

Proyecto Fin de Máster

Organización Industrial y Gestión de Empresas

Un Análisis econométrico del precio de la vivienda
en Sevilla en el año 2017

Autor: Sonia del Rey Simeón

Tutor: Fernando Núñez Hernández

Dpto. Organización Industrial y Gestión de
Empresas I

Escuela Técnica Superior de Ingeniería

Sevilla, 2018



Proyecto Fin de Máster
Organización industrial y Gestión de empresas

Un Análisis econométrico del precio de la vivienda en Sevilla en el año 2017

Autor:
Sonia del Rey Simeón

Tutor:
Fernando Nuñez Hernández
Profesor Contratado Doctor

Dpto. de Organización Industrial y Gestión de Empresas I
Escuela Técnica Superior de Ingeniería
Universidad de Sevilla
Sevilla, 2018

Proyecto Fin de Máster: Un Análisis econométrico del precio de la vivienda en Sevilla en el año 2017

Autor: Sonia del Rey Simeón

Tutor: Fernando Nuñez Hernández

El tribunal nombrado para juzgar el Proyecto arriba indicado, compuesto por los siguientes miembros:

Presidente:

Vocales:

Secretario:

Acuerdan otorgarle la calificación de:

Sevilla, 2018

El Secretario del Tribunal

A mi familia

A mis maestros

Agradecimientos

Un agradecimiento singular debo al profesor Fernando Núñez Hernández que, como tutor de este proyecto, me ha orientado, apoyado y corregido en mi labor con un interés y una entrega insuperable. Por todas las atenciones y por el tiempo que me ha dedicado.

A mis compañeros de la escuela que, aunque no les mencione de forma explícita, no les puedo negar un sincero agradecimiento.

En general a todos los profesores del MOIGE que influyeron con sus lecciones y experiencias en formarme como una persona cualificada y preparada en el sector.

A esas personas importantes en mi vida que siempre han estado listas para brindarme toda su ayuda y me han animado de una manera u otra, para terminar este proyecto.

A todos, muchas gracias.

Resumen

A diferencia de lo que ocurre en los mercados competitivos, la vivienda es un bien muy heterogéneo que determina que los compradores tengan que considerar un elevado número de atributos a la hora de comprarlo; hasta el punto, que podríamos considerar que no existen dos viviendas iguales. Uno de los primeros pasos para intercambiar una vivienda consiste en determinar y acordar su precio, el cual variará en función de los diferentes atributos del inmueble: la ubicación, la superficie que posee, el número de habitaciones, la altura, la situación del mercado en ese momento, etc. El objeto del presente estudio es analizar el precio de la vivienda en Sevilla capital aplicando la metodología de precios hedónicos propuesta por Rosen, la cual permite identificar aquellos atributos de la vivienda que más inciden en el precio de la misma mediante la estimación de un modelo de regresión lineal múltiple.

Abstract

Unlike what happens in competitive markets, housing is a very heterogeneous property that determines that buyers must consider a large number of attributes when buying it; To the point, we could consider that there are no two identical houses. One of the first steps to exchange a house is to determine and agree on its price, which will vary depending on the different attributes of the property: the location, the area it has, the number of rooms, the height, the market situation in that moment, etc. The object of the present study is to analyze the price of housing in Seville capital applying the methodology of hedonic prices proposed by Rosen, which allows identifying those attributes of the house that most affect the price of the same by estimating a model of multiple linear regression.

ÍNDICE

Agradecimientos	5
Resumen	6
Abstract	7
ÍNDICE	8
Índice de Tablas	9
Índice de Figuras	10
1 Introducción	11
<i>1.1 Objetivos y justificación del trabajo.</i>	<i>13</i>
<i>1.2 Estructura del trabajo.</i>	<i>14</i>
2 Enfoque teórico y revisión de la literatura	15
<i>2.1 Marco teórico de referencia.</i>	<i>15</i>
<i>2.2 Revisión de la Literatura.</i>	<i>23</i>
3 La metodología econométrica. Regresión lineal múltiple.	29
4 Descripción de la muestra	41
5 Estimación de un modelo hedónico del precio de la vivienda en Sevilla capital en el año 2017.	49
6 Conclusiones.	57
Referencias.	60

ÍNDICE DE TABLAS

Tabla 2.1 Resumen artículos literatura.	28
Tabla 3.1 Terminología en la regresión simple.	31
Tabla 3.2 Interpretación de los coeficientes en el modelo “lineal”.	35
Tabla 4.1 Distribución de la muestra por zonas de Sevilla capital.	42
Tabla 4.2 Distribución del número de habitaciones y de baños. Unidades y porcentajes.	46
Tabla 4.3 Distribución de la planta del piso.	47
Tabla 4.4 Distribución de otros atributos de la vivienda. Unidades y porcentajes.	48
Tabla 5.1 Estimación del precio de la vivienda en Sevilla capital.	51
Tabla 5.2 Estimación del precio de la vivienda en Sevilla capital diferenciando por tamaño.	53

ÍNDICE DE FIGURAS

Figura 2.1 La 'lista' de precios hedónicos para el atributo z_i .	16
Figura 2.2 Elección de la vivienda óptima utilizando la función de disposición a pagar por z_i .	19
Figura 2.3 Decisión óptima de oferta utilizando la función de disposición a cobrar por z_i .	20
Figura 2.4 El intercambio en el mercado.	21
Figura 2.5 RNA topología seleccionada.	25
Figura 3.1 Relación teoría económica y econometría.	29
Figura 3.2 Estimadores.	30
Figura 3.3 Homocedasticidad.	33
Figura 3.4 Autocorrelación.	34
Figura 3.5 Mínimos cuadrados ordinarios.	36
Figura 3.6 Recta poblacional con recta de regresión.	37
Figura 4.1 División de Nervión en 3 sub-zonas.	43
Figura 4.2 Distribución de las variables precio, m2 construidos y precio por m2 construido.	44
Figura 4.3 Relación entre el precio de oferta y el número de m2 construidos.	45
Figura 4.4 Relación entre el precio por m2 construido y el número de m2 construidos.	45
Figura 4.5 Relación entre el precio, los m2 construidos y el número de habitaciones.	47
Figura 5.1 Relación estimada entre el precio requerido y los m2 construidos (diferenciando por tamaño en 105 m2).	54
Figura 5.2 Relación entre el precio requerido y el precio estimado.	55

1 INTRODUCCIÓN

Durante la década previa al estallido de la crisis económica en el año 2007, algunos factores han hecho aumentar el precio de la vivienda en España a niveles inadmisibles. Entre estos factores, destacan la *edad de emancipación* de la generación más numerosa de España (año 1975), el importante *crecimiento de la renta disponible* de los hogares (debido a la reducción de las tasas de desempleo y el aumento de los niveles salariales), *los bajos tipos de interés* (variables) y las *atractivas condiciones hipotecarias*, según el Instituto de Estudios Fiscales. Todos estos factores impulsaron un aumento especulativo del precio de la vivienda en España entre los años 1998 y 2007, dando lugar a subidas registradas de más del 282% según el Instituto Nacional de Estadística.

La demanda especulativa, consistía en comprar viviendas para revenderlas años después a un precio esperado mayor; esto animó a su vez a los jóvenes a comprar una vivienda (incorporándose a la demanda), ya que si no se daban prisa “no podrían comprar jamás”. En ese momento (primer lustro de la década de los 2000), los costes de construcción de la vivienda media (81.000€ según Ministerio de Fomento) estaban muy por debajo del precio medio de venta, por tanto los beneficios de la construcción eran relativamente elevados. Ante esta tesitura, aumentaron el número de empresas constructoras y se construyó mucha más vivienda de lo normal (todo se vendía).

En el punto culminante de la burbuja, se llegaron a construir unas 800.000 viviendas en el año 2006 (mucho más que en Inglaterra, Alemania y Francia juntas) frente a una media de creación de 390.000 nuevos hogares (2001-2011) según el Instituto Nacional de Estadística. A pesar de que había mucha oferta, se producían subidas de precios, por lo que existía un antimercado (caso extremo de burbuja donde precios y cantidades de vivienda nueva están en relación directa y no inversa, como indica la ley de la oferta y la demanda) según Verges (2012).

Los bancos concedieron crédito sin revisar las necesarias garantías de solvencia de cada prestatario, muchos de ellos con recursos insuficientes.

En el año 2007 los precios de la vivienda tocan techo en la economía española y comienza una caída significativa de la demanda de vivienda, coincidiendo además con el efecto de las hipotecas subprime en EEUU, que genera problemas de estructura financiera a los bancos españoles. Además, los tipos de interés se habían mantenido altos durante la burbuja. Todo este escenario dio lugar a finales de la década de los 2000 a un crecimiento de los impagos y de la morosidad, y una caída importante de la demanda de vivienda. Esto provocó que muchas de las viviendas que acabaron de construirse entre los años 2008 y 2009 se quedaran sin vender, llegando a tener la economía española en el año 2011 un stock de 626.670 viviendas de nueva construcción sin ocupar según el Informe sobre stock de vivienda nueva del Ministerio de Fomento y un total de 2.816.695 millones de viviendas de segunda mano vacías, según el Censo de Población y Vivienda de 2011

del INE.

Esta sobreoferta de 3.443.365 millones de viviendas es uno de los datos más significativos de la crisis económica que afectó a la economía española a partir de 2008. El exceso de oferta provocó una disminución significativa del precio de la vivienda (sobre todo en determinadas zonas), llegando a tocar fondo dicho precio en el año 2012 con una caída del 12.8 %. En el segundo trimestre del 2014 por el contrario, se empezó a ver cierta recuperación del mercado inmobiliario con crecimientos de precios del 0.8% según el INE. Sin embargo, ha sido ya a partir de 2016, con el EURIBOR otra vez en niveles relativamente bajos, cuando se observan de nuevo crecimientos del precio de la vivienda significativos, llegando a crecer un 7,2% en el cuarto trimestre de 2017 en tasa interanual según el Instituto Nacional de Estadística (INE), siendo ésta la mayor revalorización de las viviendas desde el año 2007. Actualmente, el precio de la vivienda ha vuelto a tomar una rinda alcista, aunque estamos todavía a un 22.71% por debajo de 2007, ya que nos encontramos con 1587,9 €/m² en el segundo trimestre del 2018, en contraposición al 2054,5 €/m² en el segundo trimestre del 2007.

A pesar de que la situación actual en el mercado inmobiliario no es tan desfavorable como la de hace años, el futuro es incierto. En la actualidad, comprar una vivienda no es una opción para la mayoría de los ciudadanos, ya que el banco cada vez pide más requisitos a la hora de conceder hipotecas con el objeto de no cometer los mismos errores del pasado. Hoy en día, para la mayoría de los jóvenes es prácticamente inviable comprar una vivienda. La movilidad geográfica en el mercado de trabajo, la inestabilidad laboral y la precariedad salarial frenan el deseo de buena parte de la población de adquirir vivienda mediante crédito hipotecario. Este cambio de mentalidad del demandante (frente a lo acontecido en la burbuja), junto con la recuperación de la economía española y la fuerte presión turística, están provocando un aumento importante de la demanda de alquiler y del precio de alquiler de la vivienda. El año 2017 fue el año de aumento récord interanual del precio de alquiler, el crecimiento fue del 18.4% frente a los crecimientos en torno al 10-15% de una década atrás. Las regiones con mayores aumento del precio del alquiler con respecto al año anterior han sido el archipiélago canario, el balear y la comunidad de Andalucía con crecimientos en 2017 del 27.8%, 22.1% y 15.5% respectivamente según Idealista. Sin embargo, las autonomías más caras fueron Madrid y Barcelona (con precios de alquiler medios de 14.4 €/m² al mes), siguiéndole Baleares (12.5 €/m²) y Euskadi (11.2 €/m²). Al otro lado se encuentra Extremadura (4.1 €/m²), Castilla la Mancha (4.7 €/m²) y La Rioja (5.2 €/m²), siendo las comunidades más económicas. En el segundo trimestre de 2018, el precio del alquiler ya ha aumentado un 2.4%, estando el precio por metro cuadrado en 10.8 €/m² mensuales, según el Informe de Evolución de los Precios del alquiler publicado por Idealista. En tasa interanual se alcanza el 15.6%.

Con todo este panorama actual con precios de compra moderados (aunque en aumento) y precios de alquiler muy elevados (sobre todo en Madrid y Barcelona), podríamos pensar que la compra de vivienda está ganando atractivo, aunque sólo está al alcance de ciertos individuos –pensamos en una persona consolidada a nivel laboral (en torno a los 40-45 años) y con dinero suficiente para afrontar los gastos iniciales asociados a la compra del piso. La población joven que se acaba de incorporar al mercado laboral puede verse atrapada en

una especie de trampa de la precariedad consistente en tener que pagar altos alquileres por no poder afrontar la compra de vivienda.

Otro rasgo reciente de nuestro mercado de vivienda es que la vivienda de obra nueva es cada vez más demandada en España. En 2017, del total de los compradores de vivienda, el 8% habían valorado únicamente entre la nueva construcción, el 57% consideró tanto obra nueva como segunda mano y un 35% solo buscó vivienda de segunda mano. Por el contrario en 2018, estos porcentajes pasaron a ser del 15%, 52% y 33% respectivamente. En lo que va de año del 2018, el 29% de los compradores adquirió finalmente vivienda de obra nueva, siendo el restante 71% los que se decantaron por vivienda de segunda mano, según Fotocasa. Las comunidades autónomas que más interés tienen en obra nueva son Madrid (pasa del 11% en 2017 al 18% en 2018) y Andalucía (del 7% al 15% respectivamente).

Este aumento de la demanda no acababa de surgir por la situación económica, pero también por la escasez de oferta que se adecuara a las necesidades de los compradores. Una de las principales razones por las que se elige vivienda nueva es por su modernidad (gimnasios, piscina, pádel...) y por su calidad en la construcción (disminuye el consumo de energía), motivos por los cuales hacen ser muy valoradas por gente joven y con hijos.

No obstante, el comprador que no disponga del dinero necesario para comprar una vivienda de obra nueva y que se plantee pedir una hipoteca, no notará apenas diferencia con respecto a la vivienda de segunda mano; Según los últimos datos del Consejo General del Notariado, el precio por metro cuadrado de los pisos de segunda mano en julio de 2017 en España era de 1.487€ mientras que el de los pisos nuevos estaba en 1.795€. Es decir, un piso de 100 m² de segunda mano costaría casi 150.000€ frente a 180.000€ en el caso de obra nueva. Como se observa, la diferencia de 30.000€ no es una cifra alarmante y que deberíamos tener en cuenta a la hora de tomar una decisión, siempre y cuando para nosotros sea el precio la variable más determinante a la hora de comprar una vivienda y no la localización, ya que la mayoría de las viviendas de obra nueva en la actualidad, se están construyendo en el extrarradio de las ciudades, motivo de renuncia en muchas ocasiones.

1.1 Objetivos y justificación del trabajo.

Como se puede observar, el mercado de la vivienda se caracteriza por ser muy dinámico (muy ligado al ciclo de actividad) y heterogéneo, por tanto nos hacemos las siguientes preguntas: ¿qué variables son las más determinantes a la hora de demandar una vivienda u otra? ¿De qué factores depende fundamentalmente el precio de la vivienda en España en la actualidad?

En este trabajo se introduce una fórmula empírica que permite analizar el precio de la vivienda en función de las propiedades que posee en un momento determinado. El enfoque aplicado es el de precios hedónicos de Rosen (1974), que estudia cómo el conjunto de características que posee un bien heterogéneo condiciona o explica su precio de mercado. El ámbito espacial del estudio abarca algunos de los barrios más

importantes de la ciudad de Sevilla en el año 2017. Los datos empleados han sido extraídos del portal web de vivienda Idealista.com.

Los principales objetivos que se quieren conseguir con el presente Trabajo de Fin de Máster son los que se detallan a continuación:

1. Estudiar la literatura existente sobre el funcionamiento del mercado de la vivienda.
2. Exponer un modelo teórico sobre el precio de la vivienda.
3. Describir la situación actual del precio de la vivienda en Sevilla.
4. Aplicar y analizar un modelo econométrico que tiene como variable endógena el precio de la vivienda en Sevilla y como variables explicativas las características fundamentales de cada inmueble.
5. Exponer las principales conclusiones obtenidas.

1.2 Estructura del trabajo.

El presente Trabajo Fin de Máster se divide en seis capítulos. El primer capítulo tiene carácter introductorio, en él se explican los objetivos del trabajo y la situación actual del precio de la vivienda en España. En el segundo capítulo se presenta el enfoque teórico empleado para fundamentar el análisis del precio de la vivienda y una revisión de la literatura sobre los determinantes de dicho precio. En el tercer capítulo se explica la metodología econométrica utilizada en este Trabajo Fin de Máster para analizar el precio de la vivienda en Sevilla capital: la regresión hedónica (lineal múltiple). En el cuarto capítulo se describe la muestra utilizada para nuestro análisis. En el quinto capítulo se describen los resultados de la estimación del modelo hedónico de precios de la vivienda con datos de Sevilla capital en el año 2017. En el sexto y último capítulo se recogen las conclusiones del análisis desarrollado.

2 ENFOQUE TEÓRICO Y REVISIÓN DE LA LITERATURA

2.1 Marco teórico de referencia.

El marco teórico de este Trabajo sigue el enfoque de precios hedónicos propuesto por Rosen (1974). Este autor propone un modelo de diferenciación de producto basado en la hipótesis hedónica de que los bienes son valorados por la utilidad asociada a cada uno de sus atributos. Los precios hedónicos son definidos como precios implícitos asociados a cada atributo y son mostrados a los agentes a través del precio total observado del producto, el cual viene definido por un conjunto o lote de atributos¹. Empíricamente, los precios implícitos de cada atributo o característica del bien son estimados mediante un análisis de regresión múltiple que tiene como variable endógena el precio final del bien y como variables explicativas el conjunto de atributos que lo definen; precisamente, ese será el objetivo de nuestro Trabajo, explicar el precio de la vivienda en Sevilla en 2017.

El trabajo de Rosen (1974) pretende fundamentalmente aportar un modelo teórico que permita obtener y explicar los coeficientes de regresión hedónicos. Para ello, el autor propone un modelo de equilibrio competitivo que se desarrolla en espacio de varias dimensiones en el que compradores y vendedores se sitúan. Cada tipo de bien viene definido por un vector z de n características $z = \{z_1, z_2, \dots, z_n\}$ que representa un punto en el espacio — z_i mide la cantidad de la característica i -ésima que contiene cada bien—. Los oferentes tienen la posibilidad de ofrecer diferentes lotes de características dando lugar a diferenciación de producto. Cada paquete de características viene dado por un precio $p(z) = p(z_1, z_2, \dots, z_n)$ que viene dado para oferentes y demandantes dado su pequeño peso o tamaño en relación al mercado —característica esta propia de los mercados en competencia, donde los participantes son precio-aceptantes—.

La condición de equilibrio del mercado establece que el precio $p(z)$ se ajustará para que el número de unidades ofrecido en cada punto del espacio se iguale al número de unidades demandadas en ese punto, dando lugar a un equilibrio espacial de precios que determinan la igualdad de oferta y demanda para cada clase de bien ofrecido. Como veremos a continuación, este equilibrio se basa en la conducta maximizadora de ambos lados del mercado, dada una determinada distribución de los gustos de los consumidores y de los costes de los productores.

Para desarrollar el modelo, consideremos que existe un mercado para cada clase de mercancía (en adelante, viviendas) descrita por sus n atributos $z = \{z_1, z_2, \dots, z_n\}$. Los consumidores miden objetivamente los atributos de la vivienda, pero pueden diferir en el valor subjetivo que conceden a cada paquete de

¹ Se le llama precio 'hedónico' porque está determinado por las diferentes cualidades del bien diferenciado y el 'placer' (en términos de utilidad económica) que aportan al comprador.

características –por ejemplo, dos familias pueden valorar de forma diferente una casa que dispone de 4 habitaciones–. La variedad de posibles vectores de atributos z es tan amplia que podemos analizar el problema en términos continuos, podemos pensar en términos de poder aplicar diferenciales a todos los atributos. Bajo este escenario continuo, en el mercado se establece una relación entre los niveles de los atributos de una vivienda y su precio²: $p(z)=p(z_1, z_2, \dots, z_n)$, relación que como hemos comentado puede ser aproximada empíricamente aplicando la técnica de regresión al precio observado de una muestra amplia de viviendas. La Figura 2.1 asume que el precio marginal asociado a un incremento unitario de z_i no es lineal, sino decreciente, lo cual se debe a la existencia de un efecto de saciedad en dicho atributo –utilidad marginal decreciente–.

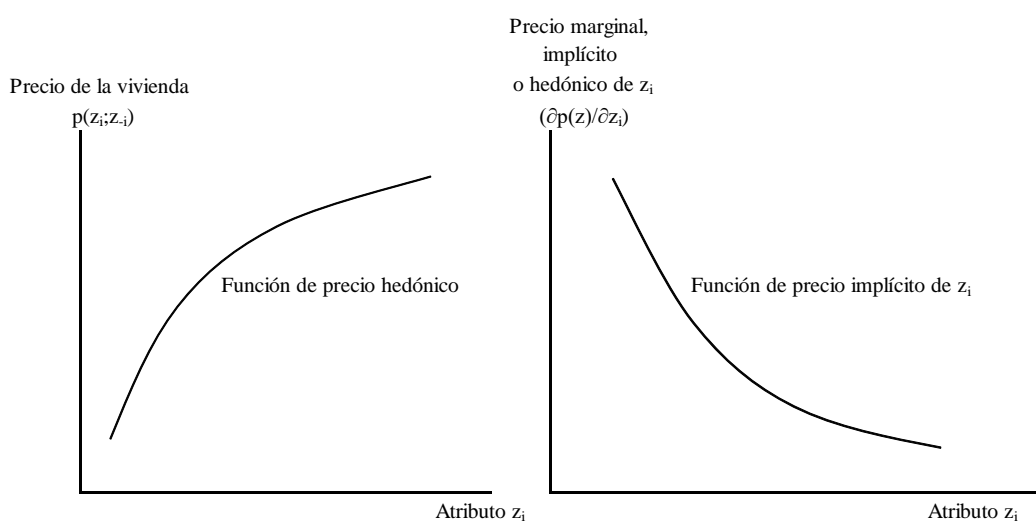


Figura 2.1 La ‘lista’ de precios hedónicos para el atributo z_i .

Como veremos a continuación, el modelo hedónico muestra como consumidores y oferentes se van a encontrar (van a intercambiar la mercancía) tomando como referencia la curva $p(z)$.

2.1.1 La decisión del consumidor.

Supongamos que el consumidor compra sólo una unidad de vivienda que tiene una determinada combinación de valores z que le conviene. La función de utilidad de este consumidor será $U(x, z_1, z_2, \dots, z_n)$; la cual se supone estrictamente cóncava –más unidades de un determinado atributo reportan cada vez menos utilidad marginal–. La variable x representa el resto de los bienes consumidos por el comprador cada período. Si fijamos el precio de x como unitario y expresamos la renta del individuo, y , en función de x , tenemos que: $y = x + p(z)$. La maximización de la función de utilidad sujeta a la restricción presupuestaria $y = x + p(z)$ da lugar al cumplimiento de las ecuaciones de Lagrange $\{L = U(z, x) + \lambda (y - x - p(z))\}$:

² Pudiera pasar en el mercado que un mismo set de características tuviera dos o más precios distintos, pero en ese caso, el precio que crearía mercado o intercambio sería el menor de ellos, el más competitivo.

$$\frac{\partial L}{\partial z_i} = U_{z_i} - \lambda p_{z_i} = 0; \quad \frac{\partial L}{\partial x} = U_x - \lambda = 0; \quad \frac{\partial L}{\partial \lambda} = y - x - p(z) = 0 \quad (2.1)$$

U_{z_i} es la derivada parcial de la función de utilidad con respecto a la característica de la vivienda z_i . Representa la utilidad adicional que proviene de elegir una vivienda con una unidad extra de la característica z_i , *ceteris paribus*. Por su parte, U_x es la derivada parcial de la función de utilidad con respecto al bien compuesto x , el cual, dado su precio unitario, representa el dinero gastado en otros bienes diferentes a la vivienda; U_x se puede interpretar como la utilidad adicional que aporta una unidad extra de dinero, *ceteris paribus*. Finalmente, la variable p_{z_i} representa el precio implícito del atributo z_i .

A partir de las dos primeras ecuaciones de Lagrange obtenemos la igualdad: $U_{z_i}/U_x = p_{z_i}$
 $\Rightarrow U_{z_i}/p_{z_i} = U_x/1$, que refleja la regla de óptimo (máxima utilidad) llamada *igualdad de las utilidades marginales ponderadas* (con $p_x = 1$), según la cual, el consumidor gastará unidades monetarias (u.m.) en x y en z_i , $i=1,2,\dots,n$, hasta que las utilidades marginales obtenidas por dichas u.m. gastadas se igualen –si fueran desiguales, sería posible coger una combinación de x y z_i que reportara un mayor nivel de utilidad al consumidor, por lo que no estaríamos maximizando su función de utilidad–. Por otro lado, la tercera ecuación de Lagrange ($\frac{\partial L}{\partial \lambda}$) representa la restricción presupuestaria del comprador: $y = x + p(z) \Rightarrow \Theta = y - x = p(z)$. ‘ Θ ’ representa la disposición a pagar por una vivienda, ya que es la renta que queda al consumidor después de haber pagado la compra del bien compuesto x .

Cuando analizamos la función de utilidad del consumidor en un determinado nivel de utilidad u_0 , obtenemos la denominada curva de indiferencia del consumidor: $U(x, z)=u_0$. La curva de indiferencia representa combinaciones de x y z que confieren el mismo nivel de bienestar o utilidad al comprador. Si despejamos x de la curva de indiferencia, obtenemos la expresión $x(z, u_0)$, que indicaría la cantidad de dinero que el hogar puede gastar en otros bienes, x , disfrutando de un nivel de utilidad u_0 y viviendo en una propiedad con las características z . Si esta relación se introduce en la restricción presupuestaria del comprador queda la siguiente expresión conocida con el nombre de función de disposición a pagar de Rosen (*bid function*):

$$\Theta = y - x(z, u_0) \quad \text{o} \quad \Theta(y, z, u_0) \quad (2.2)$$

La función de disposición a pagar (y, z, u_0) representa la cantidad máxima que un hogar pagaría por una propiedad con atributos z , de modo que pudiera alcanzar el nivel de utilidad u_0 , con sus ingresos y . Obsérvese que los aumentos de renta se traducen directamente (es decir, u.m. a u.m.) en aumentos de la disposición a pagar por el paquete de atributos de la vivienda z : $\frac{\partial \theta}{\partial y} = 1$. Asimismo, cuanto mayor sea el nivel deseado del atributo z_i , menor será la cantidad monetaria que puedo gastar en otros bienes x dado el nivel de utilidad u_0 y,

por tanto, mayor será la disposición a pagar por la vivienda $\theta = y - x : \frac{\partial \theta}{\partial z_i} > 0$. Finalmente, mayores niveles deseados de utilidad ($u_0 \rightarrow u_1$) requieren gastar más en x (dado el nivel de z_i) con la consiguiente caída de θ . $\frac{\partial \theta}{\partial u} < 0$. En el espacio (θ, z_i) , existirán tantas curvas de Rosen como niveles de utilidad queramos imaginar, por ello hablamos de una familia de curvas de disposición a pagar asociadas a los niveles de utilidad u .

La pendiente de (y, z, u_0) respecto a un atributo determinado de la vivienda z_i va a ser $\frac{\partial \theta}{\partial z_i} = -\frac{\partial x}{\partial z_i}$, donde la pendiente de la curva de indiferencia $x(z, u_0)$ respecto a z_i viene dada por

$$du_0 = 0 = \frac{\partial U}{\partial x} dx + \frac{\partial U}{\partial z_i} dz_i \Rightarrow \frac{\partial x}{\partial z_i} = -\frac{U_{z_i}}{U_x} \quad (2.3)$$

Luego la pendiente de $\theta(y, z, u_0)$ va a ser positiva y va a venir dada por el cociente de las utilidades marginales: $\frac{\partial \theta}{\partial z_i} = -\frac{\partial x}{\partial z_i} = \frac{U_{z_i}}{U_x}$. Esta pendiente es cada vez más plana por el efecto de saciedad, el comprador cada vez está dispuesto a pagar menos por tener, por ejemplo, un cuarto de baño más.

Las funciones de disposición a pagar por z (o funciones de Rosen) revelan la cantidad que un hogar está dispuesto a pagar por diferentes niveles de los diferentes atributos z . Por su parte, la función de precio hedónico $p(z)$ muestra el precio mínimo que se debe pagar en el mercado para comprar cada nivel de los atributos z . Por tanto, la restricción relevante para el hogar comprador es simplemente la función de precio hedónico; no podrán ofrecer menos dinero que lo que indica dicha curva. Las diferentes intersecciones entre las curvas de Rosen y la función de precio hedónico indican conjuntos de atributos de vivienda z en los que la disposición del hogar a pagar por una propiedad con ese conjunto de atributos es igual a su precio requerido en el mercado. Al maximizar su utilidad, el hogar elegirá el paquete de atributos de la vivienda z situado sobre aquella curva de disposición a pagar que le proporciona el mayor nivel posible de utilidad, al tiempo que es compatible con los precios vigentes en el mercado. Dicho de otra forma, el hogar maximizará su utilidad yéndose a la curva de Rosen más baja posible (de mayor utilidad posible) que tenga algún punto en común (que sea tangente) con la función de precio hedónico; ya que la disposición a pagar por los diferentes atributos z no puede ser inferior que el precio que pide el mercado.

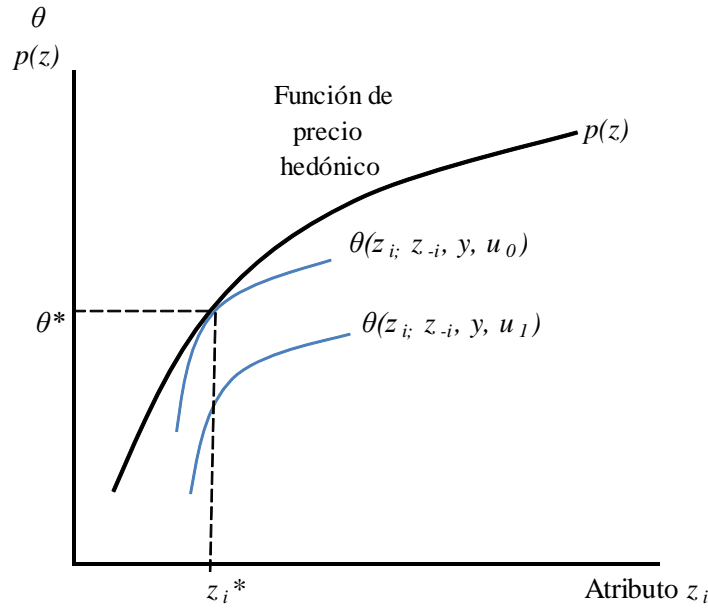


Figura 2.2 Elección de la vivienda óptima utilizando la función de disposición a pagar por z_i .

2.1.2 La decisión del oferente o productor.

Cuando se analiza la conducta del vendedor, resulta necesario definir la función de mínimo coste de ofrecer una vivienda con características z : $C(z; \beta)$, donde β representa un parámetro de desplazamiento de la función de coste debido, por ejemplo, al precio de los recursos empleados para poder ofrecer la vivienda o a la propia función de producción del vendedor. Para entender el papel de β pensemos en un propietario con conocimientos de albañilería, éste va a ser capaz por sí mismo de mejorar la calidad de su propiedad mediante la instalación de rejas en ventanas y roperos empotrados; sus costes de mejorar z pueden ser más bajos que para otro propietario que debe buscar ayuda profesional para lograr la misma mejora.

Por su parte, el beneficio obtenido por la venta del inmueble viene dado por: $\Pi(z; \beta) = p(z) - C(z; \beta)$, donde el ingreso por la venta va a venir dado por precio del mercado $p(z)$ –el vendedor no puede vender por encima del precio que marca la ‘lista’ hedónica de precios mínimos–. Al igual que analizábamos la disposición a pagar del consumidor en un determinado nivel de utilidad, podemos analizar la disposición a cobrar del vendedor de la vivienda si desea mantener un determinado nivel de beneficio Π_0 : $\Phi(z; \beta) = \Pi_0 + C(z; \beta)$, o bien, $\Phi(\Pi_0, z; \beta)$.

La función de disposición a cobrar va a crecer unidad a unidad con el nivel de beneficio deseado, dados los niveles de z y β : $\frac{\partial \Phi}{\partial \Pi} = 1$. Por otro lado, la pendiente de esta función de iso-beneficio respecto a cualquier atributo de la vivienda z_i viene dada por el coste marginal respecto a z_i : $\frac{\partial \Phi}{\partial z_i} = \frac{\partial C}{\partial z_i} = C_{z_i}$. Un supuesto importante del modelo es que el coste adicional soportado por ofrecer una unidad más del atributo z_i es cada vez mayor (coste marginal de z_i creciente); esto es, para el vendedor cada vez va a resultar más caro ofrecer, por ejemplo, una habitación adicional (especialmente, en el corto plazo, donde no se puede modificar el

tamaño total de la casa o cambiar recursos productivos fijos). Por tanto, las funciones de disposición a cobrar de Rosen van a ser crecientes en z_i ; cada vez se va a pedir más dinero por ofrecer una unidad más de z_i , ya que el coste de dicha oferta es creciente y se quiere mantener el beneficio del vendedor.

Con esta estructura de coste-beneficio del vendedor, podemos definir su conducta de la siguiente forma (Figura 2.3): para maximizar su beneficio, el propietario escogerá ofrecer aquel paquete de atributos z que lo sitúa en la curva de oferta o disposición a cobrar de mayor beneficio (más elevada) que es al mismo tiempo compatible con los precios vigentes en el mercado $p(z)$; al igual que sucedía con los hogares, esta conducta implica la tangencia entre la curva de oferta de Rosen (curva iso-beneficio) y la curva de precios hedónicos.

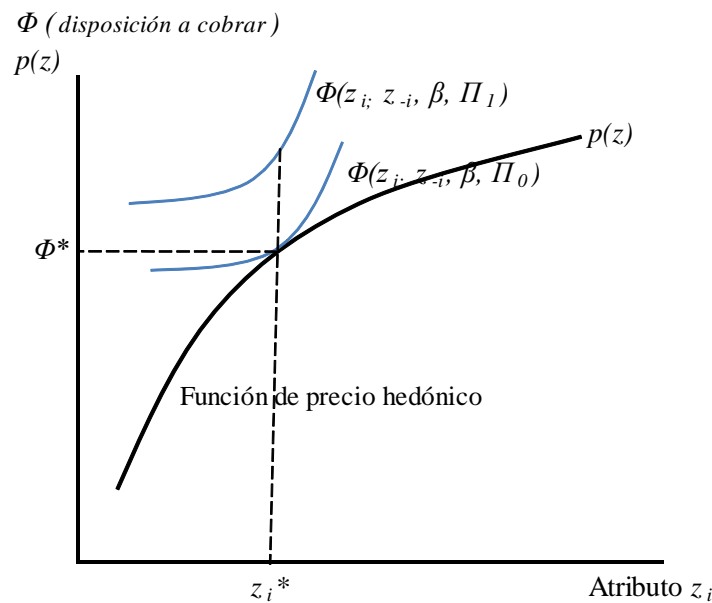


Figura 2.3 Decisión óptima de oferta utilizando la función de disposición a cobrar por z_i .

2.1.3 El equilibrio en el mercado.

Hasta ahora hemos examinado las elecciones de los consumidores (hogares) y de los vendedores (propietarios) en el mercado inmobiliario de forma independiente. La Figura 2.4 presenta ambos lados del mercado en el mismo diagrama. Los hogares definen su ubicación residencial óptima eligiendo una propiedad que cuenta con el conjunto de atributos z que les sitúa sobre la curva de disposición a pagar más baja pero con algún punto de contacto con la función hedónica de precios. El hogar no puede aumentar su utilidad haciendo una oferta por una propiedad con diferentes características a las escogidas. Simultáneamente, los propietarios maximizan sus ganancias al ofrecer una propiedad con el conjunto de atributos z que les sitúa en su curva de oferta más elevada que aún es compatible con la 'lista' de precios hedónicos del mercado. Ofrecer un conjunto alternativo de atributos no va a proporcionar al propietario un mayor nivel de beneficio.

Como se muestra en la Figura 2.4, las curvas de demanda de los hogares y las curvas de oferta de los propietarios se "besarán" a lo largo de la función de precios hedónicos. En cada coincidencia de las curvas de demanda y oferta, un propietario y un hogar se emparejan; el propietario no puede hacer otra cosa que aceptar

la oferta de adquisición del hogar, que a su vez no puede hacer otra cosa que aceptar la oferta de venta del propietario.

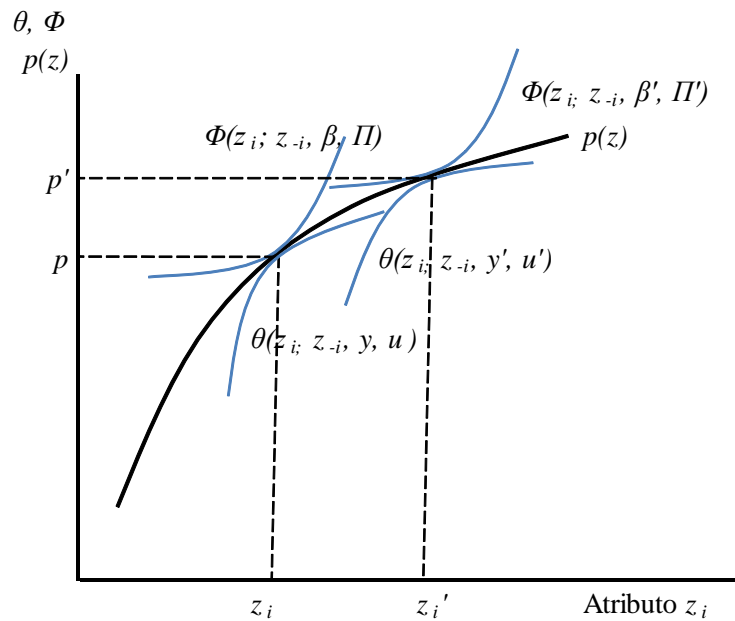


Figura 2.4 El intercambio en el mercado.

Finalmente, obsérvese que el mercado está lleno de diferentes curvas de demanda (o de disposición a pagar) y de diferentes curvas de oferta (o de disposición a cobrar). Esto se debe a que los consumidores difieren en sus niveles de renta, en sus características socio-económicas y en sus niveles deseados de utilidad, mientras que las empresas difieren fundamentalmente en sus estructuras de costes (definidas por β) y en sus aspiraciones de beneficio. Estas diferencias dan lugar a la existencia de diferentes curvas de demanda y oferta en el espacio multidimensional $(z, p(z))$.

- *El equilibrio estable en el mercado:*

Hasta ahora, nuestro análisis se ha desarrollado a nivel de compradores y vendedores de vivienda a nivel individual. Hemos asumido que cada agente individual, al ser solo un actor muy pequeño en todo el mercado, toma la 'lista' de precios hedónicos de equilibrio como dada, lo cual genera intercambios individuales (y optimizadores) en torno a esta curva de precios.

Nuestra exposición del marco teórico sobre los precios hedónicos concluye analizando el proceso de ajuste hacia el equilibrio del mercado en su conjunto, para lo cual debemos analizar la oferta y la demanda de todo el mercado. Llamemos $Q^d(z)$ a la demanda del mercado de viviendas con características z y $Q^s(z)$ a la oferta del mercado para las viviendas con características z . El mercado estará en equilibrio cuando la función de precio hedónico, $p(z)$, sea tal que $Q^d(z) = Q^s(z)$. A los precios del mercado, los vendedores y los compradores actuarán tal y como se ha descrito en los apartados anteriores, donde muestran sendas conductas optimizadoras de beneficio y utilidad respectivamente. Supongamos ahora que las cantidades demandadas y ofertadas para una combinación particular de atributos de vivienda z' no coinciden al precio vigente en el

mercado $p(z')$. En tal caso, esperaríamos que las fuerzas del mercado actuaran a fin de cambiar la 'lista' de precios hedónicos de manera tal que el mercado volviera al equilibrio en el punto espacial z' . Sin embargo, en el mercado de atributos hedónicos, el ajuste no consiste únicamente en un cambio en el precio del paquete z' , ya que dicho cambio va a inducir a que se produzcan simultáneamente cambios a toda la 'lista' de precios hedónicos. Para entender este proceso simultáneo, imaginemos que se produce un exceso de demanda del paquete de atributos z' que tiende a encarecer dicho tipo de vivienda. Los compradores, que en la 'lista' inicial de precios hedónicos habían maximizado su utilidad en dicho paquete de atributos, sustituirán ese paquete en particular por otros más baratos que le reporten mayor utilidad, por lo que aumentará la demanda de otras propiedades con características diferentes.

Por su parte, algunos vendedores optarán por modificar sus propiedades para aprovechar los precios más altos del mercado en z' .

Esto inducirá una menor oferta de propiedades con paquetes alternativos de atributos $z'' \neq z'$. El proceso de sustitución de compra y reasignación de oferta se propagará a través de todo el mercado dando lugar a una nueva curva hedónica de precios final $p'(z)$ que vuelve a igualar oferta y demanda para toda posible combinación de z .

2.2 Revisión de la Literatura.

En la mayoría de los análisis económicos sobre el mercado de la vivienda, la variable precio de la vivienda, es objeto de suma importancia en la investigación. La multitud de características que posee una vivienda, tales como tamaño, durabilidad, localización, etc., hacen complejo valorar el precio de mercado que debe tener el bien. Además, en esta valoración, también influyen el ciclo económico, las políticas macroeconómicas del momento, y las características de la población objeto de estudio.

Como hemos comentado anteriormente, el objeto de nuestro estudio es el de presentar un modelo empírico que permita explicar el precio de la vivienda según sus características en la ciudad de Sevilla en el año 2017. La línea teórica y empírica que sigue nuestro trabajo es la de Rosen (1974), explicada ya en el punto anterior, por lo que nuestra revisión de literatura se ha centrado fundamentalmente en aquellos trabajos que se han desarrollado en esta línea de investigación.

Bilbao-Terol (2000), basándose en el modelo de Rosen (1974), establece un análisis empírico para las poblaciones más importantes de la zona central del Principado de Asturias en el año 1996 (Oviedo, Gijón, Avilés, Mieres y Langreo). Su estudio se basa en una muestra que contiene 364 viviendas vendidas en el año 1996. Las variables independientes que incluye en sus ecuaciones son: m² útiles, nº de baños, disponibilidad de calefacción, antigüedad y distancia al centro de la ciudad. Estima 5 ecuaciones hedónicas (una para cada población) y ensaya con tres formas funcionales diferentes: lineal, semilogarítmica y la logartimo-lineal. En cuatro de los cinco mercados, la especificación lineal es la que funciona mejor, mientras que en uno de los mercados, es la función logaritmo lineal la que ofrece un mejor ajuste. Tras la estimación de los coeficientes por mínimos cuadrados ordinarios (MCO), se obtiene que el aumento del precio de la vivienda está directamente relacionado con los metros cuadrados y con el número de baños. En segundo lugar, estarían las variables antigüedad y calefacción.

Un estudio interesante por su metodología es el que realizan Caridad, Brañas y de la Paz (1997). Estos autores analizan el mercado de vivienda en zonas de escasa renta en Córdoba capital, haciendo una distinción de dos grupos de barrios: los de renta baja³ y los de renta media-baja⁴. El trabajo acepta por tanto la existencia de mercados segmentados; segmentación en el sentido de Kain *et al.* (1975). Utilizando la metodología econométrica de los componentes principales⁵, se agrupa una serie de variables numéricas e índices, siendo los siguientes: precio de oferta de la vivienda; superficie de la vivienda en metros cuadrados útiles; número de dormitorios; número de cuartos de baño; número de armarios empotrados; índice de instalaciones; índice de conservación; índice de localización; índice de mejoras; índice de confort; índice de parking; índice de época.

³ La zona 1 está compuesta por los siguientes barrios de la ciudad de Córdoba: Parque Azahara, Miralbaida, Sector Sur, Plaza de Andalucía y el Cerro.

⁴ La zona 2 está integrada por los siguientes barrios: Parque Cruz Conde, Carlos III, Fátima, Levante, Fuensanta, Fidiana, Cañero, Santuario, Figueroa, Vuelta de la Reina y el Naranjo.

⁵ Técnica utilizada para describir un conjunto de datos en términos de nuevas variables ("componentes"), no correlacionadas entre sí, que surgen de combinaciones lineales de las numerosas variables explicativas de partida.

Tras el análisis de cada una de los atributos, estos autores demuestran que en las zonas de renta baja, las variables más determinantes del precio para los consumidores, ordenadas según su importancia son: superficie, mejoras, confort, ubicación, conservación y edad. Como se observa el factor que prima para este tipo de zonas es la superficie de la vivienda, en contraposición a la antigüedad, que toma el factor menos relevante. Por otro lado, en las zonas de renta media-baja, también aparece la superficie del inmueble como un factor determinante; siendo el factor disponibilidad de parking negativo, lo que indica que se valora positivamente que existan problemas de aparcamiento (o que disminuye si no existen problemas de parking). Lo que se traduce en que los habitantes de esta zona prefieren en vivir en núcleos más populosos a pesar de que tengan problemas de aparcamiento, aunque en estas zonas se ha comprobado que no hay gran número de coches (no más de uno por familia). Otra conclusión de este estudio es que se observa el fenómeno del “filtering” (las viviendas más nuevas de la ciudad tienen un atractivo mayor, por lo que los más ricos las ocupan, mientras que los menos favorecidos van tomando las que éstas no quieren) en la zona dos.

Chica, *et al.* (2007) presentan un modelo para la interpretación de la variabilidad espacio-temporal del precio de la vivienda en la ciudad de Granada. Según estos autores, las características que influyen en el precio de la vivienda se clasifican en: *microlocalizativas*, *macrolocalizativas* y *generales*. Cuando hablamos de características *microlocalizativas* nos referimos a los factores físicos (forma, tamaño del solar, naturaleza geológica del terreno, atractivo de la localización...), la accesibilidad local (proximidad a escuelas, comercios...) y las variables socioeconómicas del barrio (status social, criminalidad, raza...). Estas características explican la diferencia entre el precio de una vivienda y el de su barrio. Las características *macrolocalizativas*, como la accesibilidad al centro, explican las diferencias entre los precios de un barrio y la media de la ciudad. Las características generales como la coyuntura económica, habitualmente influyen en los precios de toda la ciudad y explican las variaciones a lo largo del tiempo de los precios.

La metodología que se aplica en este caso posee dos vertientes; por un lado, instrumentos geoestadísticos con el variograma como herramienta básica, que estudia la variabilidad espacio-temporal del precio de la vivienda, y por otro lado, el modelo de regresión hedónica que explica las características constructivas, localizativas y temporales.

Como conclusión en este estudio se obtiene del análisis variográfico que el precio de la vivienda es una variable muy correlacionada espacialmente y del modelo de regresión hedónica, se observa que con sólo tres variables constructivas (antigüedad, baños y amplitud) se explica un 84,18% de las variaciones en los precios de las viviendas.

Un enfoque metodológico totalmente diferente a lo que hemos mencionado anteriormente, lo define Fernández, *et al.* (2012) con las RNA (Redes neuronales) y el mercado inmobiliario.

Las redes neuronales están compuestas por unidades simples llamadas neuronas que se pueden organizar en diferentes arquitecturas, llamadas topologías. Estas neuronas son capaces de suministrar respuestas rápidas a cualquier problema mediante un proceso de formación estableciendo correlaciones matemáticas entre los datos de entrada y salida.

En este caso, se han empleado redes neuronales artificiales para aproximar el precio por metro cuadrado de una vivienda, utilizando parametros de entrada tales como: las características internas de la vivienda, el equipamiento del barrio, lejanía de los parques, la distancia al CBD, la proximidad al tranvía o al metro, el nivel socioeconómico y el grado de los inmigrantes en el vecindario.

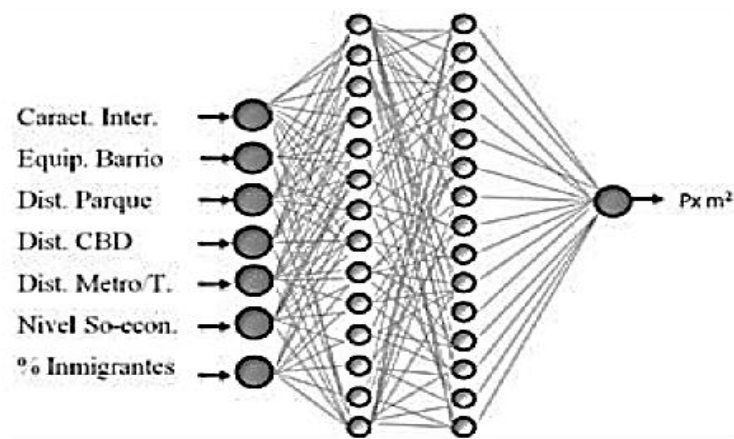


Figura 2.5 RNA topología seleccionada

Fuente: Fernández, L., Llorca, A., Valero, S., Botti, V.J. 2012. *Incidencia de la localización en el precio de la vivienda a través de un modelo de red neuronal artificial. Una aplicación a la ciudad de Valencia. Catastro.*

Como resultado, se obtiene que:

- El precio de la vivienda depende en su mayoría de la cercanía al centro.
- La inmigración interfiere con una relación negativa.
- El nivel socioeconómico tiene una relación directa
- La distancia al metro o tranvía no tiene unos resultados concluyentes.
- Cuanto mayor sea la distancia a espacios verdes con respecto a la vivienda, menor será el precio de esta.

Por otro lado, García (2008) realiza un estudio para determinar el precio de la vivienda en Málaga de 2 años de antigüedad, es decir, de segunda mano. Como la mayoría de los autores, se basa en en la metodología de precios hedónicos. Para ello, cuenta con una variable dependiente, que en este caso sería el precio de la vivienda, variables independientes que eviten multicolinealidad, y una función hedónica que, en su caso, se decanta por la forma funcional semilogarítmica.

Como conclusiones en este caso se obtiene que las características más determinantes en el precio de la vivienda de segunda mano son: superficie construida, nº de baños, garaje privado, luminosidad, localización en lo que se refiere a la cercanía al mar y la distancia al centro.

Añadimos a Núñez, *et al.* (2007) que determinan el precio de la vivienda libre y multifamiliar de un barrio no periférico utilizando de nuevo la metodología de los precios hedónicos. Referida a la muestra, establecen una serie de variables que clasifican en Internas, Externas y Ficticias. Como variables internas nos encontramos:

- Económicas: Precio del mercado
- Orientación: exterior
- Generales: Calidad (solería, cerramientos, amueblado, cocina...) y reforma.
- Básicas: superficie construida, dormitorios, baños, armarios empotrados, aseos, garaje, trastero, climatización.
- Como variables externas nos encontramos:
- Generales: año edificación, ascensor, tendedero
- Extras: Piscina, tenis, jardines.
- Localización.

Igualmente, estos autores utilizan la función hedónica funcional para estimar los coeficientes por Mínimos cuadrados ordinarios. Una vez hecho la estimación de estos coeficientes, podemos concluir que como características determinantes del precio de la vivienda destacan: ubicación, superficie y garaje. Como aspecto a destacar, mencionar que la variable antigüedad que en otras ocasiones ha sido importante, en este caso influye negativamente en el precio.

Taltavull *et al.* (2009) miden las expectativas de precios de viviendas en el mercado de siete ciudades (Alicante, Valencia, Murcia, Madrid, Barcelona, Castellón y Baleares) comprendidos entre los años 1995-2006 utilizando también el modelo hedónico para separar la parte de heterogeneidad de la vivienda y la tendencia general de los precios de la misma. Para ello utiliza un total de 210 variables divididas en tres subgrupos:

- Primer subgrupo: características comunes a la propiedad.
- Segundo subgrupo: cuestiones físicas, estatuto jurídico y vecindario, ambientales.
- Tercer subgrupo: conjunto de valores para estimar la valoración final.

Como resultado, existe un crecimiento considerable de precios entre los años citados (1995-2006) además de un aumento de la densidad de población debido a la inmigración, por lo que afecta a la capacidad de percibir los valores de la propiedad. Los coeficientes que dan como resultado positivo son: tipo casa, calidad de la construcción, área comercial, ingresos del barrio, densidad de la población, orientación y vista, ascensores, superficie de la casa, jardines, renovación, accesibilidad y transporte público. Por otro lado, los coeficientes que dan como resultado negativo son: antigüedad y densidad de la construcción. Todas estas características fundamentales son apreciadas por los propietarios de una manera diferente, dependiendo del momento y las condiciones del mercado.

Destacar a Nicodemo *et al.* (2012) que realizan estimaciones cuantitativas de los precios de las viviendas comprendidas entre 2004 y 2007 en varias ciudades como Madrid, Barcelona, Valencia, Sevilla, Zaragoza, Málaga y Bilbao. A diferencia de la mayoría de las anteriores, utilizan el modelo de precios hedónicos con una peculiaridad, la descomposición de cuantiles usando el método de Melly.

Nicodemo y Raya a través de este método, llegan a la conclusión de que conforme aumente un año en la vivienda, se disminuye un 0.04% del precio de la misma. La existencia de Ascensor por el contrario,

aumentaría un 14.93% €/m². Además, la variable “nº de dormitorios” en un piso superior con cocina separada, contribuiría positivamente cuando hablamos de percentiles inferiores y negativamente cuando contemos con percentiles superiores. Añadir además que la distribución de determinantes de cambios en la distribución de precios entre los años 2004-2007 se hace más notoria en ciudades pequeñas tales como Valencia y Bilbao, a diferencia de ciudades como Zaragoza y Sevilla que apenas se ven afectadas.

A continuación, se muestra una tabla resumen (Tabla 2.1) de todos estos artículos anteriormente mencionados para reflejar cada una de las metodologías empleadas y atributos que más se utilizan en el análisis aplicado a la determinación del precio de la vivienda. Como vemos, la mayoría de los autores de nuestra revisión centra su investigación en el modelo de precios hedónicos de Rosen (1974). En tal sentido, la Teoría de Precios Hedónicos implica un gran avance de la metodología en la configuración de mercados conformados por atributos, facilitando métodos econométricos para alcanzar precios y demandas contenidas a partir de la medición del precio del bien compuesto.

Tabla 2.1 Resumen de los artículos revisados.

AUTOR	AÑO	METODOLOGÍA	VARIABLES	CONCLUSIONES
Bilbao Terol	1996	Modelo precios hedónicos (Rosen 1974)	M ² útiles y construidos. Nº baños. Calefacción. Altura. Garaje. Obra nueva.	Permite estimar precio unidad vivienda en función de sus características hedónicas que posea. Utilidad en la resolución de problemas económicos.
Caridad, Brañas y de la Paz	1997	Modelo precios hedónicos (Rosen 1974)	PVP agencia. M ² útiles. Nº dormitorios Índice de instalaciones, localización, mejoras, confort, comodidad, parking... Época.	Población de menor renta le importa más la superficie. Población de mayor renta compran viviendas viejas para reformarlas.
Chica, Cano	2007	Modelo precios hedónicos (Rosen 1974)	Microlocalizativas, Macrolocalizativas. Observables y no observables	3 variables constructivas significativas (Antigüedad, baños y amplitud)
Núñez, Ceular y Millán	2007	Modelo precios hedónicos (Rosen 1974)	Internas: económicas, orientación, básicas y generales. Externas: localización, generales y extras. Ficticias.	Características determinantes: ubicación, superficie y garaje. La variable antigüedad influye negativamente en el precio.
García	2008	Modelo precios hedónicos (Rosen 1974)	Variables de intervención "d1" y "d2" como valores extremos.	Características determinantes: superficie construida, nº baños, garaje privado, luminosidad, localización, distancia al centro.
Taltavull y Mcgreal	2009	Modelo precios hedónicos (Rosen 1974)	Primer subgrupo: características comunes a la propiedad. Segundo subgrupo: cuestiones físicas y ambientales. Tercer subgrupo: valores para estimar valoración final.	Coefficientes positivos: calidad, área comercial, barrio, orientación, ascensores, superficie, jardines, accesibilidad. Coefficientes negativos: antigüedad, densidad de la construcción.
Fernández, Llorca, Valero y Botti	2012	Redes neuronales y mercado inmobiliario.	7 variables de entrada 1 variable de salida	Inmigración: relación negativa Nivel socioeconómico: relación directa. Mayor distancia a espacios verdes, menor precio de la vivienda.
Nicodemo y Raya	2012	Modelo precios hedónicos (Rosen 1974). Descomposición de cuantiles: Método de Melly.	Variables relevantes: Ascensor, nº de dormitorios	Más años en la vivienda, se disminuye un 0.04% del precio de la misma. La existencia de Ascensor, aumentaría un 14.93% €/m2. "Nº de dormitorios" en un piso superior con cocina separada, contribuiría positivamente cuando hablamos de percentiles inferiores y negativamente cuando contemos con percentiles superiores.

Fuente: elaboración propia.

3 LA METODOLOGÍA ECONOMETRICA. REGRESIÓN LINEAL MÚLTIPLE.

3.1 Conceptos básicos.

Entendemos por Econometría, la disciplina que integra la teoría económica, las matemáticas y las técnicas estadísticas, con el objeto de contrastar hipótesis sobre fenómenos económicos, aportar estimaciones numéricas de los coeficientes de las relaciones económicas y prever o predecir los valores futuros de las variables o fenómenos económicos.

La econometría toma como referencia un conjunto de teorías o hipótesis económicas que hay que contrastar. Si el trabajo econométrico (empírico) no respalda una determinada teoría tomada con referencia, el investigador puede experimentar con otras teorías alternativas o proponer la suya propia a la luz de la evidencia (los datos) observada. De esta forma, la investigación econométrica da lugar a la aceptación, el rechazo o la reformulación de las teorías económicas.

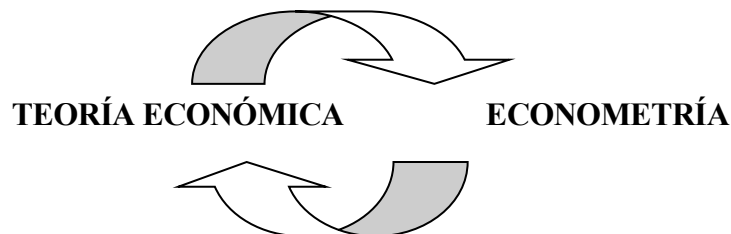


Figura 3.1 Relación teoría económica y econometría. Fuente: elaboración propia.

La investigación econométrica se desarrolla por lo general en tres etapas:

- 1ª. Especificación del modelo (o de las hipótesis realizadas) en forma de una ecuación explícita.
- 2ª. Recopilación de los datos sobre las variables del modelo y la estimación de los coeficientes de la función con las técnicas econométricas adecuadas.
- 3ª. Evaluación de los coeficientes estimados de la función utilizando criterios económicos, estadísticos y econométricos. Los coeficientes han de ser insesgados, eficientes y consistentes. Términos que se definen de la siguiente manera:

--- Un estimador $\hat{\theta}$ de un parámetro θ es **insesgado** si su valor esperado es θ :

$$E(\hat{\theta}) = \theta, \text{ o bien, } E(\hat{\theta} - \theta) = E(\text{error muestral}) = \text{Sesgo}(\hat{\theta}|\theta) = 0$$

Obsérvese que al comparar θ con $\hat{\theta}$, estaremos comparando población con muestra, algo muy habitual en econometría. $\hat{\theta}$ puede ser, por ejemplo, el estimador de la variable dependiente $\{E(y - \hat{y}) = E(\hat{u}) = 0\}$, el estimador de los coeficientes del modelo ($\hat{\beta}$), el estimador de la media de una determinada característica de una población,...

--- Un estimador insesgado $\hat{\theta}_1$, es más **eficiente** que otro estimador insesgado $\hat{\theta}_2$ si: $\text{Var}(\hat{\theta}_1) < \text{Var}(\hat{\theta}_2)$

--- Un estimador $\hat{\theta}$ de un parámetro θ es **consistente** si: $\lim_{n \rightarrow \infty} \hat{\theta} = \theta$ (n es el tamaño de la muestra).

Todo lo comentado anteriormente, se puede ver reflejado de una manera más gráfica con el llamado *Símil del tiro a la diana*: Cada tiro se corresponde con una estimación del parámetro extraída de una determinada MAS (muestra aleatoria simple).

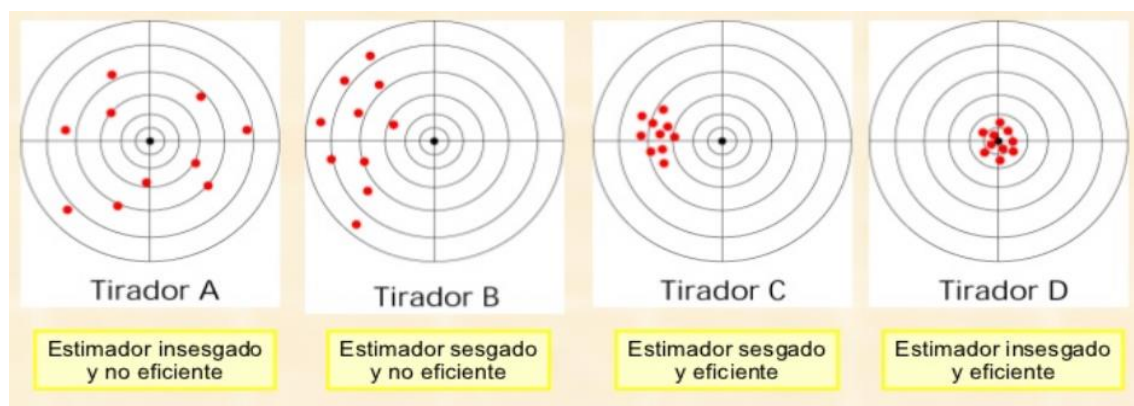


Figura 3.2 Propiedad de los estimadores. Fuente: Wikipedia.

Para realizar un estudio econométrico es necesario tener como base una serie de datos económicos que pueden ser de diferentes tipos:

- Datos **series de tiempo**: analizamos información de una o varias variables a lo largo del tiempo. Un ejemplo claro de este tipo de datos sería el precio de las acciones.
- Datos de **panel o longitudinales**: analizamos información de las mismas variables durante un intervalo de tiempo. Un ejemplo claro de este tipo de datos serían los gastos en el año 1990, los impuestos en el año 1985, datos de la inmigración en el año 1980 etc...
- Datos de **corte transversal o sección cruzada**: analizamos información de varias unidades muestrales (UM) en la misma unidad de tiempo, por lo que estaríamos hablando de un “Modelo clásico de regresión lineal” y que por tanto, estaría directamente relacionado con el proyecto en cuestión. Ejemplo:

$$\text{Precio vivienda} = \beta_0 + \beta_1 m^2 \text{const} + \beta_2 \text{Zona} + \beta_3 n^o \text{habitaciones} + \varepsilon_i$$

(Muestreo del Precio de la vivienda en Sevilla en el año 2017)

3.2 Modelo de Regresión lineal.

3.2.1 Modelo de regresión lineal simple.

Antes de adentrarnos en el estudio de la regresión múltiple, se debe introducir la *regresión simple*, aunque no es muy empleada en la econometría, sino que se usa de manera ocasional y como punto de partida al estudio en cuestión. El modelo de regresión lineal simple explica una variable en función de otra, es decir "y" en función de "x" y se denota a partir de una ecuación muy sencilla:

$$y = \beta_0 + \beta_1 \cdot x + u. \quad (3.1)$$

- Cuando las variables "y" y "x" se relacionan entre sí a través de la ecuación (3.1) se le asignan diferentes nombres según el contexto, aunque los términos “variable explicada” (variable dependiente y) y “variable explicativa” (variable independiente x) serán los más utilizados en econometría. El resto de los términos se resume en la tabla 3.1.
- La variable "u" es el error o perturbación y se identifica como todos los términos que no dependen de x pero que si afectan a y. En la regresión simple, todos estos términos se tratan como términos no observados, por lo que se estima que $u = 0$. Considerando los términos constantes y sabiendo que $u = 0$, se puede sacar la conclusión de que x tiene un efecto lineal sobre y:

$$\Delta y = \beta_1 \Delta x \text{ si } \Delta u = 0. \quad (3.2)$$

- Por tanto se puede deducir que β_1 es el parámetro de la pendiente y β_0 el parámetro del intercepto, el término constante.

Tabla 3.1 Terminología en la regresión simple.

y	x
Variable dependiente	Variable independiente
Variable explicada	Variable explicativa
Variable de respuesta	Variable de control
Variable predicha	Variable predictora
Regresando	Regresor

Fuente: Wooldridge (2010): “Introducción a la econometría. Un enfoque moderno”.

3.2.2 Modelo de regresión múltiple. Estimación.

Una vez introducido el modelo de regresión simple, avanzamos con el concepto de *modelo "poblacional" de regresión múltiple*, que tiene como objetivo analizar y explicar el comportamiento de una variable endógena, endogenous, explicada, dependiente o regresando, que designaremos como "y", utilizando la información proporcionada por los valores tomados por un conjunto de variables explicativas exógenas, explicativas, explanatory, independientes, regresores, que designaremos por " x_1 ", " x_2 " ..., " x_k " a diferencia del modelo anterior, que tenía una única variable explicativa. El modelo lineal (modelo econométrico) viene dado de la siguiente forma:

$$y = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_k \cdot x_k + u. \quad (3.3)$$

Los coeficientes (parámetros) $\beta_1, \beta_2 \dots \beta_k$ denotan la magnitud del efecto que las variables explicativas tienen sobre la variable explicada (*ceteris paribus*⁶). El coeficiente β_0 se denomina término constante o independiente (*intercepto*⁷) del modelo. El término u se denomina perturbación o término de error del modelo. "u" recoge la heterogeneidad inobservable del modelo (no relacionada con x_1, x_2, \dots, x_k); hechos o condiciones idiosincráticas de cada UM⁸, no observados directamente, pero que afectan al valor de y de cada UM. Por tanto, no se observa "u", pero se presupone que sigue una determinada distribución de probabilidad, normalmente una Normal $(0, \sigma^2)$.

Se formula el modelo lineal bajo las siguientes hipótesis clásicas (o **5 supuestos de Gauss-Markov**):

1. Las variables x_1, x_2, \dots, x_k son deterministas (no son variables aleatorias) y además no están correlacionadas con el término error "u", es decir,

$$E(u|x_1, x_2, \dots, x_k) = 0 \quad (3.4)$$

A esta última condición se le llama *hipótesis de exogeneidad*: las observaciones x_1, x_2, \dots, x_k no conllevan información sobre el valor esperado del error o perturbación, al contrario de la endogeneidad. Cuando se produce una perturbación en el modelo, ésta debe afectar a y pero no a las variables explicativas x_1, x_2, \dots, x_k .

Ejemplo de la ecuación anterior sobre el precio de la vivienda (objeto de nuestro proyecto): no por considerar diferentes tipos de vivienda el valor esperado del error va a cambiar.

⁶ Efecto que causa una variable sobre otra, si y solo si, todos los demás factores permanecen constantes. Un ejemplo en nuestro proyecto sería: Efecto causal de los m² de la vivienda sobre el precio de ésta si todos los demás factores permanecen constantes.

⁷ Valor esperado de Y si todas las X valen 0.

⁸ Unidad muestral.

- La variable y es aleatoria, ya que depende de la variable aleatoria u . Además, la relación entre y y $\{x_1, x_2, \dots, x_k\}$ es efectivamente lineal (**hipótesis de linealidad**). Frente a no linealidad. Ver ecuación 3.3.
- El término de error (u) es una variable aleatoria con esperanza nula $E(u_i) = 0$ y matriz de varianzas-covarianzas constante y diagonal; es decir, para todo i , la variable u_i tiene varianza σ^2 constante (no dependiente de i), y además $Cov(u_j, u_k) = 0$ para todo j y todo k distintos entre sí (una perturbación que afecta a un individuo no debe tener efecto en otros individuos).

El hecho de que la varianza de u_i sea constante para todo i (que no dependa de i), se denomina **hipótesis de homocedasticidad** (frente a heteroscedasticidad):

$$Var(u|x_1, x_2, \dots, x_k) = Var(u|y_1, y_2, \dots, y_k) = \sigma^2 \text{ (valor constante) } \quad (3.5)$$

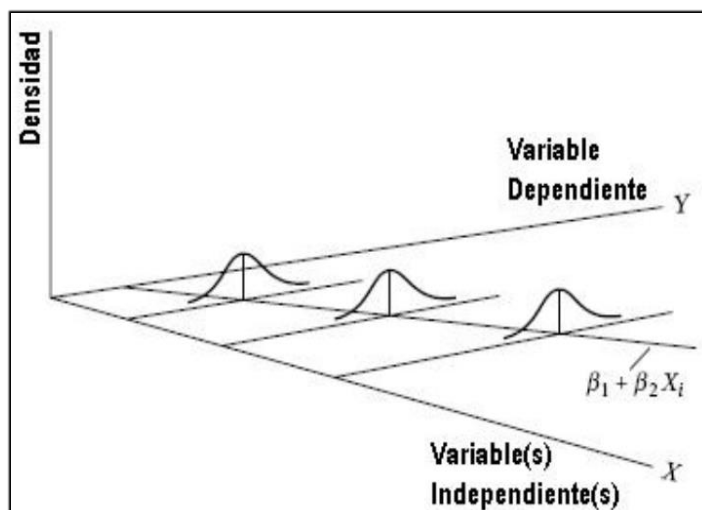


Figura 3.3 Homocedasticidad.

Fuente: Gujarati (2006): “Principios de la econometría”

El hecho de que $Cov(u_j, u_k) = 0$ para todo j y k distintos entre sí se denomina **hipótesis de no autocorrelación**, es decir no existe relación de linealidad entre las dos variables estudiadas (frente a autocorrelación).

$$E[\varepsilon\varepsilon' | \mathbf{X}] = \begin{bmatrix} E[\varepsilon_1\varepsilon_1 | \mathbf{X}] & E[\varepsilon_1\varepsilon_2 | \mathbf{X}] & \cdots & E[\varepsilon_1\varepsilon_n | \mathbf{X}] \\ E[\varepsilon_2\varepsilon_1 | \mathbf{X}] & E[\varepsilon_2\varepsilon_2 | \mathbf{X}] & \cdots & E[\varepsilon_2\varepsilon_n | \mathbf{X}] \\ \vdots & \vdots & \ddots & \vdots \\ E[\varepsilon_n\varepsilon_1 | \mathbf{X}] & E[\varepsilon_n\varepsilon_2 | \mathbf{X}] & \cdots & E[\varepsilon_n\varepsilon_n | \mathbf{X}] \end{bmatrix}$$

$$= \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix},$$

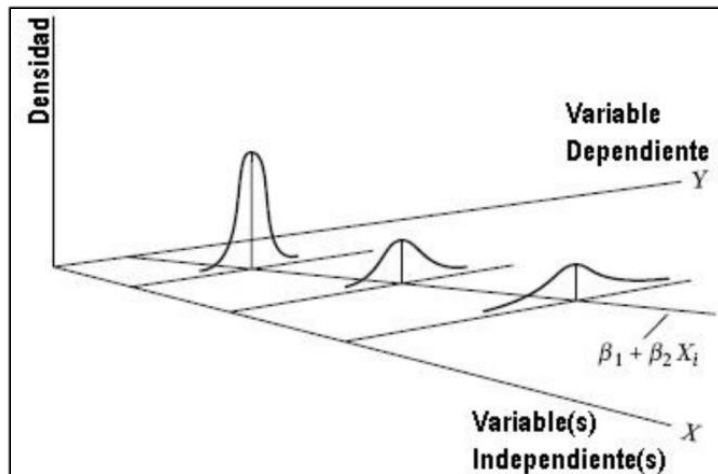


Figura 3.4 Heterocedasticidad.

Fuente: Gujarati (2006): "Principios de la econometría"

4. También se supone la ausencia de errores de especificación, es decir, que suponemos que todas las variables x_1, x_2, \dots, x_k que son relevantes para la explicación de la variable y , están bien medidas e incluidas en la definición del modelo lineal (no hay variables irrelevantes incluidas ni variables relevantes omitidas). Si hubiera variables relevantes omitidas se 'esconderían' en el error. Si incluyéramos variables irrelevantes en la ecuación aumentaría la varianza (Debido a la multicolinealidad).
5. Las variables x_1, x_2, \dots, x_k son linealmente independientes, es decir, no existe relación lineal exacta entre ellas. Esta hipótesis se denomina *hipótesis de independencia o de ausencia de multicolinealidad* (frente a multicolinealidad).
6. Incorporamos un supuesto adicional a los de Gauss-Markov: se considera la hipótesis de normalidad de los errores y , por tanto, de la variable endógena y ; es decir, los errores ϵ_i sean normales para todo i . $u_i \text{ iid} \sim N(0, \sigma^2)$. Frente a no normalidad del término de error.

Tabla 3.2 Interpretación de los coeficientes en el modelo “lineal”.

<i>Modelo Lineal</i>	Variable dependiente	Variable independiente	Interpretación del coeficiente	Relación entre las variables en nivel (Y, X)
$Y = \beta_0 + \beta_1 X + \dots + u$ (Nivel-Nivel)	Y	X	$\beta = \frac{\partial Y}{\partial X}$ o $\frac{\Delta Y}{\Delta X}$	Lineal
$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \dots + u$ (Nivel-Cuadrado)	Y	X, X ²	$\frac{\partial Y}{\partial X} = \beta_1 + 2\beta_2 X$	No Lineal

Fuente: elaboración propia.

3.2.3 Modelo de regresión múltiple. Inferencia.

La inferencia estadística es el conjunto de métodos y técnicas (fundamentalmente, estimación y predicción, intervalos de confianza y contrastes de hipótesis) que permiten inducir, a partir de la información empírica proporcionada por una muestra, cual es el comportamiento de una determinada población con un riesgo de error medible en términos de probabilidad. Por tanto, a partir de aquí, se empieza a trabajar con la muestra.

- *Estimación muestral del modelo por mínimos cuadrados ordinarios MCO:*

Gracias a nuestra MAS⁹, disponemos de un conjunto de n observaciones para cada una de las variables endógenas y exógenas del modelo poblacional:

$$y_i = \beta_0 + \beta_1 \cdot x_{1i} + \beta_2 \cdot x_{2i} + \dots + \beta_k \cdot x_{ki} + u_i, \quad i = 1, 2, 3, \dots, n \quad (3.6)$$

Vamos a estimar los parámetros (desconocidos) $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ basándonos en la información muestral. Una vez calculadas las estimaciones de los parámetros del modelo, podremos analizar el signo y la magnitud de cada parámetro estimado y hacer predicciones acerca del comportamiento (transversal y/o temporal) de la variable y. El modelo en formato matricial:

$$y_{nx1} = x_{nx(k+1)} \beta_{(k+1)x1} + u_{nx1}. \quad (3.7)$$

donde n es el número de observaciones de la muestra y k el número de variables explicativas. Ponemos ‘k + 1’ para tener en cuenta la presencia de un término constante.

Trataremos de obtener el vector $\hat{\beta}$, que representa la estimación (insesgada, eficiente y consistente), a

⁹ Muestra aleatoria simple.

partir de la muestra, de los verdaderos parámetros poblacionales β . Una vez que tenemos $\hat{\beta}$, los correspondientes valores ajustados son $\hat{y} = x\hat{\beta}$, los residuos (estimación muestral de las perturbaciones) son $\hat{u} = y - \hat{y}$, y la suma residual o suma de cuadrados de residuos (SRC) es $\hat{u}'\hat{u}$. El criterio de mínimos cuadrados ordinarios MCO (OLS) considera que la función que mejor se ajusta a los datos es la que minimiza la suma de los cuadrados de los residuos:

$$\min_{\hat{\beta}} SRC(\hat{\beta}) = \hat{u}'\hat{u} \quad (3.8)$$

$$\begin{aligned} SRC(\hat{\beta}) &= (y - x\hat{\beta})'(y - x\hat{\beta}) = y'y - y'x\hat{\beta} - \hat{\beta}'x'y + \hat{\beta}'x'x\hat{\beta} \\ &= y'y - 2\hat{\beta}'x'y + \hat{\beta}'x'x\hat{\beta} \end{aligned} \quad (3.9)$$

Las condiciones de mínimo son:

$$1^\circ \text{ orden: } \frac{\partial SRC(\hat{\beta})}{\partial \hat{\beta}} = -2x'y + 2x'x\hat{\beta} = 0 \rightarrow x'y = x'x\hat{\beta} \rightarrow \hat{\beta} = (x'x)^{-1}x'y \quad (3.10)$$

2º orden:

$$\frac{\partial^2 SRC(\hat{\beta})}{\partial \hat{\beta} \partial \hat{\beta}'} = x'x \quad x'x \text{ es una matriz definida positiva (SRC}(\hat{\beta}) \text{ es convexa)} \quad (3.11)$$

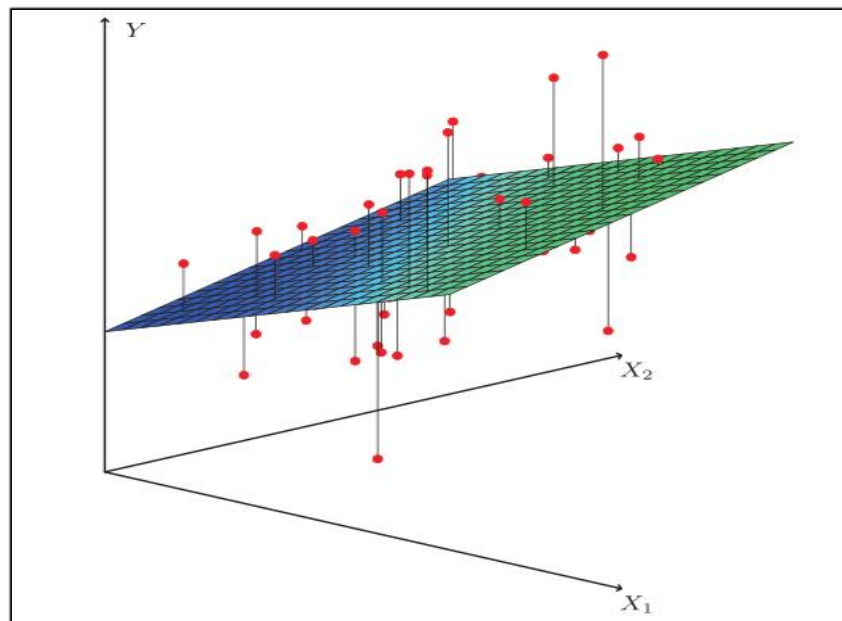


Figura 3.5 Mínimos cuadrados ordinarios.

Fuente: Hastie et al (2008): “The element of statistical learning”

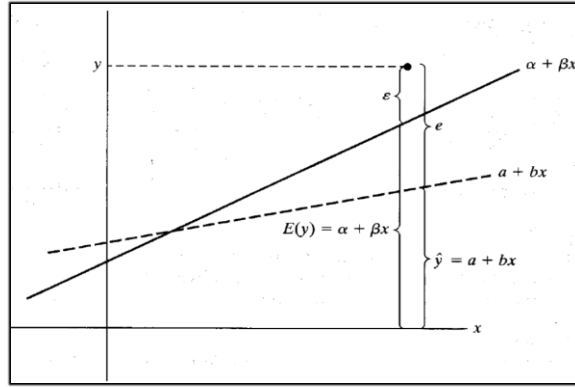


Figura 3.6 Recta poblacional y recta de regresión muestral.

Fuente: Greene (2012).

$$y(\text{población}) = x\beta + u = (\text{muestra})x\hat{\beta} + \hat{u} = \hat{y} + \hat{u} \quad (3.12)$$

Teorema de Gauss-Markov: si se cumplen los 5 supuestos de G-M, entre los estimadores lineales e insesgados de α y β , los más eficientes son los obtenidos por MCO.

3.2.4 Modelo de regresión múltiple. Bondad de ajuste y selección de regresores.

Una vez que se ha establecido los mínimos cuadrados ordinarios, es necesario tener una medida de la bondad de ajuste del modelo a los datos. En el caso de que se haya hecho un estudio de varios modelos, la bondad de ajuste nos permite seleccionar el modelo más apropiado. Por tanto, podemos definir la bondad de ajuste como aquella que resume la discrepancia entre valores observados y valores esperados.

En econometría nos encontramos con diversas medidas de bondad del ajuste. La más conocida es el coeficiente de determinación, que se denota como R^2 o R-cuadrado, y el coeficiente de determinación ajustado, que se denota como $2R$ o R-cuadrado ajustado. Como estas medidas tienen algunas limitaciones, hablaremos también de los criterios de información de AIC y BIC.

- Coeficiente de determinación R^2 :

El coeficiente de determinación parte de la siguiente ecuación:

$$STC = SEC + SRC \quad (3.13)$$

Donde STC es la suma total de cuadrados, SEC la suma explicada de cuadrados y SRC la suma de residuales cuadrados. Partiendo de esta ecuación, podemos definir R^2 de dos maneras distintas:

$$R^2 = \frac{SEC}{STC} \quad (3.14)$$

$$R^2 = 1 - \frac{SRC}{STC} \quad (3.15)$$

R^2 Mide hasta qué punto la SCE de nuestro modelo coincide con la STC obtenida con los valores reales de y . Los valores de este coeficiente en general están comprendidos entre “0” y “1”; Cuando es 0, significa que la varianza explicada es cero, y 1, cuando la varianza residual es cero, por lo que sería el modelo ideal. Un R^2 pequeño significa que la varianza del error o perturbación (σ^2) es grande con respecto a la de y , implicando que β_j no se puede predecir con exactitud, aunque en ocasiones, una varianza de perturbación grande se puede compensar con una muestra grande.

- Coeficiente de determinación R^2 ajustado:

R^2 ajustado por su parte, mide la bondad de ajuste corregida, es decir, identifica lo que se varía en el campo objetivo mediante la entrada o entradas de nuevas variables en el modelo dado. R^2 aumenta tanto como el número de variables que se incluyen en el modelo. El principal papel del R^2 ajustado es corregir esta estimación excesiva. Por el contrario, R^2 ajustado también podría disminuir si una variable no mejora el modelo. Se denota de la siguiente forma:

$$R^2_{ajustado} = 1 - \frac{SRC/n}{STC/n} \quad (3.16)$$

R^2 ajustado siempre es igual o menor que R^2 . Cuando es igual a 1 indica un modelo que predice de manera perfecta los valores del modelo. Sin embargo cuando es igual o menor que 0 significa que el modelo no tiene valor predictivo. Por lo general, este coeficiente suele estar entre estos dos valores.

- **Variables dummies:**

En estadísticas y econometría, particularmente en análisis de regresión, una variable ficticia (también conocida como variable indicadora, variable de diseño, indicador booleano, variable binaria o variable cualitativa) es aquella que toma el valor 0 o 1 a indicar la ausencia o la presencia de algún efecto categórico que se puede esperar que modifique el resultado. Las variables ficticias se usan como dispositivos para clasificar datos en categorías mutuamente excluyentes (Por ejemplo en nuestro caso, que la vivienda tenga ascensor / no ascensor, etc.).

Las variables ficticias son variables "proxy" o suplentes numéricos para hechos cualitativos en un modelo de regresión. En el análisis de regresión, las variables dependientes pueden estar influenciadas no solo por variables cuantitativas (ingreso, producto, precios, etc.), sino también por variables cualitativas (género, religión, región geográfica, etc.). Una variable ficticia independiente (también llamada variable ficticia explicativa) que para alguna observación tiene un valor de 0 hará que el coeficiente de esa variable no tenga ningún papel en influir en la variable dependiente, mientras que cuando tome un valor 1 su coeficiente actúe para alterar el intercepto.

- **Criterio de información de Akaike (AIC):**

El criterio de información de Akaike (AIC) nos proporciona la calidad relativa de un modelo estadístico para un grupo de datos, facilitándonos un modo para la elección del modelo.

AIC se rige por un trade-off, una recompensa entre la bondad de ajuste del modelo y la dificultad del mismo. Además, se fundamenta en la entropía de información, es decir, la incertidumbre de una fuente de información. De primeras, se propone una estimación relativa de la información perdida cuando se usa un modelo concreto para figurar el procedimiento que origina los datos.

AIC no nos facilita información acerca de la calidad de un modelo con total seguridad. Si todos los posibles modelos casan mal, AIC no nos notificará nada. Se denota por:

$$AIC = 2k - 2 \ln(L) \quad (3.17)$$

donde k es el nº de parámetros libres en el modelo estadístico, y L es el máximo valor de la función de veracidad para el modelo considerado.

Por tanto, podemos decir que el modelo a elegir será es el que menor valor tenga de AIC. Este criterio de información no sólo compensa la bondad de ajuste, sino también incorpora una sanción, que es una función

gradual del número de parámetros considerados. Aumentando el número de parámetros libres en el modelo reestablece la bondad del ajuste.

- **Criterio de información Bayesiano:**

El criterio de información bayesiano (BIC) o el más usual criterio de Schwarz (SBC también llamado, SBIC) es un criterio para la elección de modelos entre una agrupación finita de otros. Se sustenta de la función de probabilidad y está fuertemente relacionado con el Criterio de Información de Akaike (AIC). Bajo la hipótesis de normalidad, BIC se denota por:

$$BIC = x^2 + k \cdot \ln(n) \quad (3.18)$$

Donde k es el n° de parámetros libres en el modelo estadístico, x el n° de datos observados y n el tamaño de la muestra.

Dados dos modelos a evaluar, el modelo que tenga el menor valor de BIC es el que finalmente se elige. El BIC es un incremento de la función de σ^2 y una función creciente de k . Es decir, la variación no explicada en la variable dependiente y el n° de variables explicativas incrementan el valor de BIC. Concluyendo, un menor BIC implica un número inferior de variables explicativas, un ajuste mejorado, o los dos.

4 DESCRIPCIÓN DE LA MUESTRA

Para la realización de este trabajo se ha empleado información sobre las características de una muestra de viviendas (pisos) de Sevilla capital anunciadas en el portal web inmobiliario Idealista.com. Tras depurar la base de datos inicial, de 1.184 pisos, eliminando observaciones o viviendas poco informadas, la muestra empleada queda reducida a un total de 935 pisos. Esta muestra está compuesta, a su vez, por dos submuestras¹⁰: la primera submuestra, formada por 462 observaciones, se ha obtenido directamente (de forma manual) de la plataforma, mientras que la segunda submuestra, de 473 observaciones, ha sido proporcionada directamente por el equipo de gestión de datos de Idealista/Datos. Ambas submuestras han sido obtenidas seleccionando previamente determinadas zonas de Sevilla capital, siendo la selección de los pisos en cada una de estas zonas aleatoria. La Tabla 1 muestra la distribución por zonas de nuestros datos:

¹⁰ Agradecimientos: Agradecemos a los alumnos de la asignatura de Eficiencia y Productividad del Máster en Organización Industrial y Gestión de Empresas de la Universidad de Sevilla la ayuda prestada en la obtención de los datos. Agradecemos también la colaboración y el suministro de datos del equipo de Idealista/Datos.

Tabla 4.1 Distribución de la muestra por zonas de Sevilla capital.

Zonas de Sevilla		Frecuencia	Porcentaje
Sur	Felipe II	32	3,42
	Bami	21	2,25
	Bermejales	20	2,14
	Heliópolis	10	1,07
	El Porvenir	7	0,75
Oeste	Triana	27	2,89
	Altozano - Pagés del Corro	20	2,14
	Parque de los Príncipes	18	1,93
	Tablada	17	1,82
	Plaza de Cuba	7	0,75
	Asunción	6	0,64
Norte	Carretera de Carmona - Miraflores	17	1,82
	Macarena	12	1,28
	Manuel del Valle	9	0,96
Este	Nervión1	109	11,66
	Nervión2	106	11,34
	San Pablo	61	6,52
	Nervión3	27	2,89
	Arroyo - Santa Justa	16	1,71
	Santa Clara	2	0,21
Centro	Arenal - Museo	286	30,59
	Centro histórico	71	7,59
	El Prado	11	1,18
	Puerta de Osario	11	1,18
	Alameda	6	0,64
	Puerta de Carmona	6	0,64
Total		935	100

Fuente: Elaboración propia a partir de los datos.

Las zonas con mayor representación son Arenal-Museo (Sevilla centro) y Nervión (situada en el este de Sevilla capital), lo cual se debe a que los 473 datos procedentes de Idealista/Datos abarcan únicamente a estas dos zonas. Asimismo, hemos dividido la zona de Nervión en 3 sub-zonas –Nervión 1, Nervión 2 y Nervión 3– que, a nuestro entender, pueden tener algún efecto diferencial en el precio de la vivienda (véase dicha división en la Figura 4.1).

Figura 4.1 División de Nervión en 3 sub-zonas.



Fuente: Google Maps.

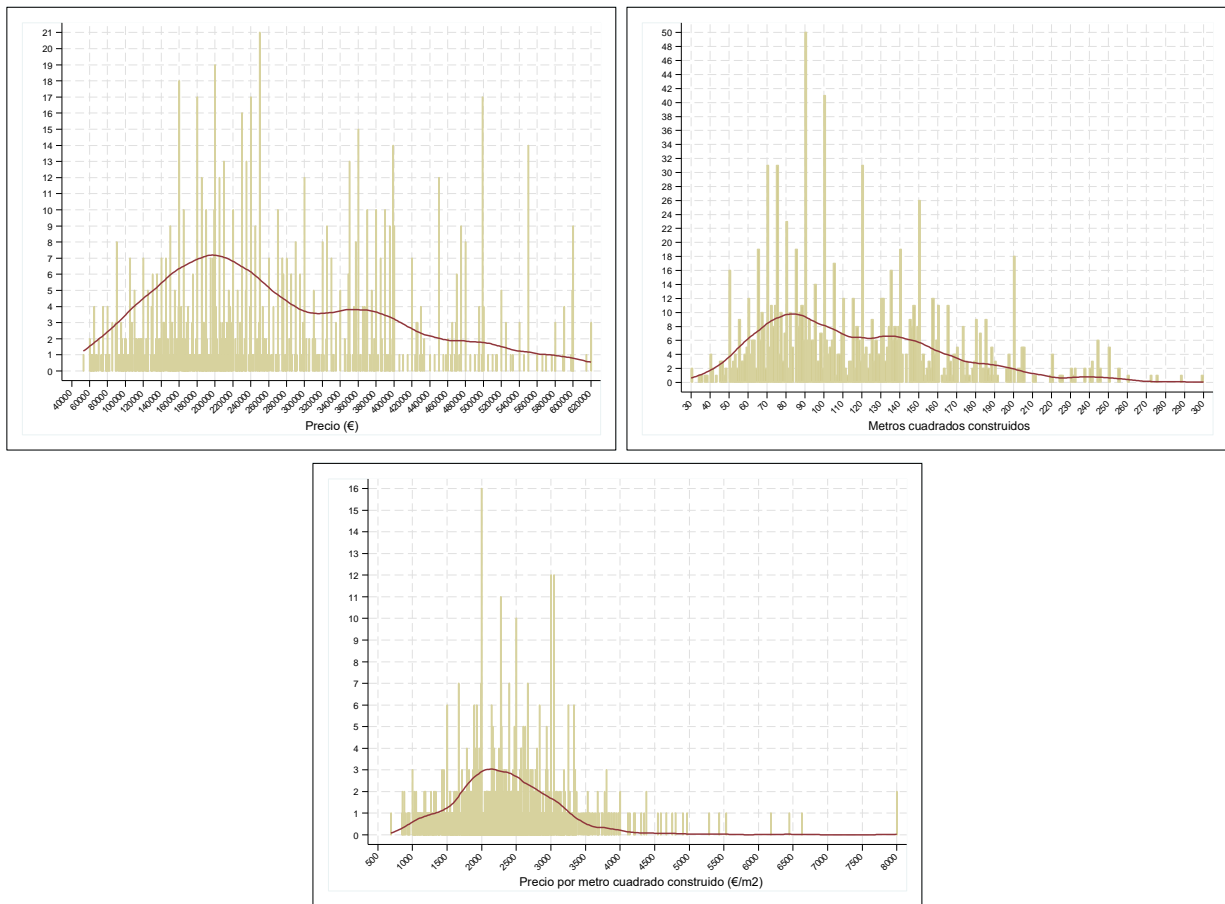
La muestra obtenida contiene la siguiente información relevante para cada inmueble¹¹: el link del anuncio en idealista.com, que hace las veces de identificador de la vivienda; su fecha de alta en idealista.com; la calle, la zona y el código postal de la vivienda; el precio de oferta (precio demandado por el comprador); los metros cuadrados construidos¹²; la planta en que se encuentra el inmueble; si el piso es nuevo o de segunda mano; el estado de conservación (si necesita o no reforma); si se entrega amueblado; el número de habitaciones y de baños/aseos; si pertenece o no a un banco; si lo anuncia en Idealista el propietario, una inmobiliaria o un banco; y si dispone de ascensor, de terraza, de garaje, de trastero, de piscina y de jardín. A partir de toda esta información resulta posible analizar la relación existente entre el precio de oferta de cada vivienda y sus diferentes determinantes.

El principal determinante del precio de un inmueble es su extensión medida en metros cuadrados. Los histogramas de la Figura 4.2 muestran la distribución de las variables precio, metros cuadrados construidos y precio por metro cuadrado construido.

¹¹ Lamentablemente, no disponemos de ciertas características que podrían influir en el precio de la vivienda, como la antigüedad del bloque, su altura en pisos, su orientación o si el piso es de Protección Oficial (VPO). Asimismo, tampoco se conoce el precio al que finalmente se intercambia la vivienda si finalmente se transmite.

¹² Para el 26,7% de los pisos se dispone también del dato de los metros cuadrados útiles.

Figura 4.2 Distribución de las variables precio, m2 construidos y precio por m2 construido.

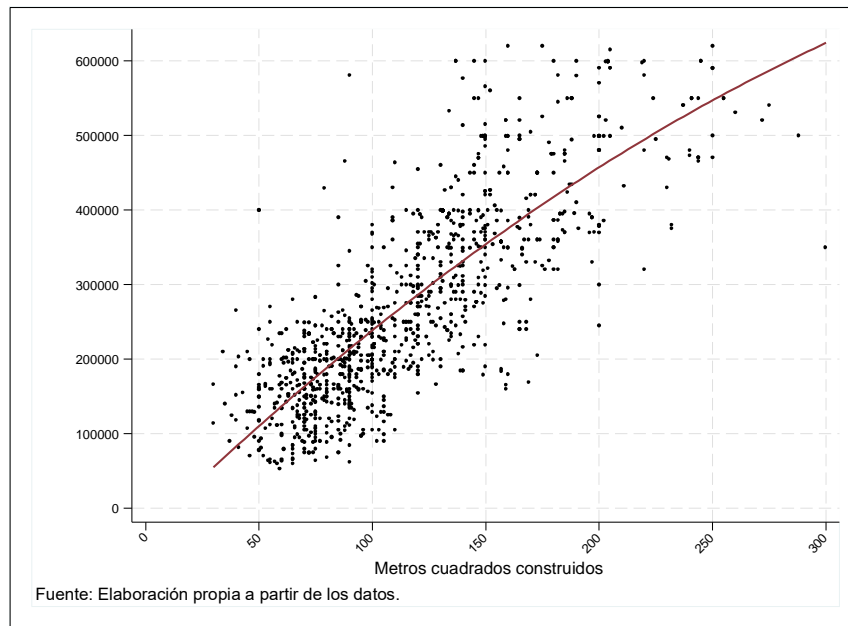


Fuente: Elaboración propia a partir de los datos.

Los precios más comunes en la muestra son 250.000 €, 200.000 €, 160.000 €, 180.000 €, 240.000 €, 499.000 €, 230.000 € y 360.000 €, aunque la mayor parte de la distribución de precios se concentra en el intervalo 150.000-250.000 €. En cuanto a los metros cuadrados de construcción, las cifras más corrientes son 90 m², 100 m², 70 m², 75 m², 120 m², 150 m² y 80 m²; la distribución de m² construidos se concentra en el intervalo 60-100 m², y en menor medida en el intervalo 130-160 m². Finalmente, el cociente de las dos variables anteriores, el precio por metro cuadrado, tiene como valores más frecuentes 2.000 €/m², 3.000 €/m², 3.046 €/m² y 2.500 €/m², estando su distribución concentrada en el intervalo 1.500-3.000 €/m².

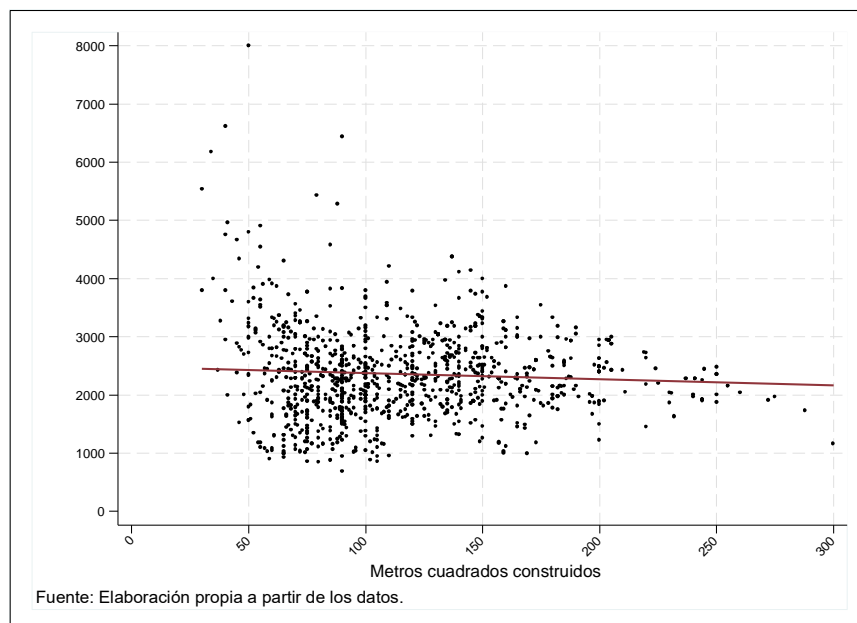
La relación entre el precio y los m² construidos se muestra en la Figura 4.3, la cual muestra claramente una relación positiva y aproximadamente lineal entre ambas variables. La dispersión de precios existente en torno a la curva de ajuste implica que existen pisos con la misma superficie que muestran precios diferentes. En principio, estas diferencias deberían venir explicadas en buena medida por el resto de atributos del piso, pero si esto no ocurre, es decir, si existen dos pisos de similares características con precios muy dispares, el más barato debería ser objeto de estudio por constituir una posible oportunidad de inversión.

Figura 4.3 Relación entre el precio de oferta y el número de m2 construidos.



Por su parte, la Figura 4.4 muestra la relación existente entre el precio por metro cuadrado del inmueble y el número de metros cuadrados construidos.

Figura 4.4 Relación entre el precio por m2 construido y el número de m2 construidos.



La dispersión de la variable 'precios por metro cuadrado construido' es mayor en los pisos pequeños, es decir, aquellos que se mueven en el entorno de 40-60 m². Además, la relación entre estas dos variables muestra cierta pendiente negativa, lo cual indicaría que en los pisos más grandes el precio por metro cuadrado construido es relativamente menor que en los pisos más pequeños.

A continuación, abordamos el análisis de otros atributos del piso que pueden añadir valor al mismo una vez que se ha controlado por la variable de metros cuadrados de construcción. La Tabla 4.2 muestra la

distribución conjunta del número de habitaciones y del número de baños que tiene la vivienda. Las cifras más habituales son pisos de 3 (33,9%) o 4 (24,5%) habitaciones y de uno (39,1%) o dos (50,9%) baños. La combinación más observada de estas dos variables es la formada por aquellas viviendas que tienen 4 habitaciones y 2 baños (19%) y aquellas que disponen de 3 habitaciones y 2 baños (17,9%).

Tabla 4.2 Distribución del número de habitaciones y de baños. Unidades y porcentajes.

		Baños					Total
		1	2	3	4	5	
Habitaciones	1	87	7	0	0	0	94
	2	108	76	2	0	0	186
	3	142	167	8	0	0	317
	4	23	178	26	2	0	229
	5	5	42	33	3	0	83
	6	1	6	9	7	0	23
	7	0	0	1	1	1	3
	Total	366	476	79	13	1	935

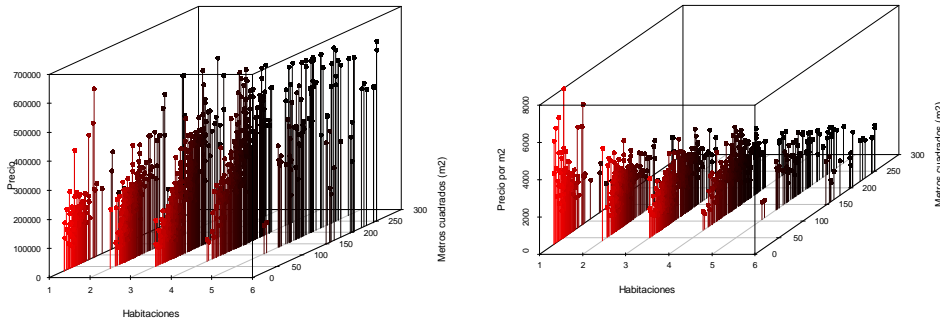
		Baños					Total
		1	2	3	4	5	
Habitaciones	1	9,3%	0,7%	0,0%	0,0%	0,0%	10,1%
	2	11,6%	8,1%	0,2%	0,0%	0,0%	19,9%
	3	15,2%	17,9%	0,9%	0,0%	0,0%	33,9%
	4	2,5%	19,0%	2,8%	0,2%	0,0%	24,5%
	5	0,5%	4,5%	3,5%	0,3%	0,0%	8,9%
	6	0,1%	0,6%	1,0%	0,7%	0,0%	2,5%
	7	0,0%	0,0%	0,1%	0,1%	0,1%	0,3%
	Total	39,1%	50,9%	8,4%	1,4%	0,1%	100,0%

Fuente: Elaboración propia a partir de los datos.

Un análisis del efecto conjunto del número de habitaciones y de los m² construidos sobre el precio de la vivienda y sobre el precio por metro cuadrado se muestra en la Figura 4.5. En el primer gráfico de la figura se observa que a medida que aumenta el número de habitaciones, el precio de la vivienda se sitúa en una escala promedio superior; al menos eso ocurre hasta la quinta habitación. Además, se observa que, para cualquier nivel dado de la variable habitaciones, los incrementos de precio debidos a incrementos de metros cuadrados parecen ocurrir de forma no lineal. El segundo gráfico muestra que los pisos de una o dos habitaciones tienen (en promedio) un precio por metro cuadrado mayor que los pisos con más de dos habitaciones cuando nos movemos en el entorno de los 50 m²; sin embargo, a partir de los 100 m² aproximadamente, el dibujo es al contrario, el precio del metro cuadrado aumenta con el número de habitaciones –al menos hasta las 5 habitaciones–.

En definitiva, cuando los pisos son pequeños se valoran más los que tienen pocas habitaciones, y a la inversa; este hecho parece aconsejar el uso de interacciones entre metros cuadrados y número de habitaciones en el modelo de regresión múltiple.

Figura 4.5 Relación entre el precio, los m2 construidos y el número de habitaciones.



Fuente: Elaboración propia a partir de los datos.

El análisis de la planta en la que se ubica el inmueble se muestra en la Tabla 4.3. Las observaciones más habituales en la muestra son las correspondientes a pisos situados en primera, segunda o tercera planta, que suponen conjuntamente más del 60% de la muestra. Los pisos situados en la planta baja muestran un precio medio y un precio medio por m² inferiores al de los pisos de plantas inmediatamente superiores. Asimismo, el precio medio por m² es relativamente menor en el entorno de las plantas más elevadas (8^a - 11^a). En cualquier caso, este análisis del precio medio por planta no está teniendo en cuenta que la asignación de plantas a zonas en una ciudad no es aleatoria; por ejemplo, en Sevilla, zonas relativamente caras, como el centro histórico, tienen más limitado el número de plantas de sus edificios, por lo que esta zona céntrica añade valor a las primeras plantas pero no a plantas relativamente elevadas. Por tanto, el análisis de la planta del edificio ha de tener en cuenta, entre otras variables, la zona en la que se encuentra ubicado el inmueble.

Tabla 4.3 Distribución de la planta del piso. Valores medios por planta del precio, los m² y el precio por m².

Planta	Frecuencia	Porcentaje	Precio medio	Media m ²	Media precio por m ²
0	102	10,9	194.701	90	2.275
1	244	26,1	277.786	117	2.466
2	192	20,5	264.822	110	2.403
3	164	17,5	303.686	129	2.356
4	73	7,8	257.818	118	2.172
5	32	3,4	299.453	128	2.321
6	40	4,3	276.276	124	2.283
7	52	5,6	323.732	131	2.387
8	14	1,5	263.250	114	2.160
9	10	1,1	214.740	101	2.293
10	9	1,0	295.867	140	2.131
11	3	0,3	310.000	136	2.125

Fuente: Elaboración propia a partir de los datos.

Nuestro análisis descriptivo finaliza con el estudio de seis atributos de la vivienda que pueden tener influencia sobre el precio de la misma (Tabla 4.4). Las variables ascensor y segunda mano se analizan de

forma conjunta, y lo mismo sucede con las variables garaje y trastero, por un lado, y piscina y jardín, por otro. Respecto a la primera pareja de variables, ascensor y segunda mano, observamos que la mayoría de los pisos de la muestra son de segunda mano (96,3%) y tienen ascensor (75,5%); el 72,3% de los pisos presentan ambos atributos, es decir, son de segunda mano y tienen ascensor. En la muestra se observan 5 pisos que a pesar de ser nuevos no disponen de ascensor; posiblemente se trate de edificaciones nuevas levantadas en espacios relativamente reducidos que dificultan la instalación de un ascensor.

En lo que se refiere a las variables garaje y trastero, se observa que solo el 29% y el 27% de los pisos tienen respectivamente garaje o trastero; aquellos pisos que tienen ambos atributos, garaje y trastero, representan solo el 15% de la muestra, mientras que el 59% de los pisos no tienen ni garaje ni trastero.

Tabla 4.4 Distribución de otros atributos de la vivienda. Unidades y porcentajes.

Unidades		Segunda Mano			Porcentajes		Segunda Mano		
		No	Sí	Total			No	Sí	Total
Ascensor	No	5	224	229	Ascensor	No	0,5%	24,0%	24,5%
	Sí	30	676	706		Sí	3,2%	72,3%	75,5%
	Total	35	900	935		Total	3,7%	96,3%	100,0%

Unidades		Trastero			Porcentajes		Trastero		
		No	Sí	Total			No	Sí	Total
Garaje	No	552	112	664	Garaje	No	59,0%	12,0%	71,0%
	Sí	131	140	271		Sí	14,0%	15,0%	29,0%
	Total	683	252	935		Total	73,0%	27,0%	100,0%

Unidades		Jardín			Porcentajes		Jardín		
		No	Sí	Total			No	Sí	Total
Piscina	No	406	36	442	Piscina	No	85,8%	7,6%	93,4%
	Sí	5	26	31		Sí	1,1%	5,5%	6,6%
	Total	411	62	473		Total	86,9%	13,1%	100,0%

Fuente: Elaboración propia a partir de los datos.

La última pareja de variables analizada permite controlar por la existencia de piscina y/o de jardín en el bloque o la comunidad de vecinos del piso. Únicamente el 6,6% de los pisos de la muestra tiene piscina, y solo el 13,1% tiene jardines; el 5,5% de los inmuebles tienen ambos atributos, piscina y jardines, y el 85,8%, la mayoría de pisos, no tienen ni piscina ni jardín.

5 ESTIMACIÓN DE UN MODELO HEDÓNICO DEL PRECIO DE LA VIVIENDA EN SEVILLA CAPITAL EN EL AÑO 2017.

En esta sección se va a llevar a cabo la estimación de un modelo de regresión lineal múltiple explicativo del precio de la vivienda en Sevilla de acuerdo con nuestra muestra. Las variables explicativas que han resultado ser significativas en el modelo, aparte del término constante, han sido las siguientes¹³:

- Metros cuadrados construidos: que se introducen en el modelo en nivel y al cuadrado para capturar la posible relación no estrictamente lineal entre dicha variable y el precio del inmueble.

- Zona de Sevilla en que se encuentra el piso. Se distinguen hasta 26 zonas diferentes controladas por variables ficticias (*dummies*): Alameda, ‘Altozano - Pagés del Corro (Triana)’, ‘Arenal Museo’, ‘Arroyo - Santa Justa’, Asunción (Los Remedios), Bami, Bermejales, ‘Carretera de Carmona - Miraflores’, Centro, El Porvenir, Felipe II, Heliópolis, Macarena, Manuel del Valle, Nervión1, Nervión2, Nervión3 (véase Figura #.1), Parque de los Príncipes (Los Remedios), Plaza de Cuba (Triana), El Prado, Puerta de Carmona, Puerta Osario, San Pablo, Santa Clara, Tablada y resto de Triana¹⁴.

- Número de habitaciones: los pisos de la muestra tienen desde 1 hasta 6 habitaciones. Se emplea una variable *dummy* para cada número de habitaciones.

- Número de baños: los pisos de la muestra tienen de 1 a 4 baños. Se emplea una variable *dummy* para cada número de baños.

- Planta del inmueble: los pisos de la muestra van desde la planta baja hasta la planta undécima. Se emplea una variable *dummy* para cada número de planta.

- Si dispone de ascensor (*dummy*): ‘1’ si tiene ascensor, ‘0’ si no lo tiene.

- El estado del piso (*dummy*): ‘1’ si está en buen estado o reformado, ‘0’ en otro caso.

- Si dispone de garaje (*dummy*): ‘1’ si tiene garaje, ‘0’ si no lo tiene.

- Si dispone de piscina (*dummy*): ‘1’ si tiene piscina, ‘0’ si no la tiene.

¹³ Otras variables como la duración (en días) del inmueble en Idealista.com, si el piso es de segunda mano o si dispone de trastero, terraza o jardín han resultado ser no significativas con nuestra muestra.

¹⁴ Resto de Triana se refiere a aquellos pisos situados en Triana pero en zonas diferentes a ‘Altozano - Pagés del Corro’ y a Plaza de Cuba.

Los resultados de la estimación se muestran en la Tabla 5.1¹⁵. La relación entre el precio y los metros cuadrados construidos es aproximadamente lineal –obsérvese que el valor del coeficiente del cuadrado de los metros cuadrados construidos (-2,6) es significativo pero relativamente pequeño–; de manera que un metro cuadrado más encarece el piso en unos 2.638 €.

En lo que se refiere a las diferentes zonas de Sevilla recogidas en la muestra, y tomando a ‘resto de Triana’ como categoría de referencia, se observa que las cinco zonas más caras (dentro de la muestra y controlando por el resto de características) son ‘Arenal - Museo’, Puerta de Carmona, Macarena, Alameda y Nervión¹. Por ejemplo, comprar un piso en ‘Arenal - Museo’ resulta casi 88.000 € más caro que hacerlo en ‘resto de Triana’; en Puerta de Carmona el sobre-precio supera los 60.000 €. Por su parte, las zonas más baratas son Manuel del Valle, Bermejales, Felipe II, Bami y San Pablo. Un piso en San Pablo se abarata unos 55.000 € respecto a ‘resto de Triana’, y un piso en Bami se abarata unos 40.000 €.

En cuanto a los baños, hay indicios de que las casas con 2 o 3 baños tienen un precio relativamente mayor que las casas con un solo baño o con cuatro baños¹⁶.

¹⁵ La estimación de los errores estándar es robusta a posibles problemas de heteroscedasticidad o de mala especificación del modelo en alguna covariable –véanse Huber (1967) y White (1980, 1982)–.

¹⁶ Se ha probado a interaccionar las variables número de habitaciones y número de baños dentro del modelo de regresión, obteniéndose alguna evidencia de que la interacción de pisos de 2 o más habitaciones y de dos baños tienen un efecto alcista en el precio. En cualquier caso, la evidencia sobre este hecho es débil dentro de la muestra, por lo que no se ha incluido la citada interacción dentro del modelo.

Tabla 5.1 Estimación del precio de la vivienda en Sevilla capital.

Variable endógena: Precio de oferta		Coefficiente	Error Estándar	t	P>t	[95% Intervalos de confianza]	
	Metros cuadrados contruidos	2.638,5	285,6	9,24	0,00	2.078,0	3.199,0
	(Metros cuadrados contruidos) ²	-2,6	0,9	-2,81	0,01	-4,5	-0,8
Zona	Arenal - Museo	87.976,0	10.153,5	8,66	0,00	68.048,1	107.903,9
	Puerta de Carmona	61.260,7	26.643,5	2,30	0,02	8.968,6	113.552,8
	Macarena	59.719,2	19.529,9	3,06	0,00	21.388,6	98.049,9
	Alameda	55.983,0	23.483,8	2,38	0,02	9.892,3	102.073,7
	Nervión1	52.163,5	10.441,7	5,00	0,00	31.670,1	72.657,0
	El Porvenir	45.889,6	17.810,7	2,58	0,01	10.933,3	80.846,0
	Centro	38.509,7	9.997,9	3,85	0,00	18.887,3	58.132,2
	Altozano-Pagés del Corro (Triana)	38.086,3	12.940,6	2,94	0,00	12.688,4	63.484,2
	Plaza Cuba (Triana)	37.818,6	16.235,3	2,33	0,02	5.954,4	69.682,9
	Puerta Osario	21.562,6	17.434,2	1,24	0,22	-12.654,8	55.780,0
	Parque de los Príncipes (Remedios)	16.376,6	15.507,6	1,06	0,29	-14.059,4	46.812,7
	Santa Clara	12.960,5	53.729,6	0,24	0,81	-92.492,4	118.413,4
	Nervión3	8.415,9	14.484,1	0,58	0,56	-20.011,6	36.843,3
	Asunción (Remedios)	7.807,9	14.054,3	0,56	0,58	-19.776,0	35.391,9
	Heliópolis	3.849,9	18.279,4	0,21	0,83	-32.026,4	39.726,2
	Nervión2	-3.568,6	9.784,5	-0,36	0,72	-22.772,2	15.635,0
	Tablada	-4.294,4	13.590,8	-0,32	0,75	-30.968,5	22.379,6
	Prado	-6.455,6	15.531,8	-0,42	0,68	-36.939,2	24.028,0
	Carr.Carmona-Miraflores	-17.075,5	10.954,0	-1,56	0,12	-38.574,5	4.423,6
	Arroyo - Santa Justa	-25.620,3	13.111,3	-1,95	0,05	-51.353,3	112,7
Manuel del Valle	-32.158,5	14.973,0	-2,15	0,03	-61.545,5	-2.771,6	
Bermejales	-34.281,7	13.070,0	-2,62	0,01	-59.933,5	-8.629,8	
Felipe II	-34.831,7	11.134,7	-3,13	0,00	-56.685,3	-12.978,0	
Bami	-39.867,8	11.316,0	-3,52	0,00	-62.077,2	-17.658,4	
San Pablo	-54.715,8	11.092,5	-4,93	0,00	-76.486,6	-32.945,0	
Habitaciones	2	-17.658,9	8.424,6	-2,10	0,04	-34.193,5	-1.124,3
	3	-34.518,2	9.207,4	-3,75	0,00	-52.589,2	-16.447,2
	4	-30.040,6	11.609,5	-2,59	0,01	-52.826,2	-7.255,1
	5	-43.172,9	13.822,7	-3,12	0,00	-70.302,1	-16.043,7
	6	-56.820,5	19.976,7	-2,84	0,01	-96.027,8	-17.613,1
Baños	2	8.301,7	5.080,2	1,63	0,10	-1.669,1	18.272,5
	3	46.572,3	11.475,0	4,06	0,00	24.050,8	69.093,8
	4	-5.172,6	22.980,2	-0,23	0,82	-50.274,9	39.929,7
Planta	p1	20.290,4	6.770,1	3,00	0,00	7.003,0	33.577,7
	p2	23.873,2	7.042,5	3,39	0,00	10.051,2	37.695,3
	p3	16.176,0	6.735,9	2,40	0,02	2.955,6	29.396,3
	p4	20.776,2	8.255,2	2,52	0,01	4.574,0	36.978,4
	p5	26.504,0	11.672,3	2,27	0,02	3.595,2	49.412,8
	p6	23.137,9	9.277,9	2,49	0,01	4.928,6	41.347,2
	p7	43.176,4	10.651,8	4,05	0,00	22.270,6	64.082,1
	p8	34.157,4	20.793,9	1,64	0,10	-6.654,0	74.968,7
	p9	23.658,2	11.817,3	2,00	0,05	465,0	46.851,5
	p10	-2.481,2	19.449,2	-0,13	0,90	-40.653,3	35.690,9
	p11	12.101,1	34.571,8	0,35	0,73	-55.751,6	79.953,8
	Ascensor	16.000,1	4.678,5	3,42	0,00	6.817,7	25.182,5
	En buen estado	15.710,0	5.488,5	2,86	0,00	4.937,9	26.482,0
	Garaje	23.992,5	5.176,5	4,63	0,00	13.832,8	34.152,3
	Piscina	58.274,7	8.465,6	6,88	0,00	41.659,5	74.889,8
	Término constante	-60.949,9	15.814,5	-3,85	0,00	-91.988,3	-29.911,5
Número de observaciones		935					
Log verosimilitud		-11484,29					
Prob > F		0,0000 (F=106,45)					
Sum. Cuadrado Residuos		2,75E+12					
R ²		83%					
R ² Ajustado		82%					
AIC		23070,58					
BIC		23317,28					

Fuente: Elaboración propia a partir de los datos.

El análisis de la planta en la que se encuentra el inmueble toma como categoría de referencia la planta baja. Moverse de la planta baja a alguna de las plantas inmediatamente superiores (1ª a 6ª) incrementa el precio, por lo general, en más de 20.000 euros. La planta mejor valorada en la muestra es la 7ª, que encarece el piso en más de 40.000 euros respecto a la planta baja.

Otros de los resultados obtenidos apuntan a que los inmuebles con ascensor y aquellos que se encuentran reformados o en buen estado son unos 16.000 € más caros que aquellos que no tienen ascensor o que necesitan reforma. Además, si el piso cuenta con plaza de garaje se encarece unos 24.000 €, y si dispone de piscina lo hace en unos 58.000 euros. Posiblemente, el que el inmueble esté situado en una urbanización con piscina vaya asociado a otras calidades adicionales del piso que expliquen también el importante efecto alcista de esta variable en el precio requerido.

Hemos dejado para el final el análisis del efecto del número de habitaciones de la vivienda. Las variables *dummy* empleadas en la regresión toman como categoría de referencia las viviendas de una sola habitación. Los coeficientes obtenidos pueden resultar poco intuitivos porque atribuyen menos valor a los pisos que más habitaciones tienen. Este resultado se debe a que el número de habitaciones está muy correlacionado con el tamaño del piso (la correlación es del 76%) y dicho tamaño está ya capturando el efecto positivo del espacio disponible sobre el precio. Los coeficientes negativos parecen indicar que para cada tamaño de piso existe un número óptimo de habitaciones; si se supera dicho número óptimo el piso comienza a perder valor porque las habitaciones (y el resto de dependencias) empiezan a ser demasiado pequeñas. Para comprobar esta hipótesis, hemos incluido en la regresión mostrada en la Tabla 5.2 una nueva variable *dummy* (llamada ‘Piso pequeño o mediano’) que distingue dos posibles tamaños del piso: la variable toma el valor 1 en los pisos que tienen menos de 105 m² construidos y el valor 0 en los que tienen una cifra de m² construidos igual o superior a dicho valor de 105 –la cifra de 105 m² construidos se corresponde con la mediana de la distribución en la muestra–.

Tabla 5.2 Estimación del precio de la vivienda en Sevilla capital diferenciando por tamaño.

Variable endógena: Precio de oferta		Coficiente	Error Estándar	t	P>t	[95% Intervalos de confianza]	
	Metros cuadrados contruidos	2.501,6	368,9	6,78	0,00	1.777,5	3.225,6
	(Metros cuadrados contruidos) ²	-2,4	1,1	-2,16	0,03	-4,6	-0,2
Zona	Arenal - Museo	85.457,9	10.271,6	8,32	0,00	65.298,1	105.617,8
	Puerta Carmona	61.499,5	25.984,3	2,37	0,02	10.500,8	112.498,2
	Macarena	58.096,8	20.308,5	2,86	0,00	18.237,8	97.955,8
	Alameda	50.989,1	24.465,0	2,08	0,04	2.972,3	99.005,9
	Nervión1	48.660,7	10.615,8	4,58	0,00	27.825,4	69.496,1
	El Porvenir	41.192,8	18.247,2	2,26	0,02	5.379,6	77.006,1
	Altozano-Pagés del Corro (Triana)	36.536,5	13.303,3	2,75	0,01	10.426,5	62.646,6
	Centro	35.808,3	10.285,6	3,48	0,00	15.621,0	55.995,6
	Plaza Cuba (Triana)	31.016,4	16.924,6	1,83	0,07	-2.201,0	64.233,8
	Puerta Osario	20.910,9	17.303,1	1,21	0,23	-13.049,5	54.871,2
	Parque de los Príncipes (Remedios)	15.201,0	15.419,5	0,99	0,32	-15.062,4	45.464,4
	Santa Clara	10.831,2	51.329,4	0,21	0,83	-89.911,7	111.574,0
	Nervión3	9.039,0	14.713,1	0,61	0,54	-19.838,0	37.915,9
	Asunción (Remedios)	2.926,1	13.861,2	0,21	0,83	-24.278,8	30.131,1
	Prado	160,9	12.680,7	0,01	0,99	-24.727,0	25.048,9
	Heliópolis	-744,6	18.250,0	-0,04	0,97	-36.563,4	35.074,1
	Tablada	-4.666,7	14.046,6	-0,33	0,74	-32.235,6	22.902,1
	Nervión2	-5.289,6	10.062,8	-0,53	0,60	-25.039,6	14.460,5
	Carr. Carmona-Miraflores	-16.353,2	10.999,7	-1,49	0,14	-37.942,1	5.235,6
	Arroyo - Santa Justa	-24.093,1	12.652,0	-1,90	0,06	-48.924,8	738,6
	Manuel del Valle	-29.653,3	15.371,8	-1,93	0,05	-59.823,0	516,5
	Felipe II	-34.386,6	10.985,9	-3,13	0,00	-55.948,4	-12.824,9
	Bami	-39.632,0	11.313,0	-3,50	0,00	-61.835,8	-17.428,2
Bermejales	-39.662,8	13.406,7	-2,96	0,00	-65.975,7	-13.349,8	
San Pablo	-55.885,7	11.290,7	-4,95	0,00	-78.045,7	-33.725,7	
	Piso pequeño o mediano (< 105 m ²)	55.623,9	20.477,9	2,72	0,01	15.432,5	95.815,2
Habitaciones	2	42.600,4	19.495,2	2,19	0,03	4.337,8	80.863,0
	3	35.349,6	15.861,3	2,23	0,03	4.219,1	66.480,1
	4	33.484,9	15.428,1	2,17	0,03	3.204,7	63.765,1
	5	20.888,5	16.745,4	1,25	0,21	-11.977,3	53.754,3
	6	7.251,1	21.793,6	0,33	0,74	-35.522,6	50.024,7
Interacción piso pequeño o med. con número de habitaciones	Pequeño o med. - 2 habitaciones	-59.889,0	21.721,3	-2,76	0,01	-102.520,7	-17.257,3
	Pequeño o med. - 3 habitaciones	-75.941,4	19.059,7	-3,98	0,00	-113.349,3	-38.533,5
	Pequeño o med. - 4 habitaciones	-69.131,8	20.014,8	-3,45	0,00	-108.414,4	-29.849,2
Baños	2	5.796,9	5.136,9	1,13	0,26	-4.285,1	15.879,0
	3	44.579,9	11.637,8	3,83	0,00	21.738,8	67.421,1
	4	-5.576,9	22.962,1	-0,24	0,81	-50.644,0	39.490,3
Planta	p2	20.047,2	6.698,4	2,99	0,00	6.900,5	33.194,0
	p3	24.463,7	7.040,6	3,47	0,00	10.645,2	38.282,1
	p4	16.195,5	6.692,6	2,42	0,02	3.060,2	29.330,9
	p5	20.578,4	8.260,9	2,49	0,01	4.365,0	36.791,9
	p6	26.280,9	11.516,4	2,28	0,02	3.678,0	48.883,9
	p7	22.489,0	9.442,1	2,38	0,02	3.957,2	41.020,7
	p8	41.147,7	10.549,8	3,90	0,00	20.442,0	61.853,5
	p9	32.911,7	20.906,6	1,57	0,12	-8.121,2	73.944,6
	p10	24.094,4	11.485,8	2,10	0,04	1.551,5	46.637,3
	p11	-1.875,7	19.662,3	-0,10	0,92	-40.466,4	36.715,0
	p12	9.953,5	32.547,8	0,31	0,76	-53.927,2	73.834,1
		Ascensor	16.877,8	4.697,7	3,59	0,00	7.657,8
En buen estado		13.941,9	5.502,7	2,53	0,01	3.141,8	24.741,9
Garaje		22.476,1	5.165,3	4,35	0,00	12.338,3	32.613,9
Piscina		56.395,9	8.443,1	6,68	0,00	39.824,8	72.967,0
Término constante		-102.895,2	32.229,4	-3,19	0,00	-166.151,0	-39.639,4
	Número de observaciones	935					
	Log verosimilitud	-11478,1					
	Prob > F	0.0000 (F=100,03)					
	Sum. Cuadrado Residuos	2,72E+12					
	R ²	83%					
	R ² Ajustado	82%					
	AIC	23066,2					
	BIC	23332,3					

Fuente: Elaboración propia a partir de los datos.

Tras introducir la nueva variable ficticia, se observa que el valor de un metro cuadrado construido adicional ha descendido algo respecto a la estimación anterior de 2.638 €, situándose en los 2.502 €. Una vez que se ha controlado por el resto de factores, se observa que el efecto de ser un piso pequeño o mediano incrementa el término constante de la regresión en más de 55.000 euros¹⁷ –véase la Figura 5.1–. Algunos factores no controlados en la regresión deben estar detrás de los mayores precios de partida (en Idealista) de los pisos pequeños y medianos; posiblemente se trate de inmuebles que experimentan una mayor demanda relativa¹⁸.

Figura 5.1 Relación estimada entre el precio requerido y los m2 construidos (diferenciando por tamaño en 105 m2).



Fuente: Elaboración propia a partir de los datos.

Una vez que se ha controlado por el tamaño del piso (pequeño o mediano vs. grande), la regresión muestra un efecto positivo, aunque ligeramente decreciente, de poseer más habitaciones, aunque dicho efecto es corregido a la baja cuando nos encontramos en pisos pequeños o medianos, donde poseer más habitaciones puede hacer decrecer demasiado el espacio por habitáculo¹⁹.

Nuestro análisis econométrico concluye con el análisis del residuo de esta última estimación (Tabla 5.2), o lo que es equivalente, con el análisis de la diferencia entre el precio requerido por la vivienda y el precio estimado por el modelo. Una diferencia negativa entre estos dos precios (que implica un residuo negativo) indicaría que nos encontramos ante un piso que podría ser considerado como una oportunidad de inversión, ya

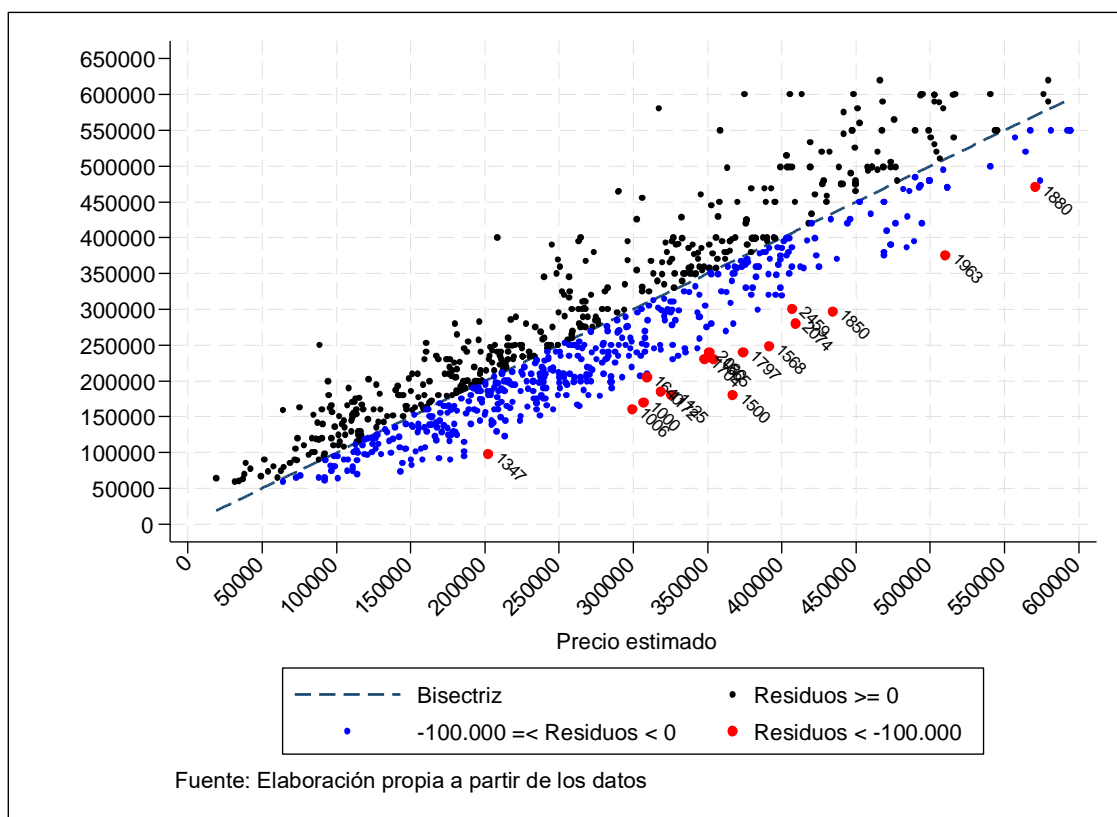
¹⁷ El segundo gráfico de la Figura 4.5 ya mostraba que, una vez que se tienen en cuenta los metros cuadrados construidos, el precio unitario del metro cuadrado era superior (en media) en las viviendas de tamaño pequeño o mediano

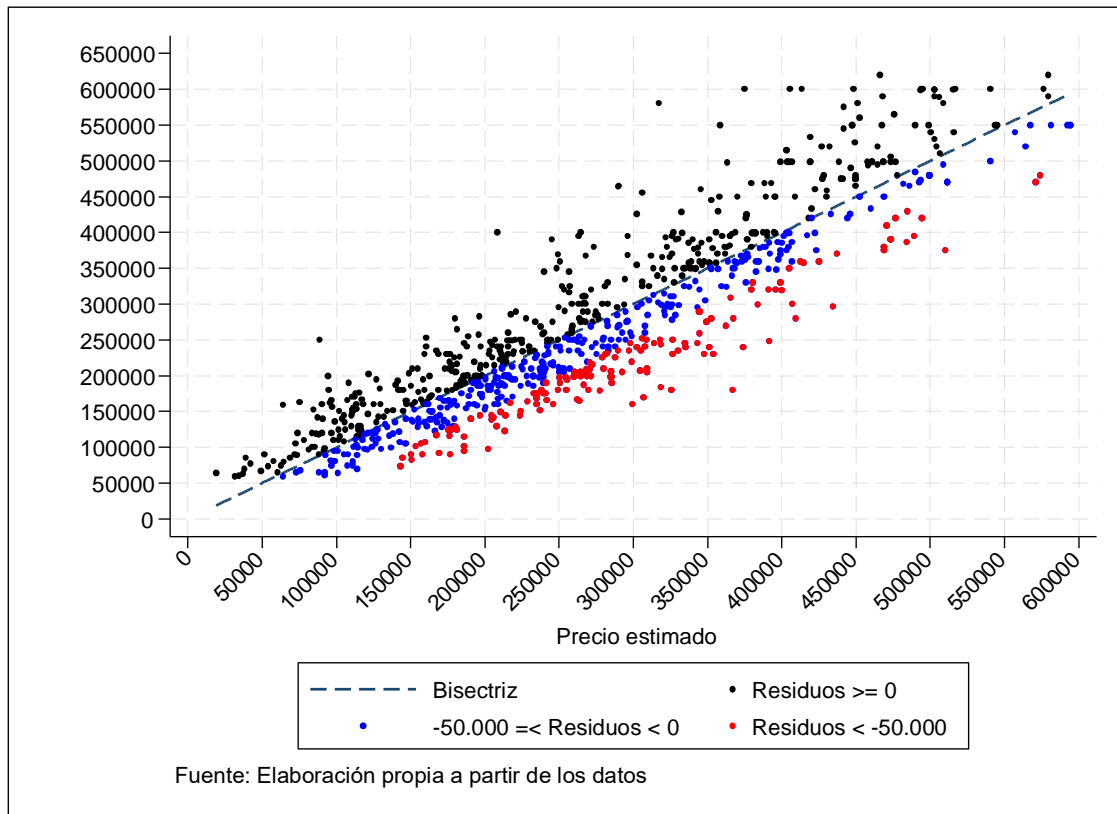
¹⁸ También podría estar pasando que los precios requeridos inicialmente por los vendedores de pisos pequeños o medianos sean relativamente superiores a los de los pisos mayores. Si esto es así, sería de esperar que los precios definitivos (los que se observan en la transmisión de la vivienda) acaben descendiendo, respecto a los inicialmente requeridos, relativamente más en el caso de los pisos pequeños y medianos.

¹⁹ Este resultado está asociado principalmente a las viviendas más pequeñas, las que se mueven en el entorno de los 50 m² construidos, tal y como se puede apreciar en el segundo panel de la Figura 4.5.

que, según el modelo, el precio que debería tener el piso (precio estimado) es superior a aquel que aparece en el anuncio registrado en Idealista. La Figura 5.2 muestra la relación entre el precio requerido y el precio estimado y señala algunos pisos que podrían constituir una oportunidad de inversión. El primer gráfico de la figura señala en color rojo aquellos pisos cuyo residuo es menor a -100.000 € –además, se indica el precio por metro cuadrado construido para cada uno de ellos–; el segundo gráfico marca en rojo aquellos pisos cuyo residuo es menor a -50.000 €. El precio medio del m^2 para los pisos con residuo menor a -100.000 € es de 1.631 €/m², mientras que dicho valor medio asciende a 1.891 €/m² en el caso de los pisos con residuo menor a -50.000 €. Ambos valores se encuentran por debajo del valor medio del precio por metro cuadrado en la muestra, que es de 2.364 €/m².

Figura 5.2 Relación entre el precio requerido y el precio estimado.





Fuente: Elaboración propia a partir de los datos.

A modo de conclusión, podemos decir que el precio de venta de un piso dentro de una determinada ciudad depende fundamentalmente de los metros cuadrados de construcción y de la zona de la ciudad donde se encuentra ubicado. Sin embargo, además de estas variables principales, existen otros muchos atributos, no todos ellos directamente observables, que pueden explicar las diferencias de precios observadas entre inmuebles que tienen el mismo tamaño y se encuentran situados en la misma zona, como el ratio ‘tamaño del piso / número de habitáculos’, la disponibilidad de garaje o la disponibilidad de más de un cuarto de baño. Una vez que se ha incorporado toda la heterogeneidad observada relevante a la ecuación de precios, las diferencias entre los precios reales y estimados pueden interpretarse en términos de posibles oportunidades de inversión.

6 CONCLUSIONES.

Mediante la metodología de precios hedónicos de Rosen (1974) se ha realizado un análisis econométrico que permite considerar cuáles son los principales determinantes del precio de un inmueble de la ciudad de Sevilla en el año 2017, centrándonos en algunos de los barrios más importantes de la capital.

El trabajo ha requerido la revisión de la literatura de varios autores como Ricardo Verges (2012) entre otros muchos, que coinciden en su mayoría en que las variables determinantes en el precio de la vivienda son los m² y la localización, centrando su investigación principalmente en el modelo de precios hedónicos de Rosen (1974).

Rosen (1974) pretende fundamentalmente aportar un modelo teórico que permita obtener y explicar los coeficientes de regresión hedónicos. Para ello, el autor propone un modelo de equilibrio competitivo que se desarrolla en espacio de varias dimensiones en el que compradores y vendedores se sitúan. Cada tipo de bien viene definido por un vector z de n características $z = \{z_1, z_2, \dots, z_n\}$ que representa un punto en el espacio — z_i mide la cantidad de la característica i -ésima que contiene cada bien—. Los oferentes tienen la posibilidad de ofrecer diferentes lotes de características dando lugar a diferenciación de producto. Cada paquete de características viene dado por un precio $p(z) = p(z_1, z_2, \dots, z_n)$ que viene dado para oferentes y demandantes dado su pequeño peso o tamaño en relación al mercado. La condición de equilibrio del mercado establece que el precio $p(z)$ se ajustará para que el número de unidades ofrecido en cada punto del espacio se iguale al número de unidades demandadas en ese punto, dando lugar a un equilibrio espacial de precios que determinan la igualdad de oferta y demanda para cada clase de bien ofrecido. Este equilibrio se basa en la conducta maximizadora de ambos lados del mercado, dada una determinada distribución de los gustos de los consumidores y de los costes de los productores.

Para la realización de dicho estudio se ha incluido información sobre las características de una muestra de 935 viviendas (exclusivamente pisos) de Sevilla capital, recogidas en el portal web inmobiliario Idealista.com. Con todos estos datos, y tras analizar la muestra, se ha observado que los precios más comunes son 250.000€, 200.000€, 160.000€, 180.000€, 240.000€, 499.000€, 230.000€ y 360.000€, aunque la mayor parte de la distribución de precios se concentra en el intervalo 150.000-250.000€; que los m² de construcción más corrientes son 90 m², 100 m², 70 m², 75 m², 120 m², 150 m² y 80 m², aunque la distribución se concentra en el intervalo de 60-100 m², y en menor medida, el intervalo 130-160 m²; finalmente el precio por metro cuadrado, tiene como valores más frecuentes 2000 €/m², 3000 €/m², 3046€/m² y 2500 €/m², estando la distribución concentrada en el intervalo 1.500-3000 €/m². En cuanto a las habitaciones y baños, las combinaciones más habituales son pisos que tienen 4 habitaciones y 2 baños (19%) y aquellas que tienen 3 habitaciones y 2 baños (17.9%). Referido al análisis de la planta en

la que se ubica el inmueble, las observaciones más habituales son los pisos situados en primera, segunda y tercera planta, que representan más del 60% de la muestra. Finalmente atributos como ascensor y garaje se observan en el 75.5% y 29% respectivamente de los pisos; garaje y trastero en el 29% y 27% respectivamente y solo el 15% conjuntamente, mientras que con piscina y jardín sólo representan el 5.5%; téngase en cuenta que el 96.3% de la muestra son viviendas de segunda mano.

En una segunda parte del trabajo, se ha estimado un modelo de regresión lineal múltiple del precio de la vivienda en Sevilla de acuerdo con dicha muestra. Al estimar el precio del bien a partir de sus características, el modelo estadístico se corresponde con un modelo hedónico en el sentido de Rosen (1974). Como principal resultado, se ha obtenido que el precio de la vivienda depende sobre todo de los metros cuadrados de construcción y de la zona de Sevilla en la que se encuentra; por ejemplo, Arenal-Museo resulta casi 88.000 € más caro que la zona llamada en el estudio “Resto Triana”, mientras que “Bami” resulta 40.000 € más barato que “Resto Triana”.

Por otro lado, existen también otras variables que pueden explicar las diferencias de precio en viviendas del mismo tamaño situadas en la misma zona, como por ejemplo el ratio “tamaño del piso / nº de habitaciones” (poseer demasiadas habitaciones en el mismo espacio puede hacer decrecer el precio del piso), la disponibilidad de garaje o que tengan más de un baño –variables todas ellas con un efecto positivo sobre el precio–. Así por ejemplo movernos de la planta baja a alguna de las plantas superiores (1ª a 6ª) encarece el precio por lo general, en más de 20.000€. La planta mejor valorada de la muestra es la 7ª, que incrementa el precio en más de 40.000€ respecto a la planta baja. Otro de los datos obtenidos apuntan a que los inmuebles con ascensor y aquellos que se encuentran reformados, son unos 16.000€ más caros que aquellos que no tienen ascensor o necesitan reforma. A su vez, si el piso cuenta con plaza de garaje, se encarece unos 24.000€ y si dispone de piscina, lo hace en 58.000€.

Nuestro análisis econométrico finaliza con la observación de la diferencia entre el precio real requerido por la vivienda y el precio estimado por el modelo. Una diferencia negativa indicaría que estaríamos ante una oportunidad de inversión, ya que el precio observado de la vivienda sería inferior al que el modelo estima que debería tener.

En definitiva, el presente estudio emplea la metodología de los precios hedónicos para estimar el precio de venta de una vivienda y los principales atributos que lo determinan. Este tipo de modelos puede ser aplicado en diversas líneas de investigación y puede generar utilidad a diferentes agentes del mercado:

- Aplicaciones al campo de la teoría urbana, ya que podemos estudiar cómo va cambiando el precio de la vivienda a medida que nos aproximamos a los centros urbanos o a los focos de actividad económica de la ciudad.
- Utilidad por los mismos agentes de la propiedad inmobiliaria, para poder calcular de una forma fácil, barata y rápida el precio de la vivienda que el cliente desea vender.
- Utilidad para los constructores, con la finalidad de conocer los atributos más importantes para determinar el

precio de la vivienda, y así prestar mayor atención a estos a la hora de construir una vivienda nueva.

- Utilidad para los organismos públicos, de cara a poder comparar tanto a nivel temporal como espacial, el precio de las transacciones inmobiliarias y de los determinantes del mismo.
- Finalmente, resulta útil para el propio demandante, como método para comprobar si está o no ante una posible “buena” inversión de compra de vivienda.

REFERENCIAS.

- Bilbao, C. 2000. Relación entre el precio de venta de una vivienda y sus características: un análisis empírico para Asturias. Universidad de Oviedo. *Revista Asturiana de Economía* nº18, 141-150. Universidad de Oviedo.
- Bover, O., y Velilla, P. 2003. *Hedonic house prices without characteristics: the case of new multiunit housing*. Working paper nº117, European Central Bank.
- Caridad, J.M., Brañas, P., y de la Paz, M. 1997. *Análisis intraurbano del precio de las características de la vivienda en Córdoba: los barrios menos favorecidos*. 780-791. Mimeo, Universidad de Córdoba.
- Chica, O., Cano, R., y Chica, M. 2007. Modelo Hedónico espacio-temporal y análisis variográfico del precio de la vivienda. *Revista internacional de ciencia y tecnología de la información geográfica* nº7, 56-7. GeoFocus.
- Fernández, L., Llorca, A., Valero, S., y Botti, V.J. 2012. *Incidencia de la localización en el precio de la vivienda a través de un modelo de red neuronal artificial. Una aplicación a la ciudad de Valencia*. Catastro.
- García, A. . 2008. Determinantes del precio de la vivienda usada en Málaga. Una aplicación de la metodología hedónica. *Revista de estudios regionales* nº82, 0213-7585(208), 135-158.
- García, J., y Raya, J.M. 2011. *Price and income elasticities of demand for housing characteristics in the city of Barcelona*, *Regional Studies*, 45(5), .597-608. .
- García, J., Raya, J.M. y García y Raya. 2015. *Use of a Gini index to examine housing price heterogeneity: a quantile approach*. *Journal of housing economics* 29(2015), 59-71.
- Huber, P. J. 1967. The behavior of maximum likelihood estimates under nonstandard conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, 221–233. Berkeley: University of California Press.
- Instituto de Estudios Fiscales [Sitio web] “Auge y crisis del sector inmobiliario reciente: interrelación con la política económica” [Consulta: 11 de octubre de 2017] Disponible en: http://portal.uned.es/pls/portal/docs/page/uned_main/launiversidad/ubicaciones/05/docente/jose_antonio_martinez_alvarez/documentos%202014_05.pdf
- Instituto nacional de estadística [Sitio web] “Índice de precios de vivienda” [Consulta: 11 de octubre de 2017] Disponible en: https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736152838&menu=ultiD atos&idp=1254735976607
- Instituto nacional de estadística [Sitio web] “Índice nacional de la vivienda en alquiler” [Consulta: 11 de octubre de 2017] Disponible en: <http://www.ine.es/jaxiT3/Tabla.htm?t=2643&L=0>
- Ministerio de Fomento [Sitio web] “Stock de vivienda nueva” [Consulta: 11 de octubre de 2017] Disponible en: <https://www.fomento.gob.es/informacion-para-el-ciudadano/informacion-estadistica/vivienda-y-actuaciones-urbanas/estadisticas/stock-de-vivienda-nueva/estadisticas-sobre-stock-de-vivienda-nueva>
- Nicodemo, C., y Raya, J.M.. 2012. *Change in the distribution of house prices across Spanish cities*. *Regional Science and urban economics* 42(2012), 739-748.
- Núñez, J., Ceular, N., y Millán, G. 2007. *Aproximación a la valoración inmobiliaria mediante la metodología de precios hedónicos (MPH)*. Conocimiento, innovación y emprendedores. Camino al futuro. 2688-2701.
- Portal inmobiliario Fotocasa [Sitio web] “Boom del alquiler turístico” [Consulta: 11 de octubre de 2017] Disponible en: <https://prensa.fotocasa.es/el-51-de-los-inquilinos-considera-que-el-boom-del-alquiler-turistico-esta-incrementando-los-precios/>

- Portal inmobiliario Fotocasa [Sitio web] “La vivienda de segunda mano en 2017” [Consulta: 11 de octubre de 2017] Disponible en: <https://prensa.fotocasa.es/la-vivienda-segunda-mano-2017/>
- Portal inmobiliario Fotocasa [Sitio web] “Perfil del comprador de obra nueva 2017-2018” [Consulta: 11 de octubre de 2017] Disponible en: <https://prensa.fotocasa.es/informe-perfil-comprador-de-obra-nueva-2017-2018/>
- Portal inmobiliario Fotocasa [Sitio web] “Perfil del hipotecado español 2017-2018” [Consulta: 11 de octubre de 2017] Disponible en: <https://prensa.fotocasa.es/informe-perfil-del-hipotecado-espanol-2017-2018/>
- Portal inmobiliario Fotocasa [Sitio web] “Radiografía del mercado de la vivienda 2017-2018” [Consulta: 11 de octubre de 2017] Disponible en: <https://prensa.fotocasa.es/radiografia-del-mercado-de-la-vivienda-2017-2018/>
- Portal inmobiliario Idealista [Sitio web] “Precio de la vivienda en alquiler” [Consulta: 11 de octubre de 2017] Disponible en: <https://www.idealista.com/news/inmobiliario/vivienda/2018/01/04/762249-el-precio-de-la-vivienda-en-alquiler-suben-un-18-4-en-2017-solo-barcelona-y-caceres>
- Portal inmobiliario Idealista [Sitio web] “Precio de la vivienda en Julio 2017” [Consulta: 11 de octubre de 2017] Disponible en: <https://www.idealista.com/news/inmobiliario/vivienda/2018/09/11/768148-los-notarios-confirman-mas-compraventas-de-viviendas-e-hipotecas-en-julio-pero-baja>
- Rosen, S. 1974. Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, vol. 82, nº1(1974), 34-55. University of Rochester and Harvard University The University of Chicago Press.
- Taltavull, P. y Mcgreal. 2009. *Measuring Price expectations: evidence from the Spanish housing market*. Journal of European real estate research.
- Verges, R. 2012. *Estadísticas de vivienda libre en España. 1987-2011. Series operativas de compraventas y precios*. Estadística española. Vol.54 nº179(2012), 357-420.
- White, H. L., Jr. 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica* 48 (4): 817–838.
- White, H. L., Jr. 1982. Maximum likelihood estimation of misspecified models. *Econometrica* 50 (1): 1–25.
- Wooldridge, J.M. 2010. *Introducción a la econometría. Un enfoque moderno*. 4ª edición. Capítulos 1-4.