



Vigilada Mineducación

Factores Determinantes del Precio de Arriendo de Viviendas en Ciudades
Colombianas: Un Enfoque de Modelos Hedónicos

Alejandro Heredia Barbosa

Felipe Cifuentes Vásquez

Trabajo de Grado

Asesora

Paula María Almonacid Hurtado

UNIVERSIDAD EAFIT

Escuela de Finanzas, Economía y Gobierno

ECONOMÍA

MEDELLÍN

2024

CONTENIDO

INTRODUCCIÓN	6
PLANTEAMIENTO DEL PROBLEMA.....	7
JUSTIFICACIÓN.....	8
OBJETIVOS.....	8
GENERAL	8
ESPECÍFICOS	8
MARCO TEÓRICO	9
METODOLOGÍA	12
RESULTADOS.....	20
Regresión líneal Múltiple	20
Modelo de regresión espacial (gwr), caso para medellín	23
CONCLUSIONES	24
REFERENCIAS	25

LISTA DE FIGURAS

Figura 1. Distribución de precios de arriendo de vivienda por ciudad.....	15
Figura 2. Gráfico de dispersión de la variable dl precio de arriendo y área en metros cuadrados.....	16
Figura 3. Matriz de correlación de variables continuas.....	17
Figura 4. Mapa de Medellín.	17
Figura 5. Mapa de Bogotá.....	18
Figura 6. Mapa de Cali.....	19
Figura 7. Regresión lineal simple para las ciudades Medellín, Bogotá y Cali.	21
Figura 8. Resultados Modelo de Regresión Espacial para la ciudad de Medellín.	24

LISTA DE TABLAS

Tabla 1. Estadísticas descriptivas de variables continuas para Medellín, Bogotá y Cali	14
Tabla 2. Distribución y precio promedio de arriendo por estratos socioeconómicos de viviendas por ciudad.....	15
Tabla 3. Estadísticas descriptivas de variables dicótomas para Medellín, Bogotá y Cali.	19

RESUMEN

Los precios de vivienda (arriendo y compra) han sido constantemente un tema de suma importancia para los *policymakers* cuando de realizar políticas de sostenibilidad urbana (y de otra índole) se trata. Recientemente dicho asunto ha tomado gran relevancia para el caso de Colombia por el crecimiento sostenido en los precios del mercado inmobiliario. Este estudio analiza los determinantes del precio de arriendo en las principales ciudades de Colombia (Bogotá, Medellín y Cali) para el año 2023, tomando un enfoque de modelos hedónicos, a través del uso de una base de datos con precios de arriendos, características estructurales de los inmuebles y variables de tipo espaciales (latitud y longitud) para dichas ciudades. Se encuentra que el estrato socioeconómico, el área del inmueble y el número de baños y garajes tienen un impacto positivo en el precio del arriendo, mientras que el número de habitaciones tiene un efecto negativo. Además, se identificaron características específicas que influyen en el aumento del precio, como la presencia de balcón o zonas verdes. Asimismo, se realiza una regresión Geographically Weighted Regression (GWR) para profundizar en el caso de la ciudad de Medellín (por la atención internacional que ha tenido en los últimos meses), y se logra observar heterogeneidad en la influencia de las variables estudiadas en diferentes áreas de la ciudad.

Palabras clave: modelo de precios hedónicos, datos espaciales, mercado inmobiliario, vivienda, GWR.

ABSTRACT

Housing prices (rental and purchase) have constantly been an issue of utmost importance for policymakers regarding urban (and other) sustainability policies. Recently, this issue has taken on great relevance in the case of Colombia due to the sustained growth in real estate market prices. This study analyzes the determinants of rental prices in the main cities of Colombia (Bogota, Medellin, and Cali) for the year 2023, taking a hedonic model approach, through the use of a database with rental prices, structural characteristics of real estate and spatial variables (latitude and longitude) for these cities. It is found that the socioeconomic stratum, the area of the property and the number of bathrooms and garages have a positive impact on the rental price, while the number of rooms has a negative effect. In addition, specific characteristics were identified that influence the increase in price, such as the presence of a balcony or green areas. A Geographically Weighted Regression (GWR) regression was also performed to study the case of the city of Medellin (due to the international attention it has received in recent months), and heterogeneity was observed in the influence of the variables studied in different city areas.

Keywords: hedonic price model, spatial data, real estate market, housing, GWR.

INTRODUCCIÓN

Cuestionar y entender la razón de los precios de bienes y servicios que las personas usamos, y aún más cuando son vitales para subsistir, se convierte en un tema que los diferentes agentes en una economía se plantean. La *vivienda* no es excepción, especialmente en países y ciudades que constantemente están experimentando crecimiento en temas de población, economía y cultura. En los últimos años (especialmente post-pandemia COVID-19) ha tomado suma relevancia los asuntos relacionados a los costos de la vivienda en diferentes ciudades de Colombia; ciudades principales donde se ha podido examinar a través de los medios de comunicación, las personas y los datos, que han experimentado un crecimiento sostenido en los costos de vida. Algunas de estas ciudades han percibido un aumento en el turismo, migración intra-nacional e internacional y la llegada de los nómadas *digitales*. Más allá de buscar una explicación por medio de estos factores, es importante también entender y tomar otra perspectiva desde aspectos estructurales y espaciales propias de estos inmuebles de vivienda.

Dado el interés en el tema, en esta investigación se busca encontrar respuesta desde una visión cuantitativa a la pregunta sobre ¿Qué factores afectan el precio de arriendo de la vivienda? Se busca también ponerle atención específicamente a los costos de arriendo de vivienda. Debería ser relevante para el caso de Colombia estudiarse el problema desde los precios de arriendos, pues según el Departamento Administrativo Nacional de Estadística (DANE, 2022, p. 6), para el año 2022 el 40.2% de la población nacional vivieron en modalidad de arriendo o subarriendo.

Con el fin de examinar el tema, se propone realizar un modelo de precios hedónicos donde se analicen las principales características estructurales de las viviendas y también se tenga en cuenta el factor geográfico intra-municipal para explicar los determinantes de los precios de arriendos para las viviendas en Medellín, Bogotá y Cali. Decidimos tomar estas tres ciudades, pues de acuerdo con cifras del DANE (2023) son las que cuentan con mayor población. Para el año 2023 la población de cada una fue de 2.4 millones, 7.7 millones y 2.2 millones respectivamente. Con el fin de realizar este estudio empírico, mediante técnicas de *webscraping* se tomaron datos de la página Fincaraiz.com (portal web donde se listan inmuebles para la venta y arriendo en Colombia) con información de inmuebles para viviendas en arriendo para el año 2023, después de una limpieza de datos obtuvimos una base de datos de 9265 observaciones para las tres ciudades.

Se encuentra que existen similitudes entre las tres ciudades con respecto a los determinantes de los precios de arriendo, variables como el área, número de habitaciones, baños y parqueaderos afectan positivamente el precio de los arriendos en estas tres ciudades.

PLANTEAMIENTO DEL PROBLEMA

Actualmente existen varias técnicas con el fin de modelar mercados tales como el de bienes raíces, y existen diferentes estudios empíricos aplicados a diferentes ciudades en el mundo. Para el año 2023 fue popular la discusión acerca de los altos precios de arriendo de los inmuebles para vivienda en las principales ciudades de Colombia, y se le atribuyen diferentes justificaciones para explicar este fenómeno. Planteando un modelo de precios hedónicos para encontrar atributos que influyan en estos precios, siendo formulado en algunas ciudades colombianas supondrá obtener resultados que ayuden a diagnosticar y entender el comportamiento del mercado inmobiliario segregado a viviendas, para el país dadas las características de los inmuebles que hacen parte de la oferta. Asimismo, proveerá una comparativa entre las ciudades a escoger, que permita entender que tan heterogéneo u homogéneo se comporta el mercado de bienes raíces dependiendo de la ciudad.

Los precios Hedónicos son definidos como los precios implícitos de los atributos de un bien, y son revelados a los agentes económicos mediante los precios de los productos diferenciados y las cantidades específicas de características asociadas a ellos (Rosen, 1974). El objetivo de los modelos de precios hedónicos en el mercado de bienes raíces es poder tener un modelo que prediga de manera precisa los precios del mercado teniendo características tangibles e intangibles de los bienes y adicionalmente otros factores influyentes (Monson, 2009, p. 64) Autores como Wing & Chin, pp. (2003, pp. 146–147) indican que la ventaja principal de usar el método de los precios hedónicos para modelar el mercado de bienes raíces se encuentra en el hecho de que solo se necesita información como el precio de la propiedad y sus características estructurales. De esta manera se encuentra una metodología sencilla donde a través de los coeficientes estimados en la regresión se pueden estimar los precios marginales de las características del inmueble.

Entonces así, teniendo en cuenta el interés constante en buscar explicaciones a los precios de un sector inmobiliario, y observando los constantes incrementos en los precios de arriendo para vivienda que se han observado en las principales ciudades de Colombia, vale la pena preguntarse ¿Cuáles son las características determinantes de los precios de arriendo de vivienda en las 3 principales ciudades de Colombia (Bogotá, Medellín y Cali) para el año 2023? Específicamente buscando entenderlos desde una perspectiva con datos estadísticos con una herramienta como el uso de un modelo de precios hedónicos. Naturalmente, de dicha interrogante se derivan otras cuestiones que son de potencial interés para diferentes agentes del sector. Teniendo los datos, es pertinente contestar a temas cómo si al comparar las ciudades seleccionadas existen determinantes que tengan más peso en algún lugar que en otro, y qué tan dispersos terminan siendo los precios en dichas comparaciones entre ciudades. Asimismo, si más allá de las características estructurales de los inmuebles, cuál es la importancia que juega el componente espacial en un modelo como el que buscamos realizar.

JUSTIFICACIÓN

Principalmente, la realización de esta investigación es de gran relevancia debido a la constante necesidad de comprender y explicar los factores determinantes que influyen en el precio de los arriendos de vivienda en las principales ciudades de Colombia. Dada la creciente preocupación social sobre los altos costos de arriendo y la escasez de vivienda asequible en el casco urbano de las ciudades más pobladas de Colombia, es fundamental identificar y analizar cómo las características estructurales y espaciales de los inmuebles impactan en los precios de alquiler. Además, este estudio proporcionará una visión alternativa respaldada por datos reales y actualizados del comportamiento del mercado inmobiliario, lo que beneficiará a los agentes demandantes del mercado e incluso a los entes públicos al ofrecerles información precisa para potencialmente les ayude a influir en la toma de decisiones.

Este trabajo de grado se justifica también en la medida en que proporcionará una comprensión alternativa al tema de los determinantes de los precios en el sector de bienes raíces con un enfoque a Colombia, lo que permitirá contribuir al existente conocimiento que se tiene sobre dicho sector y a modelos econométricos aplicados a este. La investigación, al apoyarse no solamente en un análisis exploratorio de datos reales si no también en una revisión de la literatura permitirá evaluar la validez en los hallazgos que se obtengan, y permitir que sus resultados sean utilizados como referencia para futuras investigaciones académicas en temas afines.

OBJETIVOS

GENERAL

Explicar las características propias del inmueble y las características espaciales que determinan el precio de los arriendos de vivienda en las tres principales ciudades de Colombia (Bogotá, Medellín y Cali) mediante la aplicación de un modelo de precios hedónicos usando datos de la oferta de inmuebles en arriendo en las ciudades seleccionadas para el 2023. Con el fin de proporcionar una visión respaldada por datos del comportamiento del sector de bienes raíces a los diferentes agentes del mercado inmobiliario.

ESPECÍFICOS

1. Señalar las investigaciones existentes sobre el modelo hedónicos de precios a través una revisión de la literatura y un análisis exploratorio de datos obtenidos de la página de la empresa Finca Raíz. Con el fin de comprender qué tanto los factores a considerar en la investigación influyen en los precios de alquiler para las ciudades mencionadas.
2. Establecer la relación entre variables propias de los inmuebles (como el área en metros cuadrados, número de habitaciones, número de baños, número de garajes), variables espaciales (latitud y longitud) con el precio de arriendo de estos (variable a ser explicada) específicamente para las ciudades de Medellín, Bogotá y Cali para el año 2023.

3. Evaluar los resultados de los modelos estimados para cada ciudad, para identificar similitudes y diferencias en los determinantes de los precios de alquiler entre ciudades

MARCO TEÓRICO

Buscar entender las características determinantes de los precios de arriendo en el sector de bienes raíces se complica dado el hecho que intuitivamente no se puede pensar en una fórmula precisa que explique meramente de manera numérica dicha cuestión. Los inmuebles cuentan con ciertas particularidades únicas que las hacen diferenciables entre sí, viene desde características estructurales (como el área del inmueble, número de habitaciones, número de garajes, antigüedad, entre otros) y características espaciales (como el barrio en el que se encuentra, el precio de los inmuebles cercanos, el número de establecimientos comerciales cercanos, entre otros). Estas particularidades dependiendo del área geográfica y de la función de utilidad de cada agente partícipe de este mercado pueden tener una valoración diferente. Lo anterior sugiere que los inmuebles son un bien heterogéneo que dificultan la tarea de valorizarlos. Uno de los métodos comúnmente usados en la literatura es el del modelo hedónico de precios porque permite que el valor comercial de un inmueble sea explicado por sus componentes individuales (Sirmans et al., 2005, pp. 3–43).

De acuerdo con la teoría de los precios hedónicos la utilidad de consumo de los bienes de las personas viene de los atributos o características del bien más que del bien en sí (Liang & Yuan, 2021, p. 2).

Para realizar modelos de precios hedónicos hay dos líneas de modelación que se han realizado en la literatura, econometría paramétrica y no paramétrica. Los modelos paramétricos son los más usados en cuanto a análisis de regresión de precios hedónicos, la simplicidad en la estimación e interpretación de los coeficientes es la mayor ventaja de los modelos paramétricos. Sin embargo, este método ha sido ampliamente criticado pues en ciertos casos el modelo paramétrico preseleccionado puede ser muy restringido para tomar en cuenta características inesperadas del entorno. Por otro lado, el enfoque no paramétrico es flexible y no preespecifica el modelo antes de ejecutarlo y además relaja los fuertes supuestos que conlleva ejecutar un modelo paramétrico (Owusu-Ansah, 2013, p. 21).

Malpezzi, (2002, pp. 68–69) recomienda una forma funcional log-lineal para la representación matemática del modelo de precios hedónicos aplicados al mercado de bienes raíces que sigue la forma:

$$\ln R = \beta_0 + S\beta_1 + N\beta_2 + L\beta_3 + C\beta_4 + \varepsilon$$

Donde $\ln R$ es el logaritmo natural de la renta del inmueble, y S, N, L y C pueden ser características estructurales, del sector, de la ubicación y contractuales, respectivamente y dependiendo de qué tan específicos se quiera ser. β_i y ε son los coeficientes de la regresión hedónica y el término de error respectivamente. El autor indica que esta forma funcional tiene esencialmente cinco ventajas a la hora de estimarlo. La primera es que permite la variación del valor en moneda de una característica en particular tal que el

precio de un componente dependa en parte en el resto de las características del bien inmobiliario. La segunda es la facilidad de interpretar los coeficientes obtenidos como un cambio porcentual aproximado en el valor del inmueble dado un cambio unitario en alguna de las variables independiente. La tercera es que esta forma funcional elimina problemas de heterocedasticidad. La cuarta es que es un modelo computacionalmente simple. Y por último permite flexibilidad para especificar las variables independientes que se necesiten.

Pasando a la parte espacial en los modelos de precios hedónicos, autores como Can, (1992, pp. 455–458) hace la aclaración de distinguir entre las dos externalidades (positivas o negativas) de los efectos de la locación de un inmueble. Según el autor puede haber efectos del “barrio” donde se capturan los efectos que las características del barrio tienen en la valorización del inmueble, y un efecto “derrame” (spillover) donde el precio de las propiedades adyacentes a cierto inmueble son determinantes para su valor. Bowen et al., (2001, pp. 468–469) explican que esta dependencia espacial en los precios de venta o arriendo en los inmuebles, puede tener una explicación en el comportamiento de los propietarios. En dicho comportamiento, los comportamientos se verían influenciados por los precios establecidos por sus vecinos, e incluso empresas que son agentes inmobiliarios, pues el mercado local va a tender a converger al precio escogido.

Zakaria & Fatine (2021) proponen un modelo hedónico para un índice de precios para el sector inmobiliario en Marruecos, con datos de corte transversal obtenidos para las tres principales ciudades (Rabat, Hay-Riad y Temara) de la región de la capital del país (región de Rabat) para el año 2014. Los autores proponen dicho modelo con el fin de poder identificar aquellos determinantes que explican los precios en el mercado de viviendas del país africano. Principalmente se tiene en cuenta características estructurales de los inmuebles (número de habitaciones, número del piso en caso de ser apartamento, antigüedad, si posee parqueadero o garaje, entre otros) y su ubicación. Se encontró que el determinante principal del precio del sector de venta de viviendas en esta región es la correlación espacial, es decir el barrio donde se encuentra un inmueble afecta el precio de mercado de este.

Turnbull & Dombrow, (2006, pp. 393–395) añaden a la evidencia empírica un artículo académico que busca evidenciar el comportamiento del precio de venta en el mercado de bienes raíces a través del componente espacial y las características del inmueble. Cabe notar que el estudio que hicieron en esencia busca determinar si al comparar dos inmuebles homogéneos (con características idénticas) influye la concentración de inmuebles en venta cercanos a estos. Los autores encontraron que el hecho de que un inmueble que esté a la venta esté en un sector donde se añaden más inmuebles para la venta, se genera un tráfico adicional de compradores, sin embargo, en zonas muy concentradas se da un efecto competitivo para los vendedores donde el precio de los inmuebles se termina reduciendo.

Malpezzi, (2002, pp. 78–79) sugiere que en un conjunto de datos completo incluiría variables tales como: el número de cuartos (habitaciones, baños, etc.), área de la vivienda, antigüedad, características estructurales tales como garajes o sótanos, variables representantes al barrio donde está localizado (tales como alguna variable de

carácter socioeconómico), y la fecha de recolección de los datos (en caso de que se quiera estudiar la evolución a través del tiempo).

Sirmans et al. (2005) realiza una discusión de aproximadamente 125 artículos científicos examinados donde se usaron técnicas de modelación hedónica. El autor sugiere que entre las características estructurales comúnmente usadas en este tipo de estudios son: tamaño del inmueble, área medida en pies cuadrados, antigüedad, número de habitaciones y de baños. De características internas, están la presencia de chimeneas, aire acondicionado, pisos en madera y sótano. Por el lado de características externas que usualmente explican el precio de venta de viviendas están la presencia de garajes, deck, piscinas, entre otras. También resalta que características relacionadas al medio ambiente habitualmente tienen un efecto positivo en el precio final de venta.

Sirmans et al. (2006, pp. 217–223) realizan un estudio donde hace uso de un análisis de meta-regresión (meta regression analysis o MRA en inglés), la cual tiene como fin comparar los coeficientes de regresiones estimadas en diferentes estudios, en este caso para hacer una comparación y evaluación de la literatura que relaciona diferentes atributos en una vivienda con su precio de venta. La herramienta de MRA es conveniente en análisis cuantitativo de literatura empírica y ayuda a condensar varios estudios de un tema relacionado. Para su estudio, los autores analizan nueve variables comúnmente usadas en modelos hedónicos para estimar el precio de una vivienda en Estados Unidos: área, tamaño del lote de la casa, antigüedad, habitaciones, baños, garaje, piscina, chimenea y aire acondicionado. En la literatura empírica que se tuvo en cuenta se debió haber hecho una regresión de precios hedónicos por MCO de forma cuasi-logarítmica para ser tomada en cuenta. Variables como el área (en pies cuadrados) es sensible al lugar geográfico donde se haya realizado el estudio, al igual que el tamaño del lote, antigüedad, número de baños, piscina y aire acondicionado; las otras variables no son afectadas por la ciudad de estudio. El hecho de que una variable sea sensible al lugar lo que el autor asume es que hay lugares donde una característica de la vivienda es relativamente más apreciada que en otros.

Con el fin de modelar la parte espacial, en el análisis del mercado de bienes raíces, una de las herramientas que se usa comúnmente es la Geographically Weighted Regression (GWR). Este tipo de regresión es una técnica sencilla que se deriva de la regresión lineal múltiple, pero permite observar una variación local en los coeficientes estimado más allá de asumir la existencia de un único estimado global para todas las observaciones (ignorando la localización de los datos), en otras palabras, permite estimar los coeficientes por cada punto geográfico a ser usado en los datos. Es razonable pensar en que cuando se realiza una regresión que tiene un componente localización se da la existencia de diferencias intrínsecas en cuanto a las relaciones que los datos pueden tener en la parte espacial (existencia de una heterogeneidad y autocorrelación espacial) (Brunsdon et al., 1998, pp. 433–435; Brunsdon et al., 1996, pp. 283–285) .

Sisman & Aydinoglu, (2022, pp. 2–5) realizan un estudio para analizar las dinámicas de mercado en el distrito de Pendik en Estambul (Turquía), a través de modelos de regresiones globales como MCO y con el fin de revisar las implicaciones de las variaciones geográficas usan un GWR. Cellmer et al. (2020, pp. 10–19) buscan demostrar un relación entre factores sociodemográficos, económicos y ambientales con

los precios promedios en el mercado de viviendas para el año 2018 en Polonia (datos de 380 condados, con un enfoque hacia la parte espacial; se encontró que en dicho estudio a través de modelos de tipo GWR la mayoría de variables utilizadas mostraban significancia estadística en términos de impacto de precios en el mercado y demostró para el caso de los condados de Polonia que los determinantes de los precios promedios en este mercado son espacialmente diferenciados como una consecuencia de diferencias culturales, económicas e históricas (los visualizaron a través de autocorrelación espacial y la concentración clústers high-high y low-low).

Brunsdon et al. (1996, p. 284), proponen una regresión de la forma

$$y_i = a_{i0} + \sum_{k=1,m} a_{ik}x_{ik} + \varepsilon_i$$

Donde a_{ik} es el valor del k-ésimo parámetro de una locación i . Esta forma de ecuación permite variaciones locales en tasas de cambio para los coeficientes en el modelo en lugar de asumir coeficientes globales.

METODOLOGÍA

Teniendo en cuenta la hipótesis mencionada anteriormente y la revisión de literatura realizada, se observa que en diferentes aplicaciones de los modelos de precios hedónicos como en los artículos de Potrawa & Tetereva (2022, p.57); Rey-Blanco et al. (2024, p.3-4) y Soler & Gemar (2018, p.128) se utiliza la regresión lineal simple como punto de partida para estimar este tipo de modelos. Asimismo, en los artículos de Soler & Gemar (2018, p.128) y Malpezzi (2002, p. 78-79) se explica que es mucho mejor estimar la regresión en términos del logaritmo del precio, pues así se aumenta el poder explicativo de la regresión. Por lo cual para el desarrollo de este trabajo de grado se usará una regresión lineal de este tipo para estimar los determinantes de los precios de los arriendos para las ciudades de Medellín, Bogotá y Cali.

ESPECIFICACIÓN DEL MODELO ECONÓMICO

La regresión estará dada por una ecuación en la cual tendremos como variable dependiente el logaritmo de los precios de los arriendos, y en las variables independientes se tendrán todas las características del inmueble, como número de baños y habitaciones, comodidades como piscina, ascensor, si cuenta con seguridad privada, etc.

$$\begin{aligned} \ln P = & \beta_0 + \beta_1 \text{Baños} + \beta_2 \text{Habitaciones} + \beta_3 \ln(\text{Área}) + \beta_4 \text{TipoPropiedad} + \beta_5 \text{Estrato} \\ & + \beta_6 \text{Gas} + \beta_7 \text{ParqueaderoVisitantes} + \beta_8 \text{Porteria} + \beta_9 \text{Balcón} \\ & + \beta_{10} \text{ZonasVerdes} + \beta_{11} \text{Vigilancia} + \beta_{12} \text{Bodega} + \beta_{13} \text{Garajes} \\ & + \beta_{14} \text{ConjuntoCerrado} + \epsilon \end{aligned}$$

Donde P son los precios de los arriendos, y cada una de las X desde X_1 hasta X_n son las características de los inmuebles las cuales son: Tipo de propiedad, estrato socioeconómico, número de habitaciones, número de baños, instalación de gas, parqueadero de visitantes, portería o recepción, balcón, zonas verdes, vigilancia, depósito o bodega, número de garajes, y si está en conjunto cerrado.

Adicionalmente se implementará el Modelo de Regresión Espacial GWR para así estimar las relaciones espaciales entre las variables y observar la heterogeneidad espacial de los resultados dependiendo de su ubicación geográfica. El modelo tendrá la siguiente ecuación, en la cual se tiene un vector g que es un vector de coordenadas que indica que hay un grupo de parámetros diferentes para cada una de las g observaciones, y la X es la matriz de variables independientes.

$$\ln P(g) = \beta_0 + \beta_1(g)X + \epsilon$$

DATOS

Los datos utilizados en esta investigación fueron recopilados de anuncios publicados en la página de Finca Raíz mediante un proceso de Webscraping. Esta recopilación se llevó a cabo exclusivamente para el año 2023, ya que el estudio se realizó hacia finales de dicho año. En total, se obtuvieron alrededor de 30,000 observaciones para las ciudades de Medellín, Bogotá y Cali. Estos datos fueron sometidos a un proceso de limpieza para eliminar información no relevante, datos faltantes y observaciones atípicas, como arriendos con valores extremadamente altos o inmuebles con características poco comunes. Como resultado de este proceso, se eliminaron aproximadamente 21,000 observaciones para las tres ciudades analizadas.

En cuanto a las variables, la herramienta de Webscraping extrajo información sobre precios de arriendo, área, número de habitaciones, baños y garajes, además de datos geográficos como la latitud y longitud. También se recopiló una variable con todas las comodidades o características del inmueble, de la cual se seleccionaron las 15 más frecuentes para generar variables categóricas. Además, utilizando la información de latitud y longitud, se determinaron las comunas (en Medellín y Cali) y las localidades (en Bogotá) a las que pertenecen los inmuebles. En resumen, la base de datos final cuenta con 23 variables y alrededor de 9,200 registros.

ESTADÍSTICAS DESCRIPTIVAS

Para los modelos que buscamos estimar, hicimos uso de datos de inmuebles de vivienda en arriendo para las ciudades de Medellín, Bogotá y Cali para el año 2023. En total contamos con 9265 observaciones de viviendas para las tres ciudades, e individualmente para Medellín hay 4539 observaciones, para Bogotá 2839 observaciones y para Cali 1887.

En la Tabla 1 se observan las estadísticas descriptivas de las variables continuas que usamos para los datos de Medellín, Bogotá y Cali. Como variable dependiente en nuestros modelos haremos uso del precio de arriendo de la vivienda en COP (Precio), y las variables explicativas de número de habitaciones (Habitaciones), número de baños (Baños), número de garajes (Garajes) y área en metros cuadrados (Área).

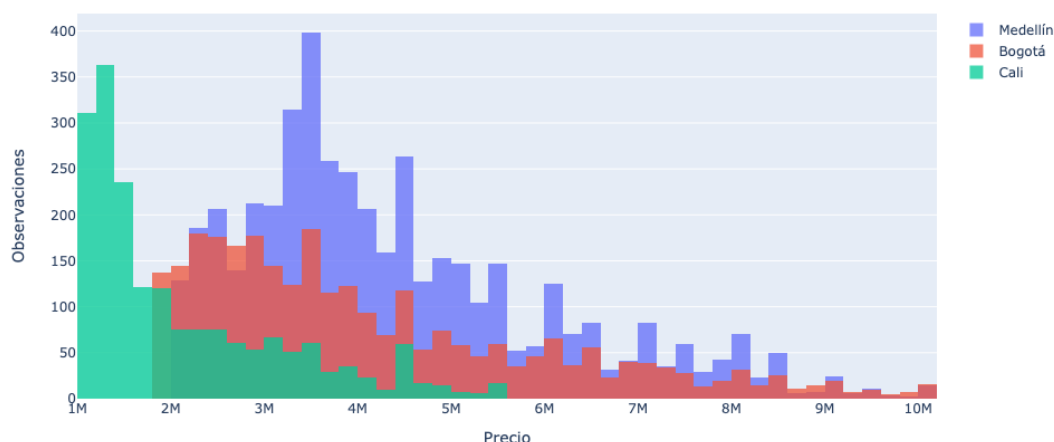
Tabla 1. Estadísticas descriptivas de variables continuas para Medellín, Bogotá y Cali

Ciudad	Variable	Media	Desviación Estándar	Valor Mínimo	Valor Máximo
Medellín	Habitaciones	2.71	0.8	1	5
	Baños	2.4	0.75	1	5
	Garajes	1.06	0.67	0	4
	Área	102.06	35.24	50	202
	Precio	4,360,210.83	1,672,274.38	2,000,000	10,147,801.00
Bogotá	Habitaciones	2.44	0.87	1	5
	Baños	2.46	0.81	1	5
	Garajes	1.49	0.69	0	4
	Área	108.83	43.59	50	225
	Precio	4,115,381.81	1,892,759.99	1,800,000	10,100,000
Cali	Habitaciones	2.62	0.78	1	5
	Baños	2.27	0.81	1	5
	Garajes	0.93	0.74	0	4
	Área	90.88	34.04	50	186.16
	Precio	2,072,191.23	1,079,610.12	1,000,000	5,500,000

Nota: Cálculos propios con base en datos extraídos para el año 2023 de <https://www.fincaraiz.com.co/>

En la Figura 1 podemos observar la distribución de los precios de arriendo en COP por ciudad. Se puede observar que en el caso Bogotá los datos se ven más dispersos, al igual que Medellín, sin embargo, visualmente Bogotá presenta picos más suaves y datos más distribuidos hacia las colas. Cali parece presentar más concentración en los precios entre \$1,000,000 y \$2,000,000.

Figura 1. Distribución de precios de arriendo de vivienda por ciudad.



Nota: Cálculos propios con base en datos extraídos para el año 2023 de <https://www.fincaraiz.com.co/>

En la Tabla 2 se tiene la distribución de los estratos socioeconómicos de las viviendas y el precio promedio de arriendo por ciudad, pero de manera proporcional a los datos de cada ciudad (en términos porcentuales). Para los datos de Medellín se presenta que hay proporcionalmente mayor cantidad de viviendas estratos 6 y menos observaciones para el estrato 3, lo mismo se observa para Bogotá, sin embargo, se puede apreciar que proporcionalmente hablando hay más viviendas estrato 6 y 4, que estrato 5 y 3. Para Cali, es interesante notar que la mayoría de las observaciones son estrato 4.

Tabla 2. Distribución y precio promedio de arriendo por estratos socioeconómicos de viviendas por ciudad.

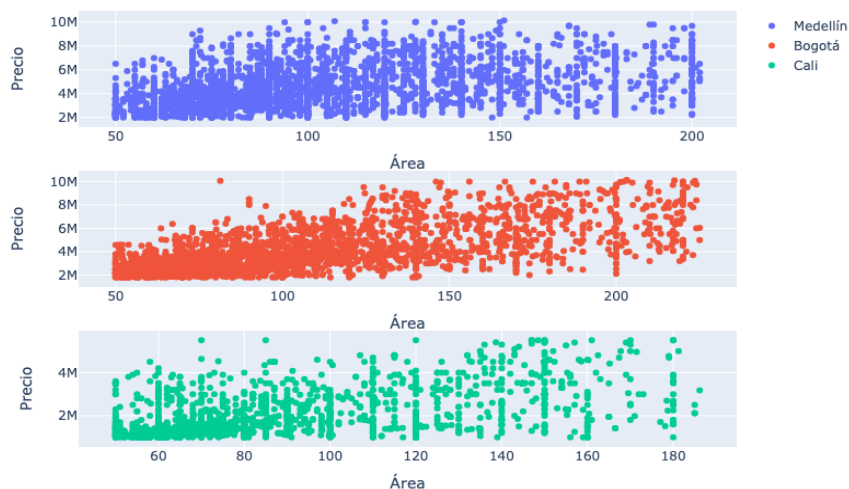
Ciudad	Estrato socioeconómico	Porcentaje	Precio promedio
Medellín	3	4.05%	\$2,811,924
	4	20.53%	\$3,196,123
	5	34.63%	\$3,918,109
	6	40.77%	\$5,475,715
Bogotá	3	7.99%	\$2,496,144
	4	31.45%	\$2,987,862
	5	23.77%	\$4,073,405
	6	36.77%	\$5,459,038
Cali	3	12.29%	\$1,293,984
	4	42.02%	\$1,512,252
	5	30.68%	\$2,491,247

	6	14.99%	\$3,421,812
--	----------	--------	-------------

Nota: Cálculos propios con base en datos extraídos para el año 2023 de <https://www.fincaraiz.com.co/>

En la Figura 2 se grafica la dispersión de la variable del precio con la variable del área en metros cuadrados. En general para Bogotá es donde se ve una tendencia un poco más clara donde a mayor área, mayor precio de la vivienda, sin embargo, para Cali y Medellín esta relación no se ve tan definida.

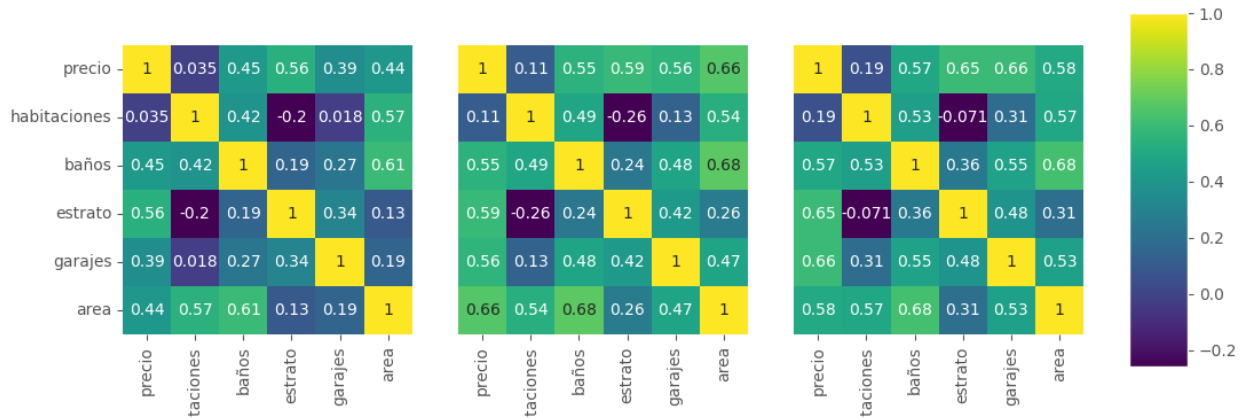
Figura 2. Gráfico de dispersión de la variable dl precio de arriendo y área en metros cuadrados.



Nota: Cálculos propios con base en datos extraídos para el año 2023 de <https://www.fincaraiz.com.co/>

En la figura 3 se tiene la matriz de correlación de las variables precio, habitaciones, baños, estrato, garajes y área para las 3 ciudades. Se aprecia que hay muy poca asociación entre el precio y el número de habitaciones para las tres ciudades. La cantidad de baños en una vivienda está débil y positivamente relacionada con el precio de esta, en todas las ciudades, pero es ligeramente más fuerte esta relación en Bogotá y Cali. Las variables que representan el estrato socioeconómico, el número de garajes y el área también muestran relación positiva con el precio del inmueble para las ciudades estudiadas.

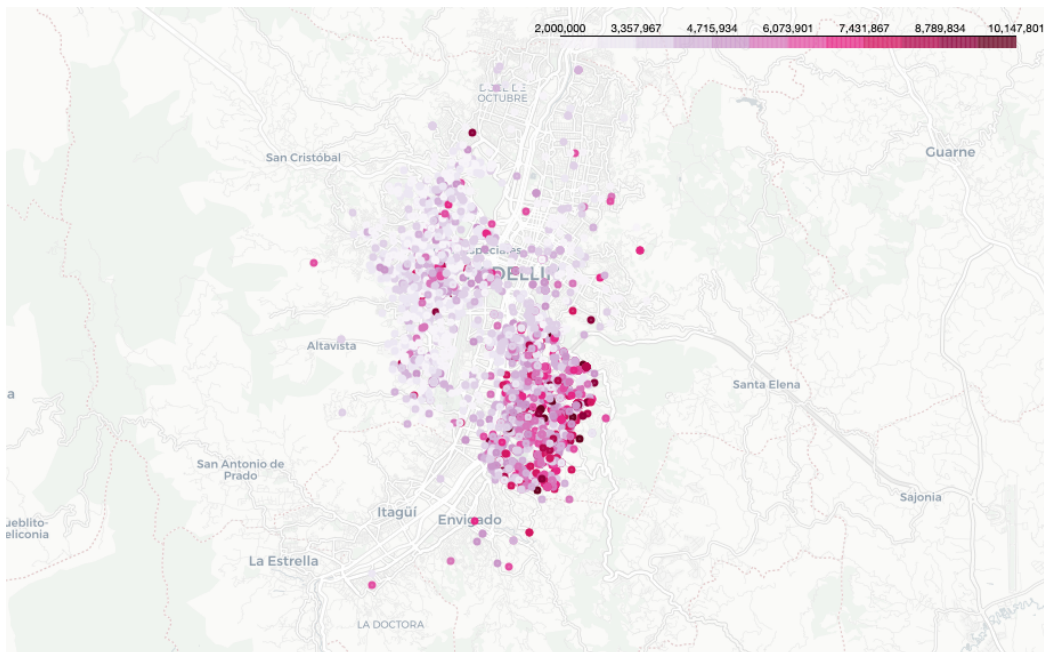
Figura 3. Matriz de correlación de variables continuas.



Nota: Cálculos propios con base en datos extraídos para el año 2023 de <https://www.fincaraiz.com.co/>

Ahora bien, es importante visualizar la locación física de las viviendas analizadas, para lo cual es relevante graficarlos en un mapa de cada una de las ciudades. Con este fin, en las Figuras 4, 5 y 6 se tienen los mapas de Medellín, Bogotá y Cali donde cada punto representa una vivienda y están discriminadas por color que representan los precios de arriendo. Para el caso de Medellín, gran parte de las viviendas están concentradas en el oriente y en el occidente de la ciudad, uno que otro está en la zona norte o sur. Asimismo, es relevante notar que las viviendas con precio de arriendo más costoso están ubicadas en la zona suroriental del municipio de Medellín.

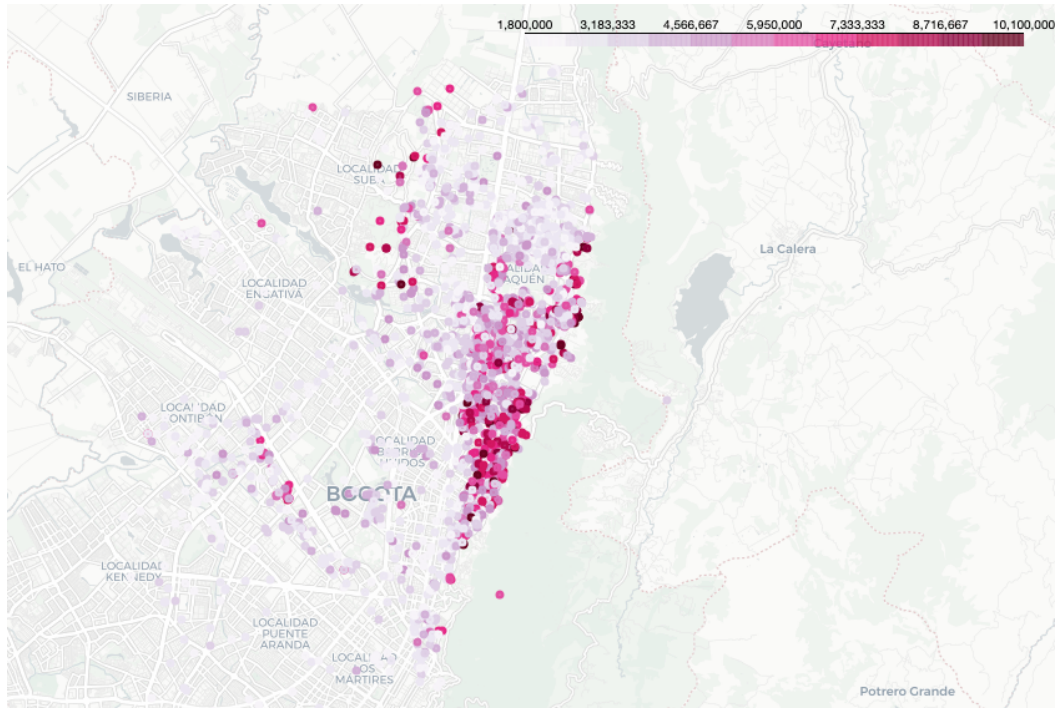
Figura 4. Mapa de Medellín.



Nota: Cálculos propios con base en datos extraídos para el año 2023 de <https://www.fincaraiz.com.co/>

En Bogotá, gran parte de las viviendas se encuentra en la zona del centro-oriente y nororiente de la ciudad, con los precios más altos concentrados en la zona oriental.

Figura 5. Mapa de Bogotá.

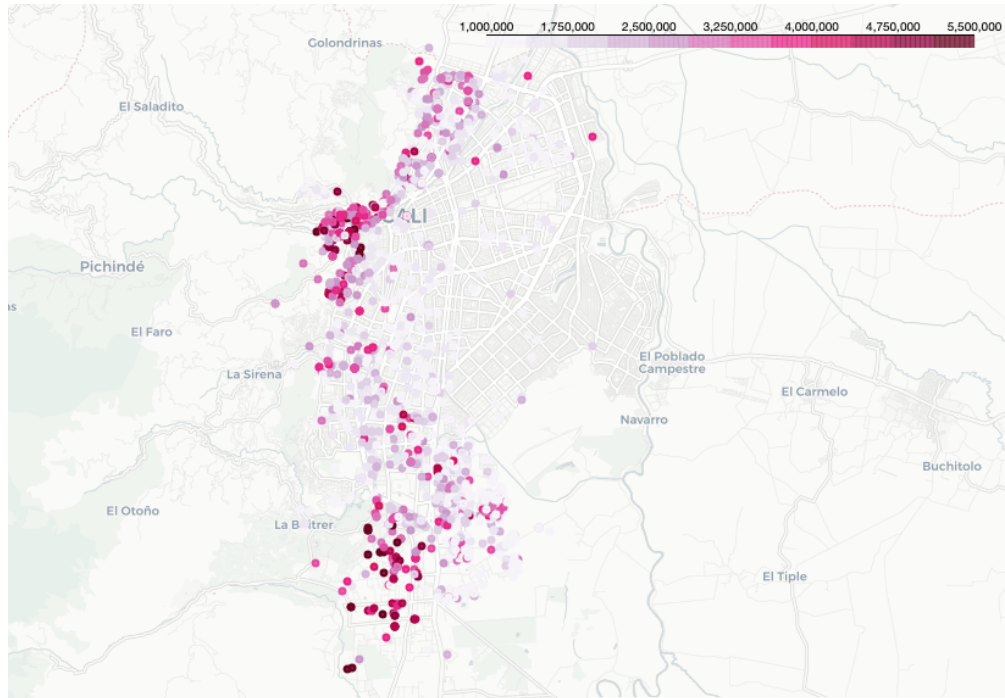


Nota: Cálculos propios con base en datos extraídos para el año 2023 de <https://www.fincaraiz.com.co/>

Para Cali, están las viviendas regadas en la zona occidental, con dos zonas concentradas con precios costosos, entre la zona norte y sur. Vea Figura 6.

Finalmente, en la Tabla 3 se observan otras variables explicativas, en este caso todas son variables dicótomas, para el caso de Medellín, Bogotá y Cali. Se presentan en términos porcentuales la presencia de dichas características en las viviendas que tenemos en nuestros datos.

Figura 6. Mapa de Cali.



Nota: Cálculos propios con base en datos extraídos para el año 2023 de <https://www.fincaraiz.com.co/>

Tabla 3. Estadísticas descriptivas de variables dicótomas para Medellín, Bogotá y Cali.

Variable	Medellín	Bogotá	Cali
Cocina Integral	88.2%	87.3%	90.8%
Zona de Lavandería	83.4%	85.9%	85.1%
Parqueadero Visitantes	73.5%	77.7%	80.8%
Portería / Recepción	72.2%	74.6%	78.9%
Balcón	81.8%	74.6%	82.5%
Calentador	76.4%	73.7%	72.6%
Zonas Verdes	68.4%	60.3%	75%
Closet	62.9%	61.3%	71.5%
Vía principal	50.7%	51.7%	57.1%
Vigilancia	74%	78.3%	82.8%
Salón Comunal	60%	65.3%	65.1%
Conjunto Cerrado	65.3%	58%	71.2%
Zona infantil	56.5%	50.9%	63.3%
Depósito / Bodega	50.3%	63%	50.5%
Zona Comercial	54.6%	50.9%	58.1%
Apartamento	91.1%	93.1%	91.3%
Casa	8.9%	6.9%	8.7%

Nota: Cálculos propios con base en datos extraídos para el año 2023 de <https://www.fincaraiz.com.co/>

RESULTADOS

REGRESIÓN LÍNEAL MÚLTIPLE

Dada la revisión de literatura revisada se realizan primero regresiones lineales múltiples para observar el efecto que tienen diferentes características de los inmuebles en su precio de arriendo. Para las tres ciudades se realizaron dos modelos, el primero en el cual se tiene las variables estrato, número de habitaciones, baños y garajes, y las características de las viviendas que se expresaron como variables dicótomas, pero no tiene componente de ubicación. Y el segundo en el que se tiene estas mismas variables y se incluyen unos efectos fijos por comuna o localidad para controlar la heterogeneidad entre ubicaciones geográficas.

En la figura 7 se observan los resultados para las regresiones lineales múltiples realizadas para cada ciudad. En la izquierda de cada una de las ciudades se encuentra la regresión estimada sin los efectos fijos espaciales, mientras que a la derecha de cada una se encuentra la regresión que si los incluye. En el análisis se usarán los coeficientes de la regresión que contiene dichos efectos fijos pues se observa que para cada ciudad el R^2 aumenta.

En el caso de la ciudad de Medellín ambos modelos tienen parámetros similares y mayormente con los signos o efectos esperados. Pero al parecer la inclusión los efectos fijos por comuna hacen que los coeficientes de ciertas de las dicótomas para las características dejen de ser significativas entre los modelos.

El modelo que incluye los efectos fijos por comunas tiene un R^2 ajustado mayor, pasando de un 59% en el modelo sin características espaciales a un 62% en el modelo con estas.

Un aumento en el estrato de la vivienda, el área y el número de baños y garajes tienen un efecto positivo en el precio del arriendo. Un incremento en el 1% del área de la vivienda genera un incremento del 0.45% en el precio del arriendo.

Un incremento de una habitación reduce el precio del arriendo en ambos modelos, una explicación de esto es que al añadir más habitaciones en un área fija se reduce el espacio para otras comodidades por lo cual el inmueble pierde valor de arriendo. Y pasando a las variables dicótomas de las características, las variables significativas tienen el signo esperado el hecho de que una vivienda contenga alguna de estas características aumenta el valor del arriendo, estas características son: Parqueadero de Visitantes, Balcón, Zonas Verdes, Closet, Sobre vía Principal, Salón Comunal, En conjunto cerrado y Zona Infantil. Se puede interpretar que el resto de las características no son tan importantes para el precio del arriendo en la ciudad de Medellín.

Con respecto a la variable Casa se puede observar que el hecho de que el inmueble sea una casa reduce su valor en un 3.1%, esto puede ser explicado porque en Medellín la mayoría de las publicaciones fueron apartamentos lo cual puede hacer que el estimador este sesgado.

Figura 7. Regresión lineal simple para las ciudades Medellín, Bogotá y Cali.

	Variable dependiente: Precio					
	Medellín		Bogotá		Cali	
	Sin efectos fijos	Con efectos fijos	Sin efectos fijos	Con efectos fijos	Sin efectos fijos	Con efectos fijos
Intercepto	12.189*** (0.062)	12.270*** (0.063)	11.565*** (0.068)	11.817*** (0.072)	11.535*** (0.110)	11.807*** (0.124)
Nro habitaciones	-0.074*** (0.006)	-0.063*** (0.006)	-0.090*** (0.007)	-0.056*** (0.007)	-0.058*** (0.011)	-0.044*** (0.012)
Nro baños	0.077*** (0.006)	0.070*** (0.006)	0.067*** (0.008)	0.080*** (0.007)	0.063*** (0.011)	0.062*** (0.012)
Nro garajes	0.055*** (0.006)	0.049*** (0.006)	0.088*** (0.008)	0.098*** (0.008)	0.174*** (0.011)	0.167*** (0.012)
Log area	0.436*** (0.017)	0.454*** (0.016)	0.601*** (0.019)	0.532*** (0.018)	0.390*** (0.028)	0.330*** (0.031)
Estrato	0.157*** (0.005)	0.112*** (0.006)	0.144*** (0.006)	0.121*** (0.006)	0.210*** (0.009)	0.194*** (0.010)
Balcón	0.072*** (0.010)	0.066*** (0.010)	0.020* (0.012)	0.023** (0.011)	0.038* (0.020)	0.047** (0.021)
Casa	-0.031** (0.015)	-0.020 (0.014)	-0.088*** (0.019)	-0.053*** (0.018)	-0.031 (0.024)	-0.034 (0.025)
Closet	0.016** (0.008)	0.013* (0.007)	-0.025*** (0.009)	-0.019** (0.009)	0.009 (0.014)	0.010 (0.015)
Cocina Integral	0.010 (0.012)	0.022* (0.012)	-0.041*** (0.015)	-0.049*** (0.014)	-0.019 (0.022)	-0.025 (0.024)
Depósito / Bodega	0.011 (0.007)	0.002 (0.007)	0.015 (0.010)	0.021** (0.009)	0.016 (0.013)	0.013 (0.014)
En conjunto cerrado	0.026*** (0.010)	0.023** (0.009)	-0.037*** (0.011)	-0.036*** (0.011)	0.005 (0.017)	0.017 (0.018)
Parqueadero Visitantes	0.028*** (0.010)	0.013 (0.010)	0.040*** (0.013)	0.039*** (0.012)	0.010 (0.020)	-0.005 (0.021)
Portería / Recepción	0.004 (0.010)	-0.005 (0.010)	-0.012 (0.013)	-0.017 (0.012)	-0.005 (0.019)	-0.001 (0.021)
Salón Comunal	0.033*** (0.009)	0.023*** (0.008)	0.031*** (0.011)	0.030*** (0.010)	-0.011 (0.015)	-0.001 (0.016)
Sobre vía principal	0.030*** (0.007)	0.026*** (0.007)	-0.016* (0.009)	-0.013 (0.008)	-0.025* (0.013)	-0.034** (0.014)
Vigilancia	0.004 (0.010)	0.010 (0.010)	0.008 (0.013)	0.015 (0.013)	0.012 (0.020)	0.004 (0.021)
Zona Comercial	-0.007 (0.008)	-0.008 (0.007)	0.002 (0.009)	-0.001 (0.008)	-0.015 (0.013)	-0.011 (0.014)
Zona Infantil	0.029*** (0.009)	0.025*** (0.009)	0.003 (0.011)	0.017 (0.010)	0.015 (0.016)	0.028* (0.017)
Zona de lavandería	-0.003 (0.011)	-0.003 (0.011)	0.042*** (0.014)	0.036*** (0.013)	0.039** (0.019)	0.021 (0.021)
Zonas Verdes	0.031*** (0.009)	0.007 (0.009)	-0.011 (0.011)	-0.010 (0.010)	-0.000 (0.017)	-0.012 (0.018)
Observaciones	4539	4475	2839	2835	1887	1518
R ²	0.595	0.627	0.721	0.755	0.685	0.719
R ² Ajustado	0.593	0.624	0.719	0.752	0.681	0.711
Error estandar Residual	0.233 (gl=4518)	0.225 (gl=4439)	0.229 (gl=2818)	0.215 (gl=2800)	0.263 (gl=1866)	0.255 (gl=1476)
Estadístico F	331.520*** (gl=20; 4518)	212.768*** (gl=35; 4439)	363.582*** (gl=20; 2818)	254.291*** (gl=34; 2800)	202.733*** (gl=20; 1866)	92.083*** (gl=41; 1476)

Nota: Las variables Balcón, Casa, Closet, Cocina Integral, Depósito / Bodega, En conjunto cerrado, Parqueadero Visitantes, Portería / Recepción, Salón Comunal, Sobre vía principal, Vigilancia, Zona Comercial, Zona Infantil, Zona de lavandería y Zonas Verdes, son variables categóricas que toman el valor de 1 si el inmueble posee estas características y 0 en caso contrario

P-values: *p<0.1; **p<0.05; ***p<0.01

Para la ciudad de Bogotá se observa que los coeficientes significativos de la regresión en su mayoría concuerdan con los resultados esperados. Nuevamente el modelo incluyendo los efectos fijos por localidades aumenta el R^2 ajustado con respecto al modelo que no las incluye, pasando de un 71% a un 75%.

Un aumento en el estrato de la vivienda, en el número de baños y en el número de garajes genera un aumento en el precio del arriendo. Un aumento de un 1% en el área del inmueble aumenta el precio del arriendo en alrededor 0.53%. Y nuevamente un aumento de una habitación disminuye el precio de la vivienda, el cual fue explicado anteriormente.

Ahora continuando con las variables dicótomas de las características, se encuentran efectos tanto negativos como positivos en las variables significativas. Las variables dicótomas zona de lavandería, parqueadero de visitantes, Balcón y Salón Comunal hacen que el precio del arriendo aumente si el inmueble tiene estas características. Al contrario, las variables Cocina Integral, Closet y En conjunto cerrado hacen que el precio del arriendo disminuya si el inmueble tiene estas características.

Con respecto a la variable Casa se puede observar que el hecho de que el inmueble sea una casa reduce su valor en un 5.3%, esto puede ser explicado porque en Bogotá al igual que en Medellín la mayoría de las publicaciones fueron apartamentos lo cual puede hacer que el estimador este sesgado.

Finalmente, para la ciudad de Cali nuevamente la inclusión de los efectos fijos por comunas hace que el R^2 aumente de un 68% a un 71% con respecto al modelo que no incluye estas variables.

Al igual que en Bogotá y Medellín un aumento en el estrato, número de baños y número de garajes aumentan el valor del precio del arriendo en un 19.4%, 6.2% y 16.7% respectivamente. Un aumento del 1% en el área de la vivienda conlleva a un aumento del 0.33% en el precio del arriendo. Adicionalmente, al igual que en las anteriores ciudades hay una relación negativa entre el número de habitaciones y el precio del arriendo.

Ahora bien, observando las variables categóricas de características significativas, al igual que en Bogotá se ven relaciones positivas y negativas. Las variables que tienen relación positiva con el precio del arriendo son: Zona de lavandería y Balcón. Mientras que la variable que tiene relación negativa es si la vivienda se encuentra sobre una vía principal, lo cual es lógico pues una vivienda puede perder atractivo si se encuentra cerca de lugares ruidosos como una vía principal. El resto de dicótomas podrían no ser relevantes a la hora de determinar el precio de un arriendo en la ciudad de Cali.

En último lugar, con respecto a las variables categóricas se observa que no todas son significativas, una explicación de esto es que los datos obtenidos de la página de finca raíz están acumulados en la zona occidental de la ciudad, por lo cual no hay suficientes datos para que los coeficientes de las otras comunas sean significativos.

MODELO DE REGRESIÓN ESPACIAL (GWR), CASO PARA MEDELLÍN

Si bien la regresión lineal múltiple es un método de estimación ampliamente aceptado en la literatura para modelar los precios hedónicos esta no tiene en cuenta ciertos aspectos espaciales de los datos como la heterogeneidad o la dependencia de estos. Por lo cual en caso de que alguno de estos fenómenos exista es más adecuado estimar un modelo de regresión espacial GWR. En este caso el modelo se aplicará para la ciudad de Medellín dado que se tiene buena información y conocimiento sobre las dinámicas de la ciudad, además del incremento que se ha venido evidenciando en los precios de arriendo de la ciudad en ciertas ubicaciones geográficas

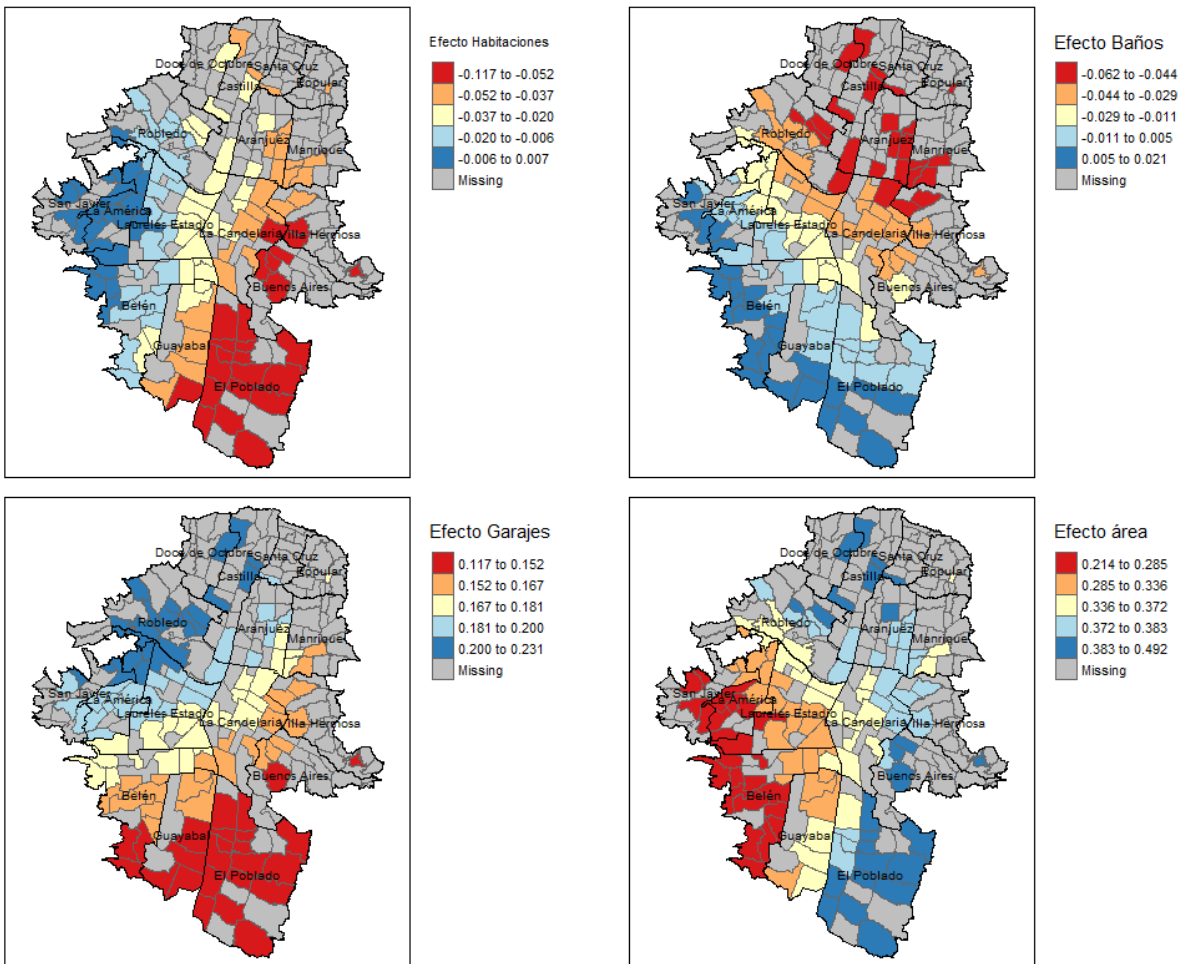
Ahora, para determinar si los datos tienen algún tipo de dependencia espacial se realiza una prueba de autocorrelación espacial llamado I de Moran, la cual si se rechaza la hipótesis nula existe evidencia de la existencia de autocorrelación espacial en los datos. Al realizar la prueba para los datos de la ciudad de Medellín la hipótesis nula se rechaza a un nivel de significancia del 5% y, por tanto, hay evidencia de autocorrelación espacial.

Una vez determinada la existencia de autocorrelación espacial se realiza la estimación de la regresión espacial, cuyos resultados son observados en la figura 8. Todos los coeficientes de la regresión estimada resultaron ser significativos con un 95% de confianza. Con respecto al número de habitaciones se observa que los coeficientes coinciden con la regresión lineal estimada en la sección anterior, pues son negativos. A su vez, se observa un efecto mayor en la zona sur occidental de la ciudad donde se encuentra la comuna 14 El poblado, por lo cual un número mayor de habitaciones representa una disminución del valor de arriendo en la vivienda en comparación al resto de la ciudad. Además, se observa que el efecto es mayor en el occidente y disminuye en el oriente de la ciudad.

Ahora analizando el número de baños se observa que un número mayor de baños aumenta en mayor proporción el valor del arriendo en el sur de la ciudad, mientras que el aumento es menor en el norte de la ciudad. Al igual que con las habitaciones el efecto tiene el mismo sentido que en la regresión lineal múltiple. Con respecto al número de garajes, se observa que en la zona nororiental de la ciudad un aumento en el número de garajes aumenta en mayor proporción el valor de arriendo de la vivienda, esto en comparación con el resto de la ciudad, mientras que hacia el suroccidente el aumento en el número de garajes aumenta en menor proporción el valor del arriendo.

Por último, con respecto al área, se observa un poco más de heterogeneidad en los resultados. Se observa que en la zona suroccidental de la ciudad un aumento de un 1% en el área de la vivienda aumenta el valor de arriendo en mayor proporción que en el resto de la ciudad, mientras que en la zona suroriental un aumento en el área tiene un impacto de menor proporción en el precio de arriendo de una vivienda.

Figura 8. Resultados Modelo de Regresión Espacial para la ciudad de Medellín.



CONCLUSIONES

En conclusión, la regresión lineal múltiple muestra como en las tres ciudades estudiadas hay consistencia con respecto al impacto del estrato, el área y el número de baños y garajes en el precio del arriendo, dado que en las tres ciudades el impacto fue positivo. Sin embargo, el número de habitaciones presenta un efecto negativo en el precio, sugiriendo que, en un espacio fijo, más habitaciones pueden reducir el valor de arriendo.

Además, se identifican características específicas que influyen en un aumento del precio, como la presencia de balcón o zonas verdes, mientras que otras, como la ubicación sobre una vía principal, pueden tener un efecto negativo. Adicionalmente se observa como la inclusión de efectos fijos espaciales en la regresión lineal múltiple hace que el ajuste de esta mejore en relación con una regresión que no los incluye.

Este resultado sugiere la importancia de considerar la variabilidad espacial inherente a cada ciudad al modelar el mercado de arriendo lo cual se realizó para la ciudad de Medellín. Para esta ciudad la influencia negativa del número de habitaciones en el valor del arriendo se confirma, destacando un efecto más pronunciado en ciertas áreas, como

el sur occidental de la ciudad (Comuna 14 el Poblado). Este hallazgo sugiere la existencia de variaciones geográficas en las preferencias de los arrendatarios y la oferta de viviendas en diferentes partes de la ciudad.

Además, se identifica una relación positiva entre el número de baños y garajes y el precio del arriendo, con variaciones en la magnitud del efecto según la ubicación. Estos resultados respaldan la idea de que las características estructurales de una vivienda influyen en su atractivo y, por tanto, en su precio de arriendo, pero también muestran cómo estos efectos pueden variar según el contexto geográfico. Además, se observa heterogeneidad en la influencia del área de la vivienda en el precio del arriendo, con impactos más significativos en algunas zonas que en otras. Este hallazgo resalta la importancia de considerar las dinámicas espaciales al analizar el mercado de arriendo y sugiere la existencia de factores contextuales específicos que pueden influir en la formación de precios en diferentes áreas de la ciudad de Medellín.

Para futuras investigaciones sería valioso profundizar en la comprensión de otros factores que podrían influir en el mercado de arriendo en estas ciudades colombianas. Por ejemplo, se podría explorar el impacto de variables socioeconómicas adicionales, como el nivel de ingresos de los residentes en diferentes áreas, la accesibilidad a servicios públicos y la calidad de la infraestructura urbana. Adicionalmente estudiar la relación espacial de estas variables para Bogotá y Cali.

REFERENCIAS

- Bowen, W. M., Mikelbank, B. A., & Prestegaard, D. M. (2001). Theoretical and Empirical Considerations Regarding Space in Hedonic Housing Price Model Applications. *Growth and Change*, 32(4), 466–490. <https://doi.org/https://doi.org/10.1111/0017-4815.00171>
- Brunsdon, C., Fotheringham, A. S., & Charlton, M. E. (1996). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geographical Analysis*, 28(4), 281–298. <https://doi.org/https://doi.org/10.1111/j.1538-4632.1996.tb00936.x>
- Can, A. (1992). Specification and estimation of hedonic housing price models. *Regional Science and Urban Economics*, 22(3), 453–474. [https://doi.org/https://doi.org/10.1016/0166-0462\(92\)90039-4](https://doi.org/https://doi.org/10.1016/0166-0462(92)90039-4)

- Cellmer, R., Cichulska, A., & Belej, M. (2020). Spatial Analysis of Housing Prices and Market Activity with the Geographically Weighted Regression. *ISPRS International Journal of Geo-Information*, 9(6), 10–19. <https://doi.org/10.3390/ijgi9060380>
- DANE. (2022). *Encuesta Nacional de Calidad de Vida (ECV) 2022. Boletín Técnico*. https://www.dane.gov.co/files/investigaciones/condiciones_vida/calidad_vida/2022/Boletin_Tecnico_ECV_2022.pdf
- DANE. (2023). *Serie nacional de población por área, para el periodo 2020-2070*. <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/proyecciones-de-poblacion>
- Liang, J., & Yuan, C. (2021). Data Price Determinants Based on a Hedonic Pricing Model. *Big Data Research*, 25, 2–3. <https://doi.org/https://doi.org/10.1016/j.bdr.2021.100249>
- Malpezzi, S. (2002). Hedonic Pricing Models: A Selective and Applied Review. In *Housing Economics and Public Policy* (pp. 67–89). <https://doi.org/https://doi.org/10.1002/9780470690680.ch5>
- Monson, M. (2009). Valuation Using Hedonic Pricing Models. *Cornell Real Estate Review*, 7, 62–72.
- Owusu-Ansah, A. (2013). A review of hedonic pricing models in housing research. *A Compendium of International Real Estate and Construction Issues*, 1, 17–38.
- Potrawa, T., & Teterewa, A. (2022). How much is the view from the window worth? Machine learning-driven hedonic pricing model of the real estate market. *Journal of Business Research*, 144, 50–65. <https://doi.org/https://doi.org/10.1016/j.jbusres.2022.01.027>
- Rey-Blanco, D., Zofío, J. L., & González-Arias, J. (2024). Improving hedonic housing price models by integrating optimal accessibility indices into regression and random forest analyses. *Expert Systems with Applications*, 235, 121059. <https://doi.org/https://doi.org/10.1016/j.eswa.2023.121059>

- Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy*, 82(1), 34–55. <https://doi.org/10.1086/260169>
- Sirmans, G. S., MacDonald, L., Macpherson, D. A., & Zietz, E. N. (2006). The Value of Housing Characteristics: A Meta Analysis. *The Journal of Real Estate Finance and Economics*, 33(3), 215–240. <https://doi.org/10.1007/s11146-006-9983-5>
- Sirmans, G. S., Macpherson, D. A., & Zietz, E. N. (2005). The Composition of Hedonic Pricing Models. *Journal of Real Estate Literature*, 13(1), 3–43. <http://www.jstor.org.ezproxy.eafit.edu.co/stable/44103506>
- Sisman, S., & Aydinoglu, A. C. (2022). A modelling approach with geographically weighted regression methods for determining geographic variation and influencing factors in housing price: A case in Istanbul. *Land Use Policy*, 119, 106183. <https://doi.org/https://doi.org/10.1016/j.landusepol.2022.106183>
- Soler, I. P., & Gemar, G. (2018). Hedonic price models with geographically weighted regression: An application to hospitality. *Journal of Destination Marketing & Management*, 9, 126–137. <https://doi.org/https://doi.org/10.1016/j.jdmm.2017.12.001>
- Turnbull, G. K., & Dombrow, J. (2006). Spatial Competition and Shopping Externalities: Evidence from the Housing Market. *The Journal of Real Estate Finance and Economics*, 32(4), 391–408. <https://doi.org/10.1007/s11146-006-6959-4>
- Wing, C. K., & Chin, T. (2003). A Critical Review of Literature on the Hedonic Price Model. *International Journal for Housing Science and Its Applications*, 27, 145–165.
- Zakaria, F., & Fatine, F. A. (2021). Towards the hedonic modelling and determinants of real estates price in Morocco. *Social Sciences & Humanities Open*, 4(1), 100176. <https://doi.org/https://doi.org/10.1016/j.ssaho.2021.100176>