



Universidad de Valladolid

Facultad de Ciencias Económicas y Empresariales

Grado en Marketing e Investigación de Mercados

Modelo Econométrico del Turismo en España

Presentado por:

Silvia Veganzones González

Tutelado por:

María Isabel Gómez Valle

Valladolid, 25 de junio de 2015

ÍNDICE

1. INTRODUCCION	3
2. ANTECEDENTES.....	4
3. METODOLOGÍA	6
3.1. Datos y software	6
3.2. Variables	6
3.2.1. Variable dependiente.....	6
3.2.2. Variables independientes	8
3.3. Matriz de correlaciones	9
3.4. Análisis de modelos econométricos.....	11
3.4.1. Modelo 1	11
3.4.2. Modelo 2.....	12
3.4.3. Modelo 3.....	13
3.5. Modelo elegido	14
3.6. Contraste sobre la estabilidad de los parámetros.....	15
3.6.1. Año 2007	16
3.6.2. Año 2008	16
3.7. Análisis de las hipótesis clásicas	22
3.7.1. Homocedasticidad	22
3.7.2. Incorrelación	23
3.7.3. Normalidad	23
3.8. Análisis del modelo finalmente elegido.....	24
3.9. Predicción.....	27
4. CONCLUSIONES	29
5. BIBLIOGRAFÍA	31

1. INTRODUCCION

El turismo, según la Organización Mundial de Turismo (OMT), “es un fenómeno social, cultural y económico relacionado con el movimiento de las personas a lugares que se encuentran fuera de su lugar de residencia habitual por motivos personales o profesional”.

El turismo genera en España el 10,2% del PIB y el 11,5% del empleo, por lo que su evolución determina, en gran medida, la evolución de la economía nacional.

En este trabajo se pretende explicar el comportamiento de la demanda turística en España, medida como número de viajeros, a través de factores relevantes, como por ejemplo el tipo de cambio de la moneda, el Índice de Precios Hoteleros, el Índice de Precios al Consumo, el paro, el personal empleado o los establecimientos abiertos estimados.

No pretendemos en esta investigación cuantificar la demanda turística, ya que como señala Sánchez (1998), “tanto el Instituto Nacional de Estadística como otros organismos nacionales y autonómicos elaboran estadísticas que permiten conocer el flujo del turismo registrado durante un período de tiempo determinado en una zona concreta” sino conocer un poco mejor su comportamiento desde 1999 hasta 2014.

El trabajo se estructura en varias partes. En la primera se hace un repaso de algunos de los modelos econométricos que se han realizado sobre el sector.

En la segunda se describe el procedimiento llevado a cabo para encontrar el modelo. Se estudiará la variable dependiente o endógena, número de viajeros, y las variables explicativas que me harán conformar el modelo más adecuado para la misma, que son: Producto Interior Bruto, Índice de Precios Hoteleros, Índice de Precios al Consumo (restaurantes y hoteles), establecimientos abiertos, empleo y personal contratado.

Se analizará la matriz de correlaciones y basándome en ella plantearé varios modelos econométricos según como estén de correlacionadas. Además se seleccionará aquel modelo que mejores resultados a priori presente y analizando si cumple las hipótesis clásicas, y por último se realizará una predicción.

En la última parte se exponen las conclusiones del trabajo y la bibliografía utilizada.

2. ANTECEDENTES

Como ya hemos comentado, el sector turístico es un pilar fundamental de muchas economías. El estudio de la demanda turística ha dado lugar a numerosos trabajos en el ámbito nacional e internacional, especialmente en los últimos treinta años, como señala Sánchez, M. (1998). Según esta autora, unos se centran en la estimación y previsión de la demanda turística, otros en analizar el grado de satisfacción del turista y otros estudian el comportamiento de los individuos ante las vacaciones.

Si nos centramos en los primeros, que son los que nos interesan de cara a nuestro trabajo, destaca, por ser el primero que utilizó técnicas econométricas para explicar la demanda turística española, el realizado por Pulido San Román (1966) "*Introducción a un análisis econométrico del turismo*", donde planteó el turismo como un fenómeno económico y demostró la utilidad de la Econometría para explicar y predecir su comportamiento.

También ha trabajado mucho en este tema Esteban Talaya con su tesis doctoral "*Análisis de la demanda: aplicación a la actividad turística de las técnicas de predicción*" (1987), cuyo objetivo fundamental consistía en estudiar los factores determinantes socioeconómicos, demográficos, pictográficos y técnicos de la demanda turística, para identificar las características de su comportamiento y sus dimensiones. Estimó varios modelos y posteriormente generó predicciones de la demanda turística española basándose en variables explicativas como visitantes extranjeros entrados por fronteras, pernoctaciones

en establecimientos hoteleros, visitantes extranjeros entrados por aeropuertos e ingresos por turismo.

Otras autoras que han modelizado la demanda turística son Guisan, M. C. y Aguayo, E., así por ejemplo, en el artículo "*Evolución del turismo hotelero en las regiones españolas*" (2001-2008) trabajan con dos modelos. El primero de ellos utiliza como variables explicativas precio y pernoctaciones hoteleras en logaritmos. El segundo modelo trata de analizar el impacto del turismo sobre el desarrollo económico y empleo en España. Se trata de un modelo econométrico donde la variable endógena es el Valor Añadido no Industrial por habitante, y las variables explicativas el Valor Añadido Industrial, las pernoctaciones hoteleras, el gasto en educación, y alguna variable ficticia con regiones europeas.

En esta misma línea, Guisán, C y Neira, I. en el trabajo "*Análisis econométrico del turismo hotelero y extrahotelero en las regiones y provincias españolas*" modelizaron la demanda turística a través de diversos modelos econométricos sobre el turismo hotelero y extrahotelero en España de 1991 a 1995.

Para ello, en un primer modelo, introdujeron como variable endógena la proporción de las pernoctaciones hoteleras anuales de la región en el total nacional, siendo las variables explicativas, la misma retardada, la proporción del número de plazas hoteleras y la proporción media ponderada del precio por plaza de cada región respecto a la media nacional correspondiente.

En cambio, el otro modelo econométrico, incluía variables ficticias donde se observó diferencias sustanciales en algunas regiones de España en comparación con el resto por variables como el clima o proximidad a grandes núcleos de población.

El último modelo analiza el turismo extrahotelero, donde se aprecian diferencias con respecto el número de viviendas secundarias, renta de la población, clima y cercanía a grandes núcleos de población sobre Alicante, Tarragona y Valencia.

Otro trabajo en esta línea es el de Espasa, A., Gómez-Churruga, R. y Jareño, J. “*Un análisis econométrico de los ingresos por turismo en la economía española*”, donde analizarán los ingresos por turismo en la economía española de dos formas: a través de los ingresos nominales y con el número de turistas que entran en España.

3. METODOLOGÍA

3.1. Datos y software

Para la realización de este trabajo los datos que he utilizado provienen del Instituto Nacional de Estadística (INE). Este organismo proporciona colecciones de datos, temporales o de corte transversal de diversas variables macroeconómicas, financieras, demográficas y culturales, etc.

La muestra de datos empleada es anual y abarca desde el año 1999 hasta el 2014, siendo los datos utilizados relativos a España.

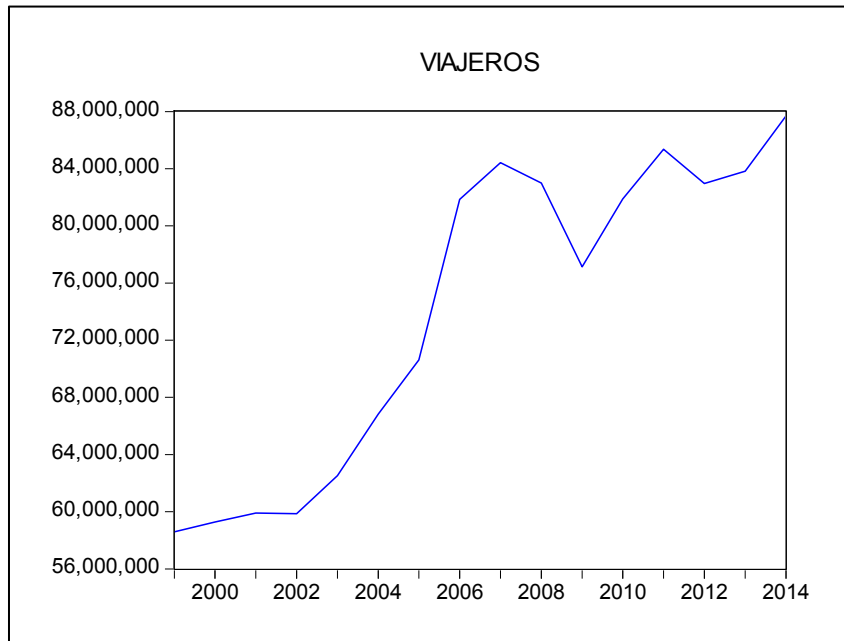
Los programas informáticos utilizados para la elaboración del modelo han sido Microsoft Excel, que se ha utilizado para organizar los datos obtenidos de las bases de datos oficiales e EViews 6.0, para realizar el tratamiento econométrico de esos datos.

3.2. Variables

3.2.1. Variable dependiente

La variable dependiente o endógena que voy a utilizar es el número de viajeros, que son todas aquellas personas que realizan una o más de una pernoctación en el mismo alojamiento y se ha extraído de la Encuesta de Ocupación Hotelera del INE.

Analizamos brevemente la evolución de esta variable en el periodo de interés, de 1999 a 2014. Como se puede observar en el gráfico que aparece a continuación, los viajeros han evolucionado de forma creciente.



Fuente: Elaboración propia con datos del INE

Cabe destacar un gran incremento del número de viajeros en el año 2006, que según el Instituto Nacional de Estadística (INE) supuso un aumento de turistas del 4,2% respecto 2005.

Con respecto a los turistas que provienen del extranjero esto se pudo deber a una crecida del uso de internet para contratar viajes, donde casi el 40% acudía a compañías de bajo coste para contratar servicios turísticos de manera directa, sin recurrir a los canales tradicionales de distribución, según la encuesta de movimientos turísticos en fronteras (Frontur).

En los siguientes años se puede observar un crecimiento de viajeros en 2007, un leve descenso en 2008 y una gran caída en 2009, que se puede atribuir a la crisis financiera de 2008.

Aparece un cambio de la tendencia en el año 2010, con respecto la tendencia decreciente de los años anteriores, que a partir de 2011, se puede justificar con el comienzo de la Primavera Árabe, donde los turistas rechazaron como destino para sus vacaciones el norte de África por estas revueltas y desviaron su destino a España.

Y por último, en el año 2014 se observa un gran incremento de viajeros tras un leve descenso en el año 2012 y 2013, España recibió un 7,1% más que en 2013, siendo el mejor año de la historia.

3.2.2. Variables independientes

A continuación voy a describir las variables explicativas o independientes que previamente he considerado para elaborar el modelo econométrico más adecuado, y estudiar cuales de ellas son las que más afectan al número de viajeros.

Las variables son las siguientes:

- Número de establecimientos abiertos estimados en España: cualquier establecimiento hotelero abierto de temporada inscrito en la Consejería de Turismo de cada Comunidad Autónoma. Son aquellos que prestan servicios de alojamiento mediante precio con o sin otros servicios complementarios, que la he denominado *Establecimientos*.
- Personal contratado: esta variable está medida en personas, y son el conjunto de personas remuneradas o no remuneradas, que contribuyen mediante su aportación a la producción de bienes y servicios en el establecimiento que he denominado *Personal*.
- Índice de Precios Hoteleros en restaurantes y hoteles (base 2005): mide el nivel de precios y servicios de consumo pagados solo por los hogares residentes en España sobre restaurantes y hoteles. Esta variable la he nombrado como *Ipcesyhotel*.
- Índice de Precios Hoteleros (Base 2008): permite medir la variación de los precios percibidos por los empresarios hoteleros considerando a todos sus clientes. Esta variable en el modelo la he denominado como *lph*.
- Producto Interior Bruto (PIB): es un indicador que mide la producción de bienes y servicios de un país, durante un periodo de tiempo

determinado. Además, la actividad turística suele influir bastante en este, y esta medida en miles de euros, denominada *Pib* en el modelo.

- Empleo: lo componen aquellas personas que tienen un trabajo y se les remunera por ello. Esta variable está medida en miles de personas y la he denominado como *Empleo*.

Todas las variables explicativas del modelo las he extraído del Instituto Nacional de Estadística (INE), más concretamente de la Encuesta de Ocupación Hotelera, excepto los dos últimos indicadores, Producto Interior Bruto (PIB) y empleo, que son de la Contabilidad Regional del INE.

3.3. Matriz de correlaciones

La matriz de correlaciones muestra las relaciones que existen entre las variables que se van a incluir en el modelo. Estas variables se relacionan de forma independiente, pero me fijaré en aquellas donde su correlación sea más próxima a 1 con respecto a viajeros. Esta matriz puede tomar valores dentro de un intervalo de (-1,1).

Su interpretación depende del valor del signo que tome y de las características de las variables analizadas. Un coeficiente de correlación muy cercano a uno (valor absoluto), indica que la relación entre las variables es muy fuerte, mientras que si es muy cercano a cero, indica que la relación es muy débil.

Para explicar la evolución del turismo en España, como ya he señalado, voy a utilizar la variable número de viajeros como variable endógena o dependiente, la relacionaré con las anteriores variables independientes para ver cuáles de ellas son las que están más correlacionadas con viajeros y tratando a su vez que entre ellas no estén demasiado correlacionadas para buscar el modelo econométrico óptimo.

La tabla que aparece a continuación recoge la matriz de correlación, en la que me basare para elaborar los modelos econométricos:

	VIAJEROS	IPH	IPCRESYHOTEL	PERSONAL	EMPLEO	ESTABLECIMIENTOS	PIB
VIAJEROS	1	0,6604	0,9479	0,5456	0,6539	0,2980	0,9510
IPH	0,6604	1	0,6021	0,3385	0,9614	-0,3532	0,7866
IPCRESYHOTEL	0,9479	0,6021	1	0,4420	0,5821	0,2916	0,9597
PERSONAL	0,5456	0,3385	0,4420	1	0,2199	0,0117	0,4206
EMPLEO	0,6539	0,9614	0,5821	0,2199	1	-0,2273	0,7862
ESTABLECIMIENTOS	0,2980	-0,3532	0,2916	0,0117	-0,2273	1	0,1526
PIB	0,9510	0,7866	0,9597	0,4206	0,7862	0,1526	1

Fuente: Elaboración propia

La variable más correlacionada con la endógena es el Producto Interior Bruto, con un coeficiente de 95,10%. La segunda variable explicativa más correlacionada es el Índice de Precios al Consumo (restaurantes y hoteles) con un 94,79%, empleo (65,39%), Índice de Precios Hoteleros (66,04%), el personal contratado (54,56%) y por último establecimientos estimados abiertos con 29,89%.

Cabe destacar que no todas las variables correlacionadas con la variable endógena serán óptimas para el modelo, ya que algunas variables están muy correlacionadas entre sí y se seleccionaría una de ellas.

Así bien, en la matriz estás incluidas variables que recogen los precios de sector, que son el Índice de Precios Hoteleros y el Índice de Precios al Consumo (restaurantes y hoteles) (60,21%), variables que recogen la oferta del sector, establecimientos abiertos estimados y personal contratado (1,17%), y las que recogen la situación económica general, Producto Interior Bruto y empleo (78,72%), que probablemente aportarían la misma información al modelo y se elegirían una de las dos para incluir en el modelo.

3.4. Análisis de modelos econométricos

A partir de las series originales obtenidas del Instituto Nacional de Estadística (INE), he estimado varios modelos partiendo de las variables que estén más correlacionadas con la variable endógena, en este caso viajeros, para seleccionar el mejor modelo y analizar si se cumplen las hipótesis clásicas sobre el mismo.

Si el modelo cumple las hipótesis clásicas, el método de estimación mejor es el de Mínimos Cuadrados Ordinarios (MCO), que obtiene los estimadores de los parámetros con la siguiente expresión:

$$\hat{\beta} = (X'X)^{-1}X'Y$$

Estos estimadores se han obtenido con el programa EViews y los resultados son los siguientes.

A continuación voy a realizar modelos econométrico de la variable endógena, viajeros, sobre las variables explicativas, empezando por la que está más correlacionada hasta la que menos, para ir comparando los modelos y observando si el R^2 ajustado mejora y si Akaike y Schwarz son menores.

3.4.1. Modelo 1

El primer modelo propuesto es el siguiente:

$$\text{Viajeros} = \beta_0 + \beta_1 \text{pib} + \varepsilon$$

Dependent Variable: VIAJEROS
 Method: Least Squares
 Date: 06/05/15 Time: 10:30
 Sample: 1999 2014
 Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PIB	0.059536	0.005171	11.51460	0.0000
C	18713187	4892991.	3.824488	0.0019
R-squared	0.904493	Mean dependent var		74111314
Adjusted R-squared	0.897671	S.D. dependent var		11145482
S.E. of regression	3565319.	Akaike info criterion		33.12787
Sum squared resid	1.78E+14	Schwarz criterion		33.22445
Log likelihood	-263.0230	Hannan-Quinn criter.		33.13282
F-statistic	132.5859	Durbin-Watson stat		0.815000
Prob(F-statistic)	0.000000			

Por lo que la ecuación sería:

$$\widehat{\text{Viajeros}} = 18.713.187 + 0,059536 \text{ pib} + \varepsilon$$

Como se puede ver en el modelo, el R^2 ajustado es de un 89,76% y Akaike y Schwarz 33,127 y 33,224, respectivamente.

3.4.2. Modelo 2

El segundo modelo sería con la segunda variable más correlacionada con número de viajeros más la primera.

$$\text{Viajeros} = \beta_0 + \beta_1 \text{ pib} + \beta_2 \text{ ipcresyhotel} + \varepsilon$$

En este caso, es el Índice de Precios al Consumo (restaurantes y hoteles), pero está muy correlacionada con Producto Interior Bruto (PIB), como dije en el apartado de la matriz de correlaciones, aun así realizaré el modelo y observaré el R^2 ajustado para ver si mejora porque podría aportar información al modelo que el PIB por sí sola no me esté aportando.

Dependent Variable: VIAJEROS
 Method: Least Squares
 Date: 06/05/15 Time: 10:42
 Sample: 1999 2014
 Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PIB	0.032802	0.017466	1.878004	0.0830
IPCRESYHOTEL	309118.7	193818.3	1.594889	0.1348
C	11403957	6524319.	1.747915	0.1040
R-squared	0.920122	Mean dependent var		74111314
Adjusted R-squared	0.907833	S.D. dependent var		11145482
S.E. of regression	3383651.	Akaike info criterion		33.07417
Sum squared resid	1.49E+14	Schwarz criterion		33.21903
Log likelihood	-261.5934	Hannan-Quinn criter.		33.08159
F-statistic	74.87441	Durbin-Watson stat		1.116601
Prob(F-statistic)	0.000000			

El modelo es el siguiente:

$$\hat{\text{Viajeros}} = 11.403.957 + 0,032802 \text{ pib} + 309.118,7 \text{ ipcresyhotel} + \varepsilon$$

En este modelo el R^2 ajustado es de 90,78%, que mejora con respecto al modelo anterior, así que ambas variables son relevantes, y Akaike y Schwarz son 33,074 y 33,219 respectivamente.

3.4.3. Modelo 3

La siguiente variable más correlacionada es el empleo, al ser una variable que me recoge la situación económica del sector, al igual que el PIB, no la incluyo en el modelo. Del mismo modo ocurre con el Índice de Precios Hoteleros (IPH), que recoge los precios del sector como el Índice de Precios al Consumo, así que no se incluye tampoco, ya que la variable de precios ya incluida está más correlacionada con la variable endógena.

Posteriormente, las siguientes variables que recoge el modelo son las relativas a la oferta del sector turístico, establecimientos abiertos estimados y personal.

La variable que está más correlacionada con la variable endógena es personal por lo que se añadiría al modelo anterior:

$$\text{Viajeros} = \beta_0 + \beta_1 \text{ pib} + \beta_2 \text{ ipcresyhotel} + \beta_3 \text{ personal} + \varepsilon$$

Dependent Variable: VIAJEROS
 Method: Least Squares
 Date: 06/06/15 Time: 10:37
 Sample: 1999 2014
 Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PIB	0.033256	0.015672	2.122003	0.0553
IPCRESYHOTEL	255230.1	175891.3	1.451067	0.1724
PERSONAL	381.1898	187.1131	2.037216	0.0643
C	-6.44E+08	3.22E+08	-2.001433	0.0685
R-squared	0.940649	Mean dependent var		74111314
Adjusted R-squared	0.925811	S.D. dependent var		11145482
S.E. of regression	3035761.	Akaike info criterion		32.90214
Sum squared resid	1.11E+14	Schwarz criterion		33.09529
Log likelihood	-259.2171	Hannan-Quinn criter.		32.91203
F-statistic	63.39574	Durbin-Watson stat		1.425750
Prob(F-statistic)	0.000000			

Por lo que la ecuación sería la siguiente:

$$\widehat{\text{Viajeros}} = -6.44\text{E}+08 + 0,033256 \text{ pib} + 255.230,1 \text{ ipcresyhotel} + 381,1898 \text{ personal} + \varepsilon$$

Como se puede observar en la tabla de resultados por el programa Eviews, el R^2 ajustado es de un 92,58% que ha mejorado notablemente con respecto los anteriores modelos. Y con respecto Akaike y Schwarz, su valor es de 32,902 y 33,095, que también han mejorado (sus valores son más pequeños).

3.5. Modelo elegido

El modelo elegido es el último descrito, ya que el R^2 ajustado ha mejorado, de un 89,76% que resultaba al solo incluir PIB, pasó a un 90,78%, PIB e Índice de Precios al Consumo (restaurantes y hoteles), y al añadir el personal el R^2 ajustado aumentó hasta el 92,58%.

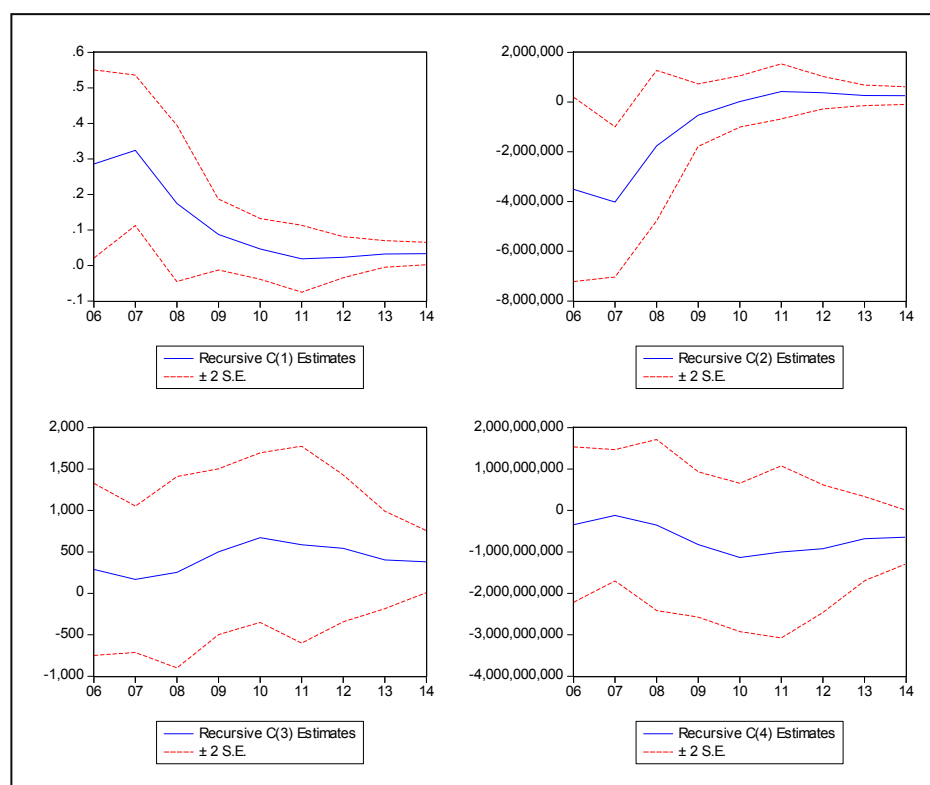
Con respecto Akaike y Schwarz, complementan información sobre los modelos, cuanto menor sean mejor, ya que se cometerá menor error por omisión de

variables que no son relevantes. En los tres modelos son similares, aunque en el modelo óptimo Akaike es menor y Schwarz ligeramente mayor.

3.6. Contraste sobre la estabilidad de los parámetros

Vamos a comprobar en este apartado si el comportamiento de los coeficientes para todo el periodo muestral se mantiene constante o no, es decir, si existe cambio estructural.

Para detectar si existe o no, como estoy utilizando datos temporales, utilizo el gráfico de coeficientes recursivos, de los regresores, calculados en $t-1$, para que me aproxime donde puede haber un salto en la muestra:



Fuente: Elaboración propia

Como se puede ver en los gráficos que aporta el programa EViews, se puede intuir que en el año 2007 o 2008 podría haber un salto en la muestra, y ser un cambio estructural, pero no se puede afirmar.

Para ello, es necesario aplicar el contraste de estabilidad de Chow:

$$H_0: \begin{pmatrix} \beta_0^1 \\ \beta_1^1 \\ \beta_2^1 \end{pmatrix} = \begin{pmatrix} \beta_0^1 \\ \beta_1^1 \\ \beta_2^1 \end{pmatrix}$$

$$H_1: \begin{pmatrix} \beta_0^1 \\ \beta_1^1 \\ \beta_2^1 \end{pmatrix} \neq \begin{pmatrix} \beta_0^1 \\ \beta_1^1 \\ \beta_2^1 \end{pmatrix}$$

Que sigue una $F_{N_1+N_2-2k-2}^{k-1}$

3.6.1. Año 2007

Para el año 2007 se aplica el contraste de Chow en el programa Eviews y se obtienen los siguientes resultados:

Chow Breakpoint Test: 2007
 Null Hypothesis: No breaks at specified breakpoints
 Varying regressors: All equation variables
 Equation Sample: 1999 2014

F-statistic	1.368621	Prob. F(4,8)	0.3262
Log likelihood ratio	8.341699	Prob. Chi-Square(4)	0.0798
Wald Statistic	5.474482	Prob. Chi-Square(4)	0.2420

En este año, el p valor es de 0,3262, que es mayor que 0,05, por lo que no se rechaza la hipótesis nula, o lo que es lo mismo, en el año 2007 no existe cambio estructural en el modelo

3.6.2. Año 2008

Como en el apartado anterior, voy a comprobar si existe cambio estructural para el año 2008 y los resultados son los siguientes:

Chow Breakpoint Test: 2008
 Null Hypothesis: No breaks at specified breakpoints
 Varying regressors: All equation variables
 Equation Sample: 1999 2014

F-statistic	6.849769	Prob. F(4,8)	0.0107
Log likelihood ratio	23.79591	Prob. Chi-Square(4)	0.0001
Wald Statistic	27.39907	Prob. Chi-Square(4)	0.0000

En este año, el p valor es de 0,0107, menor que 0,05, se rechaza la hipótesis nula y puede existir cambio estructural en el año 2008.

Este cambio estructural tendría coherencia, ya que coincide con el comienzo de la crisis financiera que afectó a España 2008 y donde el sector turístico se vio muy afectado.

Para poder corregirlo, es necesario crear una variable ficticia, como variable dicotómica, dando valor 0 a los años desde 1999 a 2007 incluido, que es donde ocurre el cambio estructural, y valor 1 desde 2008 a 2014.

$$\text{Ficticia} = \begin{cases} 0 & \text{si } 1999 \text{ a } 2007 \\ 1 & \text{si } 2008 \text{ a } 2014 \end{cases}$$

Posteriormente se incluye en el modelo la variable ficticia de forma aditiva y multiplicativa para ver si el cambio estructural se ha corregido.

Dependent Variable: VIAJEROS

Method: Least Squares

Date: 06/22/15 Time: 10:02

Sample: 1999 2014

Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PIB	0.324112	0.085882	3.773938	0.0054
IPCRESYHOTEL	-4021875.	1225681.	-3.281339	0.0112
PERSONAL	168.6199	357.9392	0.471085	0.6502
FICTICIA	-5.96E+08	6.93E+08	-0.859138	0.4153
FICTICIA*PIB	-0.014049	0.127078	-0.110557	0.9147
FICTICIA*IPCRESYHOTEL	7063657.	1477576.	4.780570	0.0014
FICTICIA*PERSONAL	-109.6203	386.3124	-0.283761	0.7838
C	-1.19E+08	6.43E+08	-0.184985	0.8578
R-squared	0.986587	Mean dependent var		74111314
Adjusted R-squared	0.974851	S.D. dependent var		11145482
S.E. of regression	1767512.	Akaike info criterion		31.91490
Sum squared resid	2.50E+13	Schwarz criterion		32.30119
Log likelihood	-247.3192	Hannan-Quinn criter.		31.93468
F-statistic	84.06239	Durbin-Watson stat		2.092298
Prob(F-statistic)	0.000001			

El modelo es el siguiente:

$$\widehat{\text{Viajeros}} = -1.19\text{E}+08 + 0,324112 \text{ pib} - 402.1875 \text{ ipcresyhotel} + 168,6199 \text{ personal} - 5.96\text{E}+08 \text{ ficticia} - 0.014049 \text{ ficticia*pib} + 7063657 \text{ ficticia*ipcresyhotel} - 109.6203 \text{ ficticia*personal} + \varepsilon$$

Para ver si el modelo ha mejorado incluyendo la variable ficticia, se observa el R^2 ajustado, que en este caso es del 97,48% y sin las variables ficticias el modelo era peor, con un 92,58%. Aún así, al incluirlas de forma multiplicativa y aditiva se observa que la única variable significativa es la que multiplica la ficticia con el Índice de Precios al Consumo (restaurantes y hoteles), por lo que procedo a suprimir las variables que sean menos significativas en el modelo para ver si este mejora.

La primera que suprimo es la variable ficticia que afecta al Producto Interior Bruto:

Dependent Variable: VIAJEROS

Method: Least Squares

Date: 06/22/15 Time: 10:03

Sample: 1999 2014

Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PIB	0.317695	0.059726	5.319206	0.0005
IPCRESYHOTEL	-3930528.	854195.9	-4.601436	0.0013
PERSONAL	185.2518	306.4496	0.604510	0.5604
FICTICIA	-5.74E+08	6.28E+08	-0.914586	0.3843
FICTICIA*IPCRESYHOTEL	7034690.	1372043.	5.127165	0.0006
FICTICIA*PERSONAL	-129.2714	323.6226	-0.399451	0.6989
C	-1.51E+08	5.41E+08	-0.278635	0.7868
R-squared	0.986567	Mean dependent var		74111314
Adjusted R-squared	0.977611	S.D. dependent var		11145482
S.E. of regression	1667699.	Akaike info criterion		31.79142
Sum squared resid	2.50E+13	Schwarz criterion		32.12943
Log likelihood	-247.3314	Hannan-Quinn criter.		31.80873
F-statistic	110.1613	Durbin-Watson stat		2.096119
Prob(F-statistic)	0.000000			

El modelo es:

$$\widehat{\text{Viajeros}} = -1.51E+08 + 0,317695 \text{ pib} - 3.930.528 \text{ ipcresyhotel} + 185,2518 \text{ personal} - 5.74E+08 \text{ ficticia} + 7.034.690 \text{ ficticia*ipcresyhotel} - 129,2714 \text{ ficticia*personal} + \varepsilon$$

El R² ajustado ha mejorado con respecto el anterior modelo, con un 97,76%, pero aún así procedo a suprimir la siguiente variable que es menos significativa con el modelo, la que afecta a personal.

Dependent Variable: VIAJEROS

Method: Least Squares

Date: 06/22/15 Time: 10:04

Sample: 1999 2014

Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PIB	0.327366	0.052255	6.264835	0.0001
IPCRESYHOTEL	-4047793.	767717.5	-5.272503	0.0004
PERSONAL	73.85042	121.5643	0.607501	0.5571
FICTICIA	-8.18E+08	1.41E+08	-5.780093	0.0002
FICTICIA*IPCRESYHOTEL	7206333.	1247063.	5.778642	0.0002
C	44968377	2.20E+08	0.204427	0.8421

R-squared	0.986328	Mean dependent var	74111314
Adjusted R-squared	0.979493	S.D. dependent var	11145482
S.E. of regression	1596081.	Akaike info criterion	31.68400
Sum squared resid	2.55E+13	Schwarz criterion	31.97372
Log likelihood	-247.4720	Hannan-Quinn criter.	31.69883
F-statistic	144.2882	Durbin-Watson stat	1.964734
Prob(F-statistic)	0.000000		

El modelo es:

$$\hat{\text{Viajeros}} = 44.968.377 + 0,327366 \text{ pib} - 4.047.793 \text{ ipcresyhotel} + 73,85042 \text{ personal} - 8.18\text{E}+08 \text{ ficticia} + 7.206.333 \text{ ficticia*ipcresyhotel} + \varepsilon$$

En este caso, si observamos el R^2 ajustado es de 97,94%, que ha mejorado con respecto el modelo anterior, y los p valores de todas las variables del modelo son significativos, a excepción de la variable personal, que suprimiré para ver si el modelo mejora.

Dependent Variable: VIAJEROS

Method: Least Squares

Date: 06/22/15 Time: 10:33

Sample: 1999 2014

Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PIB	0.338783	0.047339	7.156489	0.0000
IPCRESYHOTEL	-4199858.	704642.7	-5.960266	0.0001
FICTICIA	-8.55E+08	1.24E+08	-6.910326	0.0000
FICTICIA*IPCRESYHOTEL	7530515.	1094320.	6.881459	0.0000
C	1.78E+08	26557057	6.686173	0.0000
R-squared	0.985824	Mean dependent var		74111314
Adjusted R-squared	0.980669	S.D. dependent var		11145482
S.E. of regression	1549631.	Akaike info criterion		31.59524
Sum squared resid	2.64E+13	Schwarz criterion		31.83667
Log likelihood	-247.7619	Hannan-Quinn criter.		31.60760
F-statistic	191.2370	Durbin-Watson stat		1.877272
Prob(F-statistic)	0.000000			

El modelo es:

$$\widehat{\text{Viajeros}} = 1.78\text{E}+08 + 0,338783 \text{ pib} - 4.199.858 \text{ ipcresyhotel} - 8.55\text{E}+08 \text{ ficticia} + 7.530.515 \text{ ficticia*ipcresyhotel} + \varepsilon$$

Finalmente, el modelo mejora considerablemente, donde el R^2 ajustado es de un 98,06%, y donde todas las variables son significativas, con lo que seguiré el estudio con este último modelo.

3.7. Análisis de las hipótesis clásicas

El modelo seleccionado finalmente es el siguiente:

$$\hat{\text{Viajeros}} = 1.78\text{E}+08 + 0,338783 \text{ pib} - 4.199.858 \text{ ipcresyhotel} - 8.55\text{E}+08 \text{ ficticia} + 7.530.515 \text{ ficticia*ipcresyhotel} + \varepsilon$$

Como he comentado en apartados anteriores, este es el modelo que mayor R^2 ajustado tiene y con respecto a la bondad del ajuste, el 98,58% de la variabilidad del modelo es explicada por los regresores.

A continuación analizaré si en el modelo se cumplen las hipótesis clásicas: homocedasticidad, incorrelación y normalidad.

3.7.1. Homocedasticidad

En un modelo de regresión lineal por Mínimos Cuadrados Ordinarios, la heteroscedasticidad hace referencia a distinta varianza de las perturbaciones. El siguiente contraste, que voy a realizar es el de White que tiene por hipótesis:

H_0 : Homocedasticidad

H_1 : Heteroscedasticidad

El estadístico que sigue una χ^2 de Pearson:

Heteroskedasticity Test: White

F-statistic	0.673537	Prob. F(9,6)	0.7151
Obs*R-squared	8.041012	Prob. Chi-Square(9)	0.5300
Scaled explained SS	5.255206	Prob. Chi-Square(9)	0.8115

Como se puede comprobar en la tabla del programa EViews, tiene un p valor de 0,5300 que es mayor que 0,05, con lo que se rechaza la hipótesis nula del modelo, así que se acepta que las varianzas de las perturbaciones del modelo son todas iguales, con lo que no hay heteroscedasticidad.

3.7.2. Incorrelación

La autocorrelación rompe con la hipótesis de incorrelación entre las perturbaciones del modelo.

El contraste que voy a realizar en el Breuch-Godfrey y tiene por hipótesis:

H_0 : Incorrelación

H_1 : Autocorrelación

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	2.359634	Prob. F(2,9)	0.1500
Obs*R-squared	5.503813	Prob. Chi-Square(2)	0.0638

El contraste de hipótesis sigue una χ^2 de Pearson que resulta 0,0638 que es mayor que 0,05, con lo que no rechazo la hipótesis nula del modelo, luego hay incorrelación, la covarianza de las perturbaciones es igual a cero, son independientes.

3.7.3. Normalidad

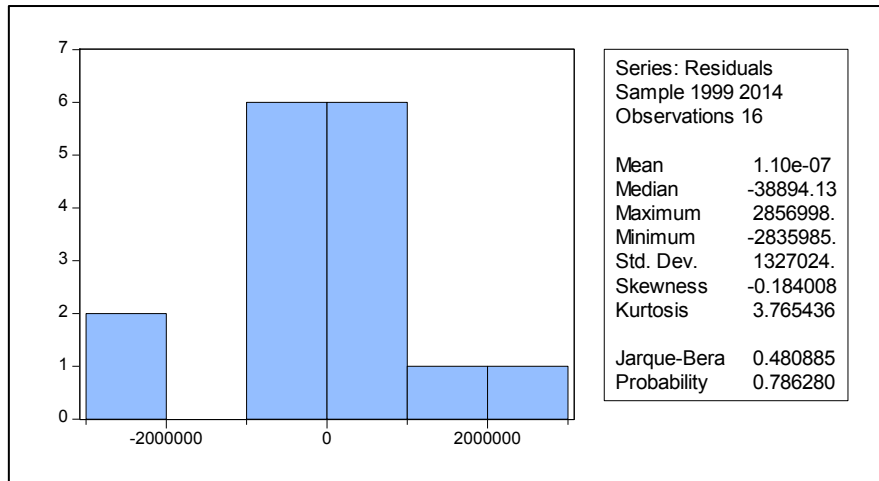
La hipótesis de normalidad se estudia a través del contraste de Jarque-Bera, que analiza las perturbaciones a partir de la forma de la distribución de los residuos, ya que las perturbaciones son variables no observables.

H_0 : Normalidad

H_1 : No normalidad

El estadístico del contraste sigue una χ^2 de Pearson y es el siguiente:

$$d_{JB} = \frac{N + k - 1}{6} \left(g_1^2 + \frac{1}{4}(g_2 - 3)^2 \right)$$



Como se puede observar en la salida de EViews, el estadístico es 0,4808, con un p valor de 0,7862, que es mayor que 0,05 por lo que se rechaza la hipótesis nula y se acepta que las perturbaciones del modelo siguen una Normal.

3.8. Análisis del modelo finalmente elegido

En el modelo finalmente elegido analizaré los coeficientes, la bondad del ajuste y la significación individual y conjunta.

El modelo era:

Dependent Variable: VIAJEROS
 Method: Least Squares
 Date: 06/22/15 Time: 10:33
 Sample: 1999 2014
 Included observations: 16

Variable	Coefficient	Std. Error	t-Statistic	Prob.
PIB	0.338783	0.047339	7.156489	0.0000
IPCRESYHOTEL	-4199858.	704642.7	-5.960266	0.0001
FICTICIA	-8.55E+08	1.24E+08	-6.910326	0.0000
FICTICIA*IPCRESYHOTEL	7530515.	1094320.	6.881459	0.0000
C	1.78E+08	26557057	6.686173	0.0000
R-squared	0.985824	Mean dependent var		74111314
Adjusted R-squared	0.980669	S.D. dependent var		11145482
S.E. of regression	1549631.	Akaike info criterion		31.59524
Sum squared resid	2.64E+13	Schwarz criterion		31.83667
Log likelihood	-247.7619	Hannan-Quinn criter.		31.60760
F-statistic	191.2370	Durbin-Watson stat		1.877272
Prob(F-statistic)	0.000000			

Cuya ecuación es la siguiente:

$$\hat{\text{Viajeros}} = 1.78\text{E}+08 + 0.338783 \text{ pib} - 4.199.858 \text{ ipcresyhotel} - 8.55\text{E}+08 \text{ ficticia} + 7.530.515 \text{ ficticia} * \text{ipcresyhotel} + \varepsilon$$

El coeficiente de la constante representa el valor de la variable endógena cuando todos los regresores son nulos.

El valor estimado de β_1 indica que, si aumenta el Producto Interior Bruto en una unidad, los viajeros aumentarán en 0,3387. Este resultado indica una relación positiva entre las variables, con lo que a medida que se incremente el PIB en España supondrá un incremento del número de viajeros en menos de una persona.

El valor estimado que resulta de β_2 nos indica que, cuando aumenta en una unidad el Índice de Precios al Consumo (restaurantes y hoteles), los viajeros disminuyen en 4.199.858 personas. Esto indica que hay una relación negativa entre las variables, cuando se incrementa la variable explicativa en función de la oferta del mercado de turismo, supone una disminución del número de viajeros. Esta relación de entrada es coherente, ya que a medida que se incrementen los precios sobre restaurantes y hoteles, disminuirá la demanda turística.

El valor estimado que resulta de β_3 , vale $- 8.55\text{E}+08$ de la aditiva, significa que a partir de 2008 el término constante del modelo vale 178.000.000 más 855.000.000, que resulta 1.033.000.000, mientras que antes de 2008 era de 178.000.000.

Por último, el valor estimado que resulta de β_4 es de 7.530.515 de la multiplicativa, significa que a partir de 2008 si el Índice de Precios al Consumo de restaurantes y hoteles se incrementa en un punto, el número de turistas se incrementará en la suma de $- 4.199.858$ y 7.530.515, que resulta 3.330.657, mientras que antes de 2008 disminuía en 4.199.858 viajeros.

Para estudiar la significación individual de las variables exógenas utilizaré el estadístico t-Student, a través del p valor para el contraste siguiente:

$$H_0: \beta_j = 0 \quad \text{siendo } j = 1,2,3,4$$

$$H_1: \beta_j \neq 0 \quad \text{siendo } j = 1,2,3,4$$

Con un nivel de significación del 5%, es decir $p < 0,05$, como se puede comprobar en la tabla del programa Eviews expuesta anteriormente, como dicho p valor para cada una de las variables exógenas los valores son 0,0000, 0,0001, 0,0000 y 0,0000 del PIB, IPC, ficticia y ficticita*IPC respectivamente, se concluye que, en todos los casos se rechaza la hipótesis alternativa, por lo tanto, todas las variables exógenas aparentemente son significativas individualmente.

Para comprobar la significación global del modelo se utiliza el estadístico F-Statistic, que toma un valor de 191,2370, obtenido mediante la fórmula siguiente:

$$F = \frac{R^2(n - K - 1)}{(1 - R^2)K}$$

Siendo $K= 4$, que es el número de regresores; y $n = 16$, que es la muestra.

En este caso el planteamiento del contraste sería:

$$H_0 = \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

$$H_1 = \text{Alguno de los parámetros } \beta_1, \beta_2, \beta_3, \beta_4 \neq 0$$

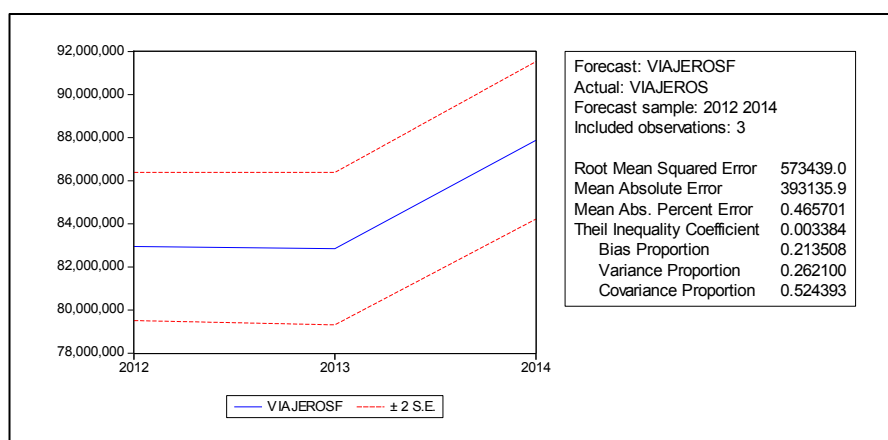
Con un nivel de significación del 5%, es decir, $p < 0,05$, como su valor es inferior, se rechaza la hipótesis nula. Por lo tanto se puede afirmar que las variables son, conjuntamente, significativas.

En referencia a la bondad de ajuste, con un valor de $R^2 = 0,985824$, el modelo tiene buen ajuste, ya que el 98,58% de la variabilidad del número de viajeros es explicada por los regresores del modelo, Producto Interior Bruto, Índice de Precios al Consumo (restaurantes y hoteles) y personal.

3.9. Predicción

Para poder obtener valores fuera de la muestra, he realizado una predicción de la variable endógena. Es necesario un buen comportamiento del modelo durante el periodo de la muestra para poder realizar una predicción fiable. Además se necesitará conocer los posibles valores que tomarán las variables explicativas, que estimaré en función de una valoración de la base de datos de 1999 a 2014. Y por último, es necesario que el modelo mantenga la misma estructura de los parámetros, que sean similares y el efecto de la predicción sea el mismo al de la muestra.

Primero se deberá determinar si el modelo es adecuado para predecir, estudiando la capacidad predictiva. Esta se estudia con el coeficiente de desigualdad de Theil, que toma valores entre 0 y 1, cuanto más próximo a cero más adecuado será el modelo para predecir. Lo analizamos en cuatro años, de 2012 a 2014.



Como se puede comprobar el coeficiente Theil es 0,003384, próximo a 0, lo que indica una buena capacidad predictiva, además, la proporción del sesgo (0,2135) y la varianza (0,2621) son también muy pequeñas, mientras que la mayor proporción recae en la covarianza (0,5243).

Lo siguiente que hay que hacer es ampliar el rango o la muestra. Como sólo voy a predecir un año más, lo amplio hasta 2015, y doy valores a las variables explicativas: Producto Interior Bruto, Índice de Precios al Consumo (restaurantes y hoteles) y la ficticia, donde para el año 2015 va a tomar valor 1.

obs	PIB	IPCRESYHOTEL	FICTICIA
1999	5.95E+08	76.96650	0.000000
2000	6.46E+08	80.22934	0.000000
2001	7.00E+08	83.76779	0.000000
2002	7.49E+08	88.40509	0.000000
2003	8.03E+08	92.21451	0.000000
2004	8.61E+08	95.94944	0.000000
2005	9.31E+08	99.99237	0.000000
2006	1.01E+09	104.4680	0.000000
2007	1.08E+09	109.5295	0.000000
2008	1.12E+09	114.6927	1.000000
2009	1.08E+09	116.9139	1.000000
2010	1.08E+09	118.0983	1.000000
2011	1.08E+09	119.9574	1.000000
2012	1.06E+09	121.0093	1.000000
2013	1.05E+09	121.5863	1.000000
2014	1.06E+09	122.1525	1.000000
2015	1.07E+09	122.0000	1.000000

Para el año 2015, he estimado que el Producto Interior Bruto en España será de 1.070.000.000 millones de euros, ya que el Fondo Monetario Internacional (FMI) que ha revisado al alza las perspectivas de crecimiento para España augura un aumento de 2,5%.

He estimado el Índice de Precios al Consumo (restaurantes y hoteles), en 122,000, que disminuye muy poco tras observar la apenas variación en los años anteriores. Además según un estudio realizado por Trivago Hotel Price Index, que es un indicador que muestra los precios actuales del mercado, el precio de los hoteles en España ha subido un 10% de media.

Y por último, para la variable ficticia he dado valor 1 al año 2015, ya que el cambio estructural realizado en apartados anteriores a partir del año 2008 tomaban valor 1.

Posteriormente he realizado la predicción para el año 2015. Como se puede comprobar en la tabla que aparece a continuación, los valores observados en el periodo muestral son similares a los predichos para esos años.

obs	VIAJEROS	VIAJEROSF
1999	58588944	58588944
2000	59282522	59282522
2001	59904301	59904301
2002	59868812	59868812
2003	62531380	62531380
2004	66831268	66831268
2005	70629025	70629025
2006	81855900	81855900
2007	84423433	84423433
2008	82998877	82998877
2009	77140316	77140316
2010	81888870	81888870
2011	85366976	85366976
2012	82962481	82962481
2013	83820919	83820919
2014	87687000	87687000
2015	NA	91277701

El resultado es de 9.191.277.701, es decir, el número de viajeros que se estima en 2015, con los datos de las variables explicativas estimados anteriormente, sería de 9.191.277.701.

4. CONCLUSIONES

A lo largo del presente trabajo y mediante los modelos econométricos planteados, se ha podido comprobar las diferentes relaciones que existe entre algunas de las variables del sector turístico.

Para ello se procedió a formular un modelo donde la variable dependiente o endógena era el número de viajeros, que posteriormente fui relacionando con las variables endógenas que supuse en un primer momento iban a estar más correlacionadas, como el Índice de Precios Hoteleros (restaurantes y hoteles), el Índice de Precios al Consumo, el Producto Interior Bruto, establecimientos abiertos, personal contratado y empleados.

Algunas de las conclusiones que se pueden extraer son las siguientes:

- La situación económica del país medida a través del Producto Interior Bruto (PIB) es un factor que influye mucho en la demanda turística.
- Los precios de los servicios turísticos también afectan de forma importante, ya que a medida que estos aumentan, afectan de forma negativa al número de viajeros, relación demanda precio, donde a medida que se incrementa la demanda, disminuye el precio, pero esa relación cambia de signo a partir de 2008, como se observa con la introducción de una ficticia multiplicativa con los precios.
- Se ha observado también que la crisis económica ha afectado directamente al turismo en España. Los coeficientes del modelo no han permanecido estables para todo el periodo muestral sino que han cambiado a partir de 2008.
- La primavera árabe, aunque ha afectado al turismo español, ya que toda esa demanda turística cuyo destino era el norte de África se desvió, no han alterado la estabilidad de los parámetros del modelo estimado.

5. BIBLIOGRAFÍA

ABC (2015): “*La Primavera Árabe ha desviado a España más de seis millones de turistas*”. Disponible en: <http://www.abc.es/economia/20150402/abci-turismo-espana-primavera-arabe-201504012203.html>, [consulta: 12/05/2015].

Balance del Turismo 2010 del Instituto de Estudios Turísticos (IET), Ministerio de Industria, Turismo y Comercio.

Encuesta de Gasto Turístico (EGATUR). Disponible en: <http://www.iet.tourspain.es/es-es/estadisticas/egatur/paginas/default.aspx>, [consulta:15/05/2015].

Encuesta de Movimientos Turísticos de los españoles (Familitur). Disponible en: <http://www.iet.tourspain.es/es-es/estadisticas/familitur/paginas/default.aspx>, [consulta: 02/06/2015].

Encuesta de Movimientos Turísticos en Fronteras (FRONTUR). Disponible en: <http://www.iet.tourspain.es/es-es/estadisticas/frontur/paginas/default.aspx>, [consulta: 12/05/2015].

Espasa, A., Gómez-Churruca, R. y Jareño, J. (1990): *Un análisis econométrico de los ingresos por turismo en la economía española*, Documento de Trabajo 9002, Madrid, Banco de España.

Esteban Talaya, A., (1987): *Análisis de la demanda: Aplicación a la actividad turística de las técnicas de predicción*, Madrid.

Esteban Talaya, A., Reinares Lara, E., (1996): *La investigación de la demanda turística en España: recopilación y análisis*.

Gardella, J. A., y Aguayo, E.: *Análisis econométricos de la demanda turística internacional en la Can*.

Guisán, C., Neira I., (2001): *Un análisis econométrico del turismo hotelero y extrahotelero en las regiones y provincias españolas.*

Guisán, M., Aguayo, E., (2001-2008): *Evolución del turismo hotelero en las regiones españolas.*

Hosteltur (2015): *“El turismo incrementa su aportación en el PIB”*. Disponible en: http://www.hosteltur.com/155130_turismo-incrementa-su-aportacion-al-pib.html, [consulta: 02/06/2015]

Instituto de Estudios Turísticos del Ministerio de Industria, Turismo y Comercio: Resultados de la actividad turística en España (2010). Disponible en: <http://www.iet.tourspain.es/es-ES/estadisticas/analisis-turistico/balantur/anuales/Balance%20turismo%20en%20Espa%C3%B1a%20en%202010.pdf>, [consulta: 02/06/2015].

Instituto Nacional de Estadística (INE). Disponible en: <http://www.ine.es/> [consulta: 02/06/2015].

Martin, C.A. y Witt, S.F. (1987): *“Tourism Demand Forecasting Models: Choice of Appropriate Variable to Represent Tourists’ Cost of Living, International Journal of Tourism Management”*.

Movimientos Turísticos en Fronteras (FRONTUR) (2006). Disponible en: <http://www.iet.tourspain.es/es-ES/estadisticas/frontur/Anuales/Movimientos%20Tur%C3%ADsticos%20en%20Fronteras%20%28Frontur%29%202006.pdf>, [consulta: 15/05/2015].

Organización Mundial del turismo (OMT). Disponible en: <http://www2.unwto.org/es>
http://www.hosteltur.com/121080_turismo-rural-cifras-debate.html, [consulta: 08/05/2015].

Ortega, E. y Peñalosa, J., (2012): *Claves de la crisis económica española y retos para crecer en la UEM.*

Pulido San Roman, A., (1996): *Introducción a un análisis econométrico del turismo.*

Sánchez Rivero, M., (1998): *Modelización estadística de tablas de contingencia: aplicación al análisis de la demanda turística española.*

Wooldridge, J.: *Introducción a la econometría. Un enfoque moderno.*