Grado de Digitalización* para cada individuo, agregándolo para cada Comunidad Autónoma. Esto nos permitirá evaluar la magnitud de la brecha tecnológica a nivel de Comunidad Autónoma, o por zonas geográficas según se determinen.

¹¹ Es posible que casi todos tengan televisión y teléfono fijo o móvil en casa, aunque la conexión a Internet en el hogar, según hemos observado antes, no llega al 30% de los hogares españoles.

Definición de variables

<i>Variable</i>	<i>Definición</i>
<i>USE</i>	Cantidad de sitios donde el individuo usa Internet (hasta 4)
(*) <i>IDD</i>	Índice de digitalización
(*) <i>ITIC</i>	Gasto Individual en tecnología
<i>HOUSINT</i>	Ficticia =1 si el individuo accede a Internet en el hogar
<i>EQ</i>	Cantidad de equipamiento informático del individuo en el hogar
<i>MALE</i>	Ficticia =1 si el individuo es de sexo masculino
<i>AGE</i>	Edad del individuo
<i>POP</i>	Población provincial
<i>STUDYLEVEL</i>	Grado alcanzado en nivel de estudios (en años)
<i>BESTUDYING</i>	Ficticia =1 si el individuo aún se encuentra estudiando
<i>SIZE1</i>	Ficticia =1 si es de una capital de provincia con más de 500.000 habitantes
<i>SIZE2</i>	Ficticia =1 si es de una capital de provincia
<i>SIZE3</i>	Ficticia =1 si es de una municipio con más de 100.000 habitantes
<i>SIZE4</i>	Ficticia =1 si es de un municipio entre 50.000 y 100.000 habitantes
<i>SIZE5</i>	Ficticia =1 si es de un municipio entre 20.000 y 50.000 habitantes
<i>SIZE6</i>	Ficticia =1 si es de un municipio entre 10.000 y 20.000 habitantes
<i>SIZE7</i>	Ficticia =1 si es de un municipio con menos de 10.000 habitantes
<i>COM1</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Andalucía
<i>COM2</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Aragón
<i>COM3</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Asturias
<i>COM4</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Baleares
<i>COM5</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Canarias
<i>COM6</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Cantabria
<i>COM7</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Castilla y León
<i>COM8</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Castilla – La Mancha
<i>COM9</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Cataluña
<i>COM10</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Valencia
<i>COM11</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Extremadura
<i>COM12</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Galicia
<i>COM13</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Madrid
<i>COM14</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Murcia
<i>COM15</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de Navarra
<i>COM16</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de País Vasco
<i>COM17</i>	Ficticia =1 si pertenece a la Comunidad Autónoma de La Rioja
<i>COM18</i>	Ficticia =1 si pertenece a las Comunidades de Ceuta o Melilla
<i>WORK0</i>	Ficticia =1 si se encuentra trabajando por cuenta propia
<i>WORK1</i>	Ficticia =1 si se encuentra trabajando por cuenta ajena
<i>WORK2</i>	Ficticia =1 si se encuentra parado
<i>WORK3</i>	Ficticia =1 si es estudiante
<i>WORK4</i>	Ficticia =1 si se dedica a las labores del hogar
<i>WORK5</i>	Ficticia =1 si es pensionista
<i>IMILLS</i>	Inversa del Ratio de Mills

- CLUSTER1* Ficticia =1 si pertenece a las Comunidades Autónomas de Madrid, Baleares, Canarias, País Vasco, Cataluña, Ceuta o Melilla.
- CLUSTER2* Ficticia =1 si pertenece a las Comunidades Autónomas de Aragón, Asturias, Valencia, Castilla-La Mancha o Murcia.
- CLUSTER3* Ficticia =1 si pertenece a las Comunidades Autónomas de Cantabria, Andalucía, Navarra, Castilla y León, Extremadura, Galicia o La Rioja.
-

(*) Nota: La construcción de estas dos variables fueron explicadas en detalle en el apartado anterior.

Resultados del modelo de Poisson (Sección 4)

Regresores	Andalucía			Aragón		
	$\hat{\beta}$	$dy/dx^{(*)}$	ε	$\hat{\beta}$	$dy/dx^{(*)}$	ε
Constante	-3.776 (8.58)	---	---	-4.263 (5.73)	---	---
<i>ITIC</i>	1.189 (5.07)	.144 (5.44)	1.905 (5.07)	1.655 (4.28)	.218 (4.65)	2.770 (4.28)
<i>HOUSINT</i>	.515 (3.54)	.074 (2.90)	.105 (3.54)	.620 (2.91)	.096 (2.37)	.162 (2.91)
<i>EQ</i>	.152 (1.79)	.028 (1.76)	.035 (1.79)	-.109 (.81)	-.014 (.81)	-.032 (.81)
<i>MALE</i>	.207 (2.83)	.026 (2.74)	.087 (2.83)	.255 (2.58)	.034 (2.50)	.115 (2.58)
<i>AGE</i>	-.038 (11.44)	-.005 (12.13)	-1.803 (11.44)	-.039 (8.71)	-.005 (8.75)	-2.017 (8.71)
<i>HABITAT</i>	.017 (1.06)	.002 (1.06)	.077 (1.06)	-.212 (.67)	-.003 (.67)	-.129 (0.67)
<i>STUDYLEVEL</i>	2.306 (13.41)	.279 (11.57)	1.211 (13.41)	2.332 (9.61)	.306 (8.12)	1.342 (9.61)
<i>BESTUDYING</i>	.389 (4.63)	.055 (3.74)	.048 (4.63)	.211 (1.75)	.030 (1.54)	.022 (1.75)
Observaciones	2147			996		
LR	1,280.74 ($\chi^2 = 0.00$)			662.03 ($\chi^2 = 0.00$)		
Pseudo R²	.3724			.3809		

(*) Para el caso de las variables ficticias, el efecto marginal representa el cambio discreto de 0 a 1.

Regresores	Asturias			Balears		
	$\hat{\beta}$	$dy/dx^{(*)}$	ε	$\hat{\beta}$	$dy/dx^{(*)}$	ε
Constante	-3.828 (5.18)	---	---	-3.253 (4.26)	---	---
<i>ITIC</i>	1.459 (3.83)	.201 (4.12)	2.377 (3.83)	1.002 (2.44)	.168 (2.57)	1.625 (2.44)
<i>HOUSINT</i>	.491 (2.00)	.079 (1.67)	.098 (2.00)	.776 (2.79)	.157 (2.26)	.230 (2.79)
<i>EQ</i>	.202 (1.23)	.028 (1.22)	.044 (1.23)	.073 (.44)	.012 (.43)	.024 (.44)
<i>MALE</i>	.062 (.55)	.008 (.55)	.024 (.55)	.274 (2.14)	.047 (2.08)	.124 (2.14)
<i>AGE</i>	-.034 (7.62)	-.005 (7.79)	-1.799 (7.62)	-.034 (6.13)	-.006 (6.40)	-1.633 (6.13)
<i>HABITAT</i>	-.006 (.18)	-.0008 (.18)	-.019 (.18)	-.013 (.42)	-.002 (.42)	-.043 (.42)
<i>STUDYLEVEL</i>	1.919 (7.36)	.264 (6.69)	1.096 (7.36)	2.003 (6.75)	.336 (6.16)	1.113 (6.75)
<i>BESTUDYING</i>	.230 (1.74)	.035 (1.54)	.025 (1.74)	.234 (1.61)	.043 (1.43)	.029 (1.61)
Observaciones	893			579		
LR	536.80 ($\chi^2 = 0.00$)			378.19 ($\chi^2 = 0.00$)		
Pseudo R²	.3628			.3598		

(*) Para el caso de las variables ficticias, el efecto marginal representa el cambio discreto de 0 a 1.

Regresores	Canarias			Cantabria		
	$\hat{\beta}$	dy/dx (*)	ε	$\hat{\beta}$	dy/dx (*)	ε
Constante	-3.077 (4.93)	---	---	-3.350 (3.61)	---	---
<i>ITIC</i>	1.135 (3.38)	.273 (3.57)	1.930 (3.38)	.793 (1.70)	.094 (1.76)	1.30 (1.70)
<i>HOUSINT</i>	.365 (1.68)	.095 (1.52)	.107 (1.68)	1.139 (3.80)	.196 (2.56)	.248 (3.80)
<i>EQ</i>	.151 (1.02)	.037 (1.02)	.049 (1.02)	-.069 (.39)	-.008 (.40)	-.017 (.39)
<i>MALE</i>	.267 (2.63)	.065 (2.56)	.121 (2.63)	-.342 (2.30)	.042 (2.17)	.141 (2.30)
<i>AGE</i>	-.037 (8.25)	-.009 (9.21)	-1.646 (8.25)	-.034 (5.15)	-.004 (5.48)	-1.75 (5.15)
<i>HABITAT</i>	-.024 (.83)	-.006 (.83)	-.075 (.83)	-.046 (1.35)	-.006 (1.33)	-.164 (1.35)
<i>STUDYLEVEL</i>	1.928 (7.92)	.465 (7.78)	1.010 (7.92)	2.537 (7.00)	.300 (5.99)	1.429 (7.00)
<i>BESTUDYING</i>	.429 (3.80)	.120 (3.13)	.072 (3.80)	.267 (1.45)	.035 (1.26)	.025 (1.45)
Observaciones	787			544		
LR	513.98 ($\chi^2 = 0.00$)			366.43 ($\chi^2 = 0.00$)		
Pseudo R²	.3257			.4131		

(*) Para el caso de las variables ficticias, el efecto marginal representa el cambio discreto de 0 a 1.

Regresores	Castilla - León			Castilla – La Mancha		
	$\hat{\beta}$	dy/dx (*)	ε	$\hat{\beta}$	dy/dx (*)	ε
Constante	-2.994 (5.16)	---	---	-3.951 (4.90)	---	---
<i>ITIC</i>	.845 (2.82)	.102 (3.00)	1.350 (2.82)	1.550 (3.86)	.125 (4.24)	2.417 (3.86)
<i>HOUSINT</i>	.773 (4.11)	.120 (3.05)	.164 (4.11)	.695 (2.08)	.073 (1.56)	.100 (2.08)
<i>EQ</i>	.150 (1.33)	.018 (1.31)	.036 (1.33)	-.058 (.20)	-.005 (.20)	-.009 (.20)
<i>MALE</i>	.336 (3.67)	.041 (3.46)	.152 (3.67)	.038 (.33)	.003 (.33)	.017 (.33)
<i>AGE</i>	-.036 (8.70)	-.004 (8.97)	-1.877 (8.70)	-.041 (7.53)	-.003 (7.14)	-2.128 (7.53)
<i>HABITAT</i>	-.022 (1.01)	-.003 (1.01)	-.791 (1.01)	-.008 (.26)	-.0006 (.26)	-.034 (.26)
<i>STUDYLEVEL</i>	2.002 (9.02)	.243 (7.79)	1.098 (9.02)	2.093 (7.71)	.170 (6.17)	1.017 (7.71)
<i>BESTUDYING</i>	.320 (2.80)	.044 (2.34)	.040 (2.80)	.497 (3.52)	.049 (2.58)	.053 (3.52)
Observaciones	1304			1012		
LR	916.90 ($\chi^2 = 0.00$)			590.69 ($\chi^2 = 0.00$)		
Pseudo R²	.4145			.4110		

(*) Para el caso de las variables ficticias, el efecto marginal representa el cambio discreto de 0 a 1.

Regresores	Cataluña			Valencia		
	$\hat{\beta}$	dy/dx (*)	ε	$\hat{\beta}$	dy/dx (*)	ε
Constante	-2.894 (6.23)	---	---	-4.347 (7.34)	---	---
<i>ITIC</i>	1.081 (4.53)	.243 (4.74)	1.832 (4.53)	1.494 (4.76)	.176 (5.26)	2.420 (4.76)
<i>HOUSINT</i>	.733 (5.27)	.1941 (4.34)	.223 (5.27)	.703 (3.82)	.103 (2.92)	.157 (3.82)
<i>EQ</i>	.073 (0.92)	.016 (.91)	.025 (.92)	.050 (.44)	.006 (.43)	.013 (.44)
<i>MALE</i>	.103 (1.57)	.024 (1.55)	.042 (1.57)	.233 (2.68)	.028 (2.58)	.100 (2.68)
<i>AGE</i>	-.319 (11.29)	-.007 (11.81)	-1.58 (11.29)	-.036 (8.97)	-.004 (9.13)	-1.763 (8.97)
<i>HABITAT</i>	.005 (.32)	.001 (.32)	.029 (.32)	.015 (.70)	.001 (.70)	.071 (.70)
<i>STUDYLEVEL</i>	1.42 (9.37)	.320 (9.34)	.805 (9.37)	2.177 (10.31)	.257 (8.38)	1.172 (10.31)
<i>BESTUDYING</i>	.277 (3.43)	.070 (2.99)	.028 (3.43)	.331 (3.27)	.045 (2.71)	.038 (3.27)
Observaciones	1833			1377		
LR	1223.31 ($\chi^2 = 0.00$)			962.93 ($\chi^2 = 0.00$)		
Pseudo R²	.3339			.4141		

(*) Para el caso de las variables ficticias, el efecto marginal representa el cambio discreto de 0 a 1.

Regresores	Extremadura			Galicia		
	$\hat{\beta}$	dy/dx (*)	ε	$\hat{\beta}$	dy/dx (*)	ε
Constante	-3.17 (3.91)	---	---	-3.662 (5.74)	---	---
<i>ITIC</i>	.996 (2.36)	.084 (2.52)	1.553 (2.36)	1.361 (4.09)	.122 (4.41)	2.126
<i>HOUSINT</i>	.743 (1.92)	.083 (1.44)	.109 (1.92)	.909 (3.92)	.114 (2.67)	.150
<i>EQ</i>	.054 (.18)	.005 (.18)	.009 (.18)	-.203 (1.22)	-.018 (1.21)	-.388
<i>MALE</i>	.170 (1.21)	.015 (1.18)	.071 (1.21)	.099 (.97)	.009 (.95)	.042
<i>AGE</i>	-.043 (6.79)	-.004 (7.07)	-2.174 (6.79)	-.044 (9.45)	-.004 (9.11)	-2.260
<i>HABITAT</i>	.022 (.66)	.002 (.66)	.103 (.66)	-.020 (.72)	-.002 (.72)	-.084
<i>STUDYLEVEL</i>	1.976 (6.34)	.166 (5.36)	.982 (6.34)	2.428 (10.02)	.217 (7.58)	1.285
<i>BESTUDYING</i>	.492 (3.02)	.051 (2.22)	.046 (3.02)	.240 (1.93)	.024 (1.66)	.026
Observaciones	831			1171		
LR	385.54 ($\chi^2 = 0.00$)			780.07 ($\chi^2 = 0.00$)		
Pseudo R²	.3557			.4252		

(*) Para el caso de las variables ficticias, el efecto marginal representa el cambio discreto de 0 a 1.

Regresores	Madrid			Murcia		
	$\hat{\beta}$	dy/dx (*)	ε	$\hat{\beta}$	dy/dx (*)	ε
Constante	-3.748 (7.89)	---	---	-4.515 (5.75)	---	---
<i>ITIC</i>	1.221 (4.89)	.337 (5.20)	2.114 (4.89)	1.899 (4.60)	.163 (5.08)	2.97 (4.60)
<i>HOUSINT</i>	.568 (4.44)	.176 (3.82)	.181 (4.44)	.563 (2.09)	.058 (1.66)	.113 (2.09)
<i>EQ</i>	.086 (1.36)	.024 (1.35)	.034 (1.36)	.0145 (.08)	.001 (.08)	.003 (.08)
<i>MALE</i>	.152 (2.31)	.042 (2.28)	.067 (2.31)	.342 (2.66)	.030 (2.46)	.149 (2.66)
<i>AGE</i>	-.030 (10.99)	-.008 (11.83)	-1.418 (10.99)	-.042 (7.38)	-.004 (7.03)	-2.117 (7.38)
<i>HABITAT</i>	.033 (1.99)	.009 (1.98)	.172 (1.99)	-.010 (.24)	-.001 (.24)	-.028 (.24)
<i>STUDYLEVEL</i>	1.956 (11.86)	.540 (11.83)	1.238 (11.86)	1.733 (5.79)	.148 (4.92)	.904 (5.79)
<i>BESTUDYING</i>	.277 (3.77)	.084 (3.35)	.048 (3.77)	.420 (2.86)	.042 (2.19)	.055 (2.86)
Observaciones	1567			773		
LR	1135.60 ($\chi^2 = 0.00$)			491.66 ($\chi^2 = 0.00$)		
Pseudo R²	.3360			.4092		

(*) Para el caso de las variables ficticias, el efecto marginal representa el cambio discreto de 0 a 1.

Regresores	Navarra			País Vasco		
	$\hat{\beta}$	dy/dx (*)	ε	$\hat{\beta}$	dy/dx (*)	ε
Constante	-3.015 (5.36)	---	---	-4.029 (5.38)	---	---
<i>ITIC</i>	.892 (3.06)	.149 (3.21)	1.488 (3.06)	1.120 (2.78)	.141 (2.93)	1.889 (2.78)
<i>HOUSINT</i>	.423 (2.09)	.079 (1.84)	.111 (2.09)	1.110 (3.88)	.183 (2.87)	.335 (3.88)
<i>EQ</i>	.220 (1.53)	.037 (1.52)	.062 (1.53)	.190 (.85)	.024 (.84)	.060 (.85)
<i>MALE</i>	.258 (2.83)	.044 (2.75)	.121 (2.83)	.257 (2.73)	.033 (2.69)	.121 (2.73)
<i>AGE</i>	-.035 (9.30)	-.006 (9.95)	-1.731 (9.30)	-.029 (6.84)	-.004 (6.97)	-1.504 (6.84)
<i>HABITAT</i>	-.040 (2.00)	-.007 (1.98)	-.164 (2.00)	-.003 (.10)	-.0003 (.10)	-.011 (.10)
<i>STUDYLEVEL</i>	2.174 (10.32)	.363 (9.42)	1.271 (10.32)	1.949 (8.48)	.245 (7.32)	1.051 (8.48)
<i>BESTUDYING</i>	.482 (5.36)	.097 (3.66)	.067 (5.36)	.201 (1.46)	.245 (7.32)	.015 (1.46)
Observaciones	1202			1181		
LR	747.24 ($\chi^2 = 0.00$)			808.83 ($\chi^2 = 0.00$)		
Pseudo R²	.3539			.4043		

(*) Para el caso de las variables ficticias, el efecto marginal representa el cambio discreto de 0 a 1.

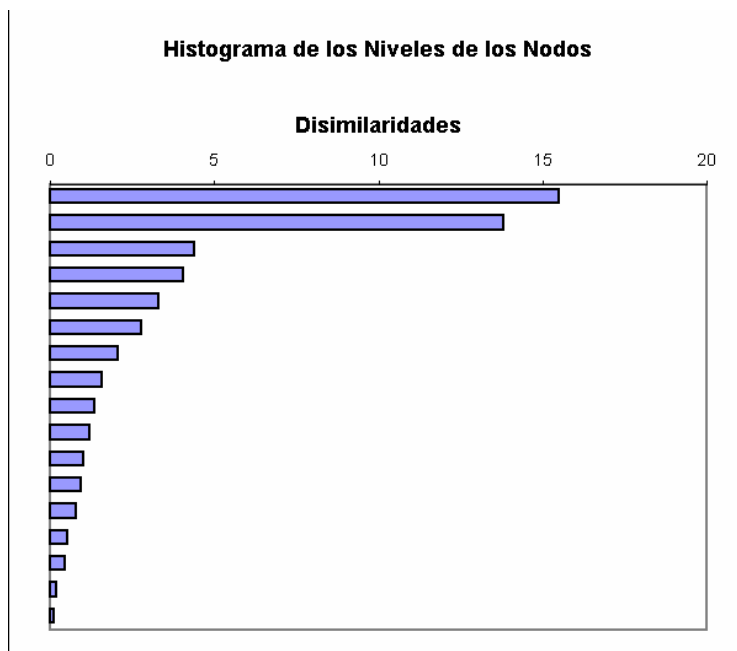
Regresores	La Rioja			Ceuta y Melilla		
	$\hat{\beta}$	$dy/dx^{(*)}$	ε	$\hat{\beta}$	$dy/dx^{(*)}$	ε
Constante	-2.870 (3.06)	---	---	-3.624 (3.07)	---	---
<i>ITIC</i>	.879 (1.82)	.108 (1.92)	1.439 (1.82)	1.296 (2.03)	.283 (2.23)	2.158 (2.03)
<i>HOUSINT</i>	.947 (2.84)	.159 (1.97)	.204 (2.84)	.591 (1.18)	.147 (1.00)	.182 (1.18)
<i>EQ</i>	-.097 (.39)	-.012 (.39)	-.023 (.39)	.320 (1.03)	.070 (1.02)	.110 (1.03)
<i>MALE</i>	.263 (1.95)	.033 (1.89)	.119 (1.95)	.041 (.21)	.009 (.21)	.019 (.21)
<i>AGE</i>	-.042 (6.94)	-.005 (6.73)	-2.124 (6.94)	-.033 (3.41)	-.007 (3.62)	-1.412 (3.41)
<i>HABITAT</i> ⁽⁺⁾	-.063 (1.85)	-.008 (1.85)	-.209 (1.85)	---	---	---
<i>STUDYLEVEL</i>	2.286 (7.02)	.282 (5.97)	1.298 (7.02)	1.601 (3.66)	.351 (3.39)	.940 (3.66)
<i>BESTUDYING</i>	.538 (3.70)	.082 (2.77)	.072 (3.70)	.580 (2.40)	.154 (1.83)	.106 (2.40)
Observaciones	549			203		
LR	398.89 ($\chi^2 = 0.00$)			149.54 ($\chi^2 = 0.00$)		
Pseudo R²	.4109			.3681		

(*) Para el caso de las variables ficticias, el efecto marginal representa el cambio discreto de 0 a 1.

(+) Esta variable no tiene variabilidad en Ceuta y Melilla. De considerarla en la regresión provocaría multicolinealidad exacta.

Conglomerados jerárquicos de Comunidades Autónomas para elasticidades de uso de Internet con respecto al gasto en equipamiento, la edad y el nivel de estudios (Sección 4)

Nodo	Ramificación a la izquierda	Ramificación a la derecha	Frecuencia	Peso	Nivel
35	33	34	18	18.000	15.498
34	32	30	11	11.000	13.793
33	27	31	7	7.000	4.377
32	29	24	5	5.000	4.060
31	25	26	4	4.000	3.292
30	28	23	6	6.000	2.792
29	2	19	3	3.000	2.058
28	13	21	4	4.000	1.568
27	6	22	3	3.000	1.340
26	12	17	2	2.000	1.215
25	7	11	2	2.000	1.006
24	8	14	2	2.000	0.934
23	9	18	2	2.000	0.793
22	1	15	2	2.000	0.516
21	4	20	3	3.000	0.450
20	5	16	2	2.000	0.186
19	3	10	2	2.000	0.123



Modelo Probit para el uso de Internet para la generación de la inversa del Ratio de Mills (valores absolutos de los estadísticos *t* entre paréntesis) – (Sección 5)

Variable Dependiente	Use (Sí=1)
Observaciones	18,948
Log-verosimilitud	-5368.82
Chi-cuadrado (dgf)	11457.19 (6)
P- valor	.000
Pseudo R ²	.516
Constante	-2.458 (18.71)
<i>ITIC</i>	.968 (13.57)
<i>HOUSINT</i>	.909 (11.94)
<i>EQ</i>	.130 (2.13)
<i>MALE</i>	.263 (9.54)
<i>AGE</i>	-.043 (41.91)
<i>STUDYLEVEL</i>	2.475 (38.33)

Resultados de las Estimaciones del MPI (Sección 5)

Variable	Coeficiente
Constante	-0,94 (57,34)
Interacciones	
$\ln(STUDYLEVEL) * \ln(EDAD)$	0,05 (18,15)
$\ln(EDAD)^2$	-0,004 (2,91)
$\ln(EQ)^2$	0,11 (7,15)
Constantes específicas	
<i>HOUSINT</i>	0,60 (54,14)
<i>WORK2</i>	-0,23 (0,25)
<i>WORK4</i>	0,26 (4,66)
Pendientes específicas	
$CLUSTER1 * \ln(EQ) * \ln(STUDYLEVEL)$	-0,15 (4,13)
$CLUSTER1 * [1/\ln(EDAD)^2]$	-0,15 (1,83)
$CLUSTER2 * \ln(AGE)^2$	-0,04 (6,39)
$CLUSTER2 * \ln(EQ) * \ln(STUDYLEVEL)$	-0,16 (4,30)
$CLUSTER3 * \ln(EQ) * \ln(STUDYLEVEL)$	-0,17 (4,64)
$CLUSTER3 * \ln(STUDYLEVEL)^2$	-0,03 (2,21)
$CLUSTER3 * [\ln(STUDYLEVEL)/\ln(EDAD)]$	0,17 (1,98)
$CLUSTER3 * \ln(STUDYLEVEL) * \ln(AGE)$	-0,006 (1,01)
$HOUSINT * \ln(STUDYLEVEL)^2$	0,06 (2,44)
$WORK1 * \ln(EQ)$	-0,02 (1,70)
$WORK2 * \ln(EQ)$	-0,20 (8,71)
$WORK2 * \ln(AGE)$	0,25 (0,45)
$WORK2 * \ln(STUDYLEVEL) * \ln(EDAD)$	-0,01 (0,77)
$WORK2 * \ln(EDAD)^2$	-0,05 (0,52)
$WORK3 * \ln(EQ)$	0,06 (3,01)
$WORK3 * \ln(STUDYLEVEL) * \ln(EDAD)$	0,009 (0,94)
$WORK3 * \ln(STUDYLEVEL)^2$	0,02 (1,01)
$WORK3 * \ln(EDAD)^2$	-0,02 (1,02)
$WORK4 * \ln(EQ)$	1,37 (2,88)
$WORK4 * \ln(STUDYLEVEL) * \ln(EDAD)$	-0,02 (5,03)
$WORK4 * \ln(EDAD)^2$	-0,03 (7,81)
$WORK4 * [\ln(EQ)/\ln(EDAD)]$	-4,84 (2,43)
$SIZE6 * \ln(STUDYLEVEL) * \ln(AGE)$	-0,004 (1,46)

$SIZE6 * \ln(EDAD)^2$	-0,005 (7,38)
Observaciones	18673
R ² ajustado	0,704
Error estándar de la estimación	0,257
RCMSPE	0257
AIC	-2,718

Resultados de las Estimaciones del MPC (Sección 5)

Variable	Coefficiente
Constante	-0,92 (68,91)
Interacciones	
$\ln(STUDYLEVEL)^2$	-0,05 (4,51)
$\ln(AGE)^2$	-0,004 (4,09)
Constantes específicas	
<i>HOUSINT</i>	0,54 (55,92)
Pendientes específicas	
$SEX_H * \ln(EQ)^2$	0,002 (4,23)
$VIVINT * \ln(STUDYLEVEL)$	-0,07 (4,58)
$VIVINT * (\ln(EQ))^2$	0,12 (11,76)
$SIZE6 * [1/\ln(AGE)^2]$	-0,52 (7,32)
$WORK1 * [\ln(STUDYLEVEL)/\ln(AGE)]$	0,21 (5,52)
Observaciones	5.204
R ² ajustado	0,750
Error estándar de la estimación	0,122
RCMSPE	0,181
AIC	-3,422