



UNCUYO
UNIVERSIDAD
NACIONAL DE CUYO



FACULTAD DE
**CIENCIAS
ECONÓMICAS**

Formación de precios en el **sector inmobiliario** de C.A.B.A. en 2023: aplicación de un enfoque de **precios hedónicos** y “**non-market values**”

*Tesis de grado para
Licenciatura en Economía*

Por

Guillén Juan Cruz

N° de registro: 30.672

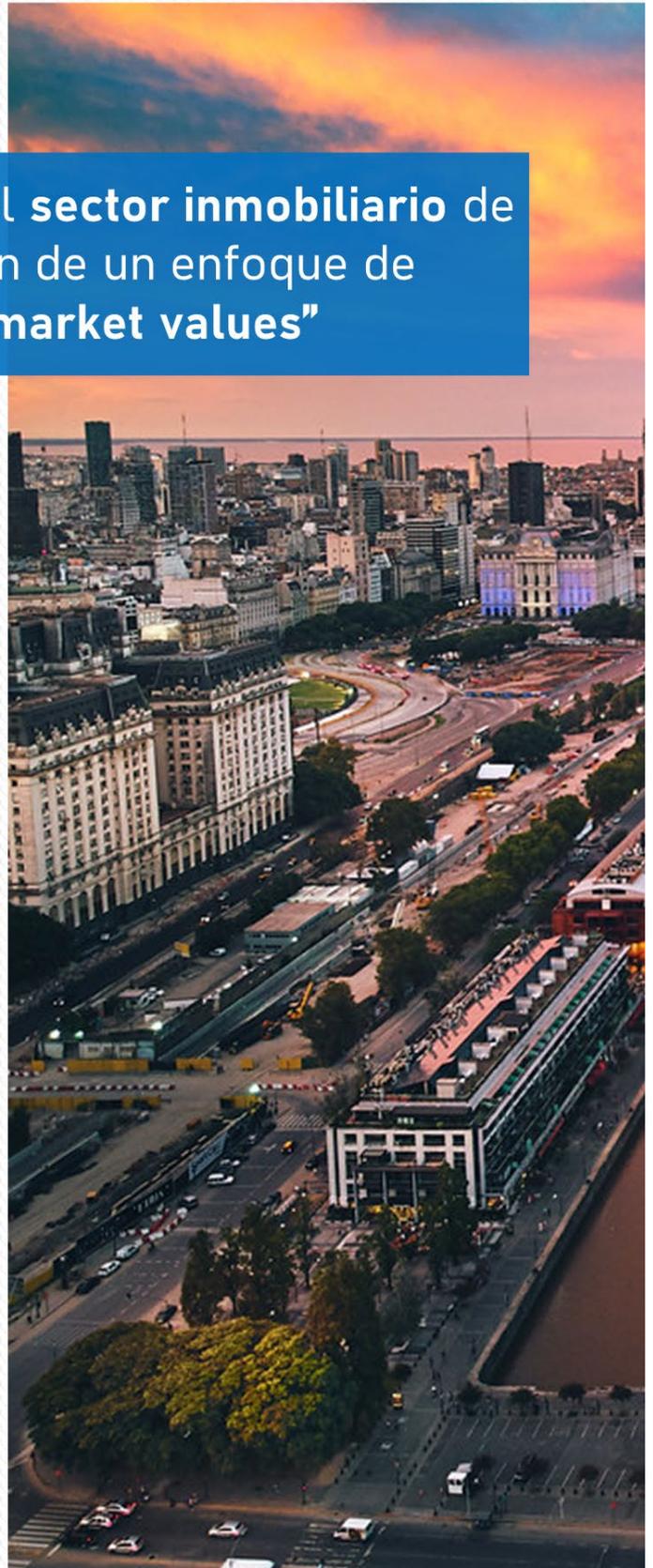
Squartini Santiago

N° de registro: 31.129

Tutoría de

Mg. Maradona, Gustavo

Mendoza, 2024



Resumen Técnico

El mercado inmobiliario de la Ciudad Autónoma de Buenos Aires se encuentra en un punto de inflexión previo a un cambio de gobierno nacional luego de un período de estancamiento. En este contexto, resulta clave analizar los factores que determinan la formación de precios de oferta de los inmuebles.

La presente investigación tiene como objetivo determinar la relación entre el precio de oferta de los inmuebles en venta en la Ciudad Autónoma de Buenos Aires en diciembre de 2023, y sus atributos estructurales, comunitarios, ambientales y de accesibilidad; mediante la aplicación de modelos hedónicos, un enfoque econométrico que descompone los precios de un bien o servicio según sus características individuales.

Además, esta investigación innova al incorporar en Argentina, "valores no de mercado" como variables explicativas adicionales, evaluando cómo factores ambientales como la calidad del aire y características escénicas, entre otros, influyen en las valuaciones.

Esta investigación adopta una metodología innovadora de recolección de datos a través de APIs comerciales y datos abiertos. Combinando información de portales inmobiliarios, organismos oficiales y fuentes no tradicionales, se construyó una base integral que representa un avance significativo para este tipo de análisis. El estudio brinda información oportuna y relevante para la toma de decisiones en el sector privado y el progreso del conocimiento académico en este campo, como así también para el diseño de políticas públicas.

Palabras claves

Precios Hedónicos – Economía espacial – Ciudad Autónoma de Buenos Aires – Valores de no mercado

Clasificación JEL

C31 – Q53 – R31

INTRODUCCIÓN

La presente investigación se sumerge en el análisis del mercado inmobiliario de la Ciudad Autónoma de Buenos Aires (C.A.B.A.) en diciembre de 2023. En este contexto, surge el interés de los autores por explorar los factores que determinan la formación de precios de oferta de los inmuebles en esta urbe.

El objetivo principal de este estudio es determinar la relación entre el precio de oferta de las propiedades en venta en C.A.B.A. y una amplia gama de atributos que las caracterizan, abarcando aspectos estructurales, ambientales, comunitarios y de accesibilidad. Para alcanzar este objetivo, se aplica la metodología de precios hedónicos, un enfoque econométrico que permite descomponer el precio de un bien en función de sus características individuales. Además, esta investigación innova al incorporar "valores no de mercado" como variables explicativas adicionales, evaluando cómo factores como la calidad del aire y las características escénicas del entorno pueden influir en los precios de las propiedades.

La tesis se estructura en cuatro capítulos principales, además de una sección final de apéndices. El primer capítulo, "Planteamiento del problema", se presenta el problema de investigación, las preguntas y objetivos que guían el análisis y las hipótesis a contrastar.

En el segundo capítulo, "Antecedentes, estado actual del tema y marco teórico", se realiza una exhaustiva revisión bibliográfica para contextualizar la evolución y los aportes de la teoría de precios hedónicos en el ámbito inmobiliario, explorando las contribuciones de diversos autores que han marcado el desarrollo de esta teoría. Asimismo, se detalla el marco teórico que sustenta el estudio, con énfasis en la aplicación de la teoría de precios hedónicos al mercado inmobiliario.

En el tercer capítulo, "Estimación de precios de oferta en el sector inmobiliario", se describe en detalle la metodología empleada, desde la selección de la muestra y la extracción de datos, hasta la especificación de los modelos econométricos y la interpretación de los resultados. Este capítulo constituye el núcleo analítico de la tesis.

El cuarto capítulo, "Conclusiones", presenta los hallazgos más relevantes del estudio, las implicancias para el mercado inmobiliario de C.A.B.A. y las reflexiones finales sobre el alcance y las limitaciones de la investigación.

Se presenta una sección de apéndices que complementan y profundizan aspectos técnicos y metodológicos, permitiendo al lector interesado conocer los detalles más específicos del análisis realizado.

La relevancia de esta investigación radica en su potencial para generar conocimiento valioso tanto para el ámbito académico como para los actores del mercado inmobiliario. Los hallazgos de este estudio pueden orientar las decisiones de inversión de los agentes privados y abrir nuevas líneas de investigación en el campo de la economía espacial y la valuación de bienes raíces.

TABLA DE CONTENIDOS

I. PLANTEAMIENTO DEL PROBLEMA	5
II. ANTECEDENTES, ESTADO ACTUAL DEL TEMA Y MARCO TEÓRICO	8
1. PERSPECTIVAS POST-POSITIVISTAS EN ECONOMÍA Y LA SINERGIA ENTRE MICROECONOMÍA Y ECONOMETRÍA.	8
2. INTRODUCCIÓN A LA VALUACIÓN DE PRECIOS DEL SECTOR INMOBILIARIO	11
3. INTRODUCCIÓN A LA VALUACIÓN DE “NON-MARKET VALUES”	16
4. EVOLUCIÓN DE LA TEORÍA DE PRECIOS HEDÓNICOS	20
5. ANTECEDENTES EN ARGENTINA	22
6. MARCO TEÓRICO	23
III. ESTIMACIÓN DE PRECIOS DE OFERTA EN EL SECTOR INMOBILIARIO	28
1. MUESTRA	29
2. EXTRACCIÓN DE DATOS Y CREACIÓN DE VARIABLES	36
3. ESPECIFICACIÓN DEL MODELO DE PRECIOS HEDÓNICOS	48
4. VERIFICACIÓN DE LAS HIPÓTESIS DE LA INVESTIGACIÓN.	59
IV. CONCLUSIONES	61
V. BIBLIOGRAFÍA	63
APÉNDICES	67

I. PLANTEAMIENTO DEL PROBLEMA

El sector inmobiliario tiene una gran relevancia económica y social en Argentina. Para las familias, la vivienda propia representa usualmente su activo más valioso y un componente primordial de su patrimonio y bienestar. Para las empresas, los espacios comerciales e industriales son factores clave en su producción y competitividad. El gobierno asimismo obtiene ingresos significativos de impuestos a las propiedades. Sin embargo, se conoce poco acerca de los determinantes del precio de los inmuebles y su importancia relativa, especialmente con respecto a atributos no transados en el mercado como la localización, accesibilidad y ambiente.

Por su parte, el mercado inmobiliario de C.A.B.A. se encuentra en 2023 en un punto de inflexión, previo a un cambio de gobierno que probablemente tendrá sus efectos sobre el mismo. Por otra parte, la elección de C.A.B.A. como ciudad objeto de análisis se fundamenta en el tamaño y dinamismo de su mercado inmobiliario. Asimismo, al ser la capital del país, cuenta con mejores fuentes de información y bases de datos más completas.

En este contexto, se consideró relevante analizar en profundidad los factores subyacentes en la formación de precios, para entender la dinámica del sector y es por lo que nuestra investigación buscó responder la siguiente pregunta:

Pregunta General : ¿Cuál es la relación que existe entre el precio de inmuebles en venta y sus atributos estructurales, ambientales, comunitarios y de accesibilidad en C.A.B.A. en diciembre de 2023?

De este interrogante, se desprendieron las siguientes preguntas específicas:

Pregunta Específica 1: ¿Cuál es la influencia de los atributos estructurales (metros cuadrados, número de ambientes, antigüedad, entre otros) en el precio de las propiedades?

Pregunta Específica 2: ¿En qué medida la cercanía a puntos de interés como escuelas, hospitales, parques y acceso al transporte público se relaciona con los precios de los inmuebles en venta?

Pregunta Específica 3: ¿Cuál es el impacto de aspectos comunitarios como la seguridad, calidad de vida y nivel cultural por comuna¹ sobre los precios de las propiedades en venta?

Pregunta Específica 4: ¿Cómo se vinculan aspectos ambientales como la calidad del aire y características escénicas con el precio de los inmuebles en venta?

Las preguntas de investigación planteadas buscaron examinar desde diversos ángulos la relación entre el precio de las propiedades en venta y sus atributos estructurales, ambientales, comunitarios y de

¹ Se presentan los conceptos de comunas y barrios de C.A.B.A., como así también mapas ilustrativos en el Capítulo III, subsección 1.

accesibilidad en el mercado inmobiliario de C.A.B.A. hacia fines de 2023. Con el propósito de abordar los distintos cuestionamientos surgidos en torno a esta temática, se estableció el siguiente objetivo general que engloba el propósito del estudio:

Objetivo General : Determinar la relación entre el precio de inmuebles en venta y sus atributos estructurales, geográficos, de accesibilidad y espaciales en C.A.B.A. en diciembre de 2023.

Asimismo, se desprendieron los siguientes objetivos específicos que dieron propósitos específicos al trabajo de investigación:

Objetivo Específico 1: Analizar la influencia de los atributos estructurales en el precio de las propiedades en venta.

Objetivo Específico 2: Estimar la relación entre los atributos de cercanía y de accesibilidad, y los precios de los inmuebles en venta.

Objetivo Específico 3: Determinar el impacto de atributos comunitarios sobre los precios de venta de propiedades.

Objetivo Específico 4: Evaluar la asociación entre atributos ambientales con el precio de venta de los inmuebles.

Nuestro trabajo de investigación se justificó por la escasez relativa de trabajos empíricos que hayan aplicado la metodología de precios hedónicos al mercado inmobiliario de Argentina. Si bien en este momento existen algunos estudios previos, como el de Domínguez Prost (2019) que utiliza precios hedónicos para la construcción de índices de precios inmobiliarios, son limitados en número y alcance de variables explicativas. Asimismo, los trabajos encontrados en Argentina se encuentran desactualizados. Dada la relevancia económica y social de la vivienda, resultó clave generar nueva evidencia sobre la formación de precios utilizando técnicas econométricas robustas.

Asimismo, el momento actual representa una oportunidad única para estudiar el mercado inmobiliario de C.A.B.A. Realizar un análisis en profundidad sobre los determinantes del precio brindará información oportuna y relevante tanto para hacedores de política económica como para participantes del sector.

Esta investigación fue viable dado que utilizó datos que se encuentran disponibles públicamente, no requiriendo acceso especial ni financiamiento. Además, la disponibilidad de grandes volúmenes de datos en plataformas digitales y portales de gobierno abierto posibilitó un análisis empírico robusto aplicando metodologías econométricas sólidas.

Se deriva de la pregunta general, el objetivo general y el marco teórico, la siguiente hipótesis sustantiva:

Hipótesis Sustantiva : Los atributos estructurales, ambientales, comunitarios y de accesibilidad de los inmuebles en C.A.B.A. en diciembre de 2023, tienen un efecto significativo y diferenciado sobre su precio de venta.

Esta hipótesis sustantiva buscó integrar la relación que plantea la pregunta general entre el precio de las propiedades y sus distintos atributos (estructurales, ambientales, comunitarios y de accesibilidad). Se propone que dichos atributos inciden significativamente en el precio, conformando un paquete de características favorables que en conjunto lo determinan.

Por otro lado, las hipótesis específicas a partir de las preguntas y objetivos específicos fueron las siguientes:

Hipótesis Específica 1: Existe una relación positiva entre los atributos estructurales metros cuadrados, cantidad de ambientes, número de habitaciones, número de baños, número de espacios de estacionamiento, piscina y jardín de las propiedades y su precio de venta, y una relación negativa entre el atributo estructural de antigüedad de las propiedades y su precio de venta.

Hipótesis Específica 2: La cercanía a puntos de interés como escuelas, hospitales, parques y transporte público se relaciona positivamente con el precio de los inmuebles en venta.

Hipótesis Específica 3: Mejor seguridad, calidad de vida, y nivel cultural en la comuna del inmueble tiene un efecto positivo sobre su precio de venta.

Hipótesis Específica 4: Existe una relación significativa entre los aspectos ambientales y escénicos del entorno en donde está ubicado el inmueble y su precio de venta.

II. ANTECEDENTES, ESTADO ACTUAL DEL TEMA Y MARCO TEÓRICO

Se llevó a cabo una revisión bibliográfica exhaustiva, la cual reviste una importancia fundamental para contextualizar la evolución y los aportes de la teoría de precios hedónicos en el ámbito del sector inmobiliario. Reconociendo la trascendencia de este análisis, se ha optado por un procedimiento que aborda los conceptos desde sus raíces más filosóficas y abstractas hasta sus manifestaciones más concretas y aplicadas. En este sentido, se exploraron las contribuciones de diversos autores y obras que han delineado el camino de esta teoría, desde los fundamentos epistemológicos del paradigma de la teoría del conocimiento hasta las formas funcionales específicas aplicadas en la valoración de bienes inmuebles.

La revisión bibliográfica se erige como un pilar indispensable para comprender el nacimiento y el desarrollo de la teoría de precios hedónicos. Permite rastrear la evolución de las ideas y los debates que han moldeado su conceptualización y aplicación en el ámbito inmobiliario. Al sumergirse en las obras y los aportes de diferentes autores, se pudo apreciar cómo los enfoques teóricos han ido evolucionando en respuesta a los desafíos y las demandas del análisis de mercado. Desde los trabajos pioneros que sentaron las bases conceptuales hasta las investigaciones contemporáneas que refinan las metodologías de estimación, esta revisión proporciona una panorámica completa de la teoría de precios hedónicos y su relevancia en el estudio de la oferta inmobiliaria en la Ciudad Autónoma de Buenos Aires ("C.A.B.A") a fines del año 2023.

En este capítulo se abordan conceptos fundamentales que preceden al análisis aplicado en esta tesis. Las subsecciones siguientes, como piezas de un rompecabezas, se unen para proporcionar un contexto previo esencial. Se presenta en primer lugar el paradigma de la teoría del conocimiento y la evolución de la Microeconomía y la Econometría, por una parte, la transformación de las metodologías de valuación en el mercado inmobiliario por otra, y la transformación de metodologías de valuación de valores de no mercado. Consiguientemente se da paso a un repaso acerca de la evolución de la teoría de precios hedónicos, y posteriormente a un repaso de ejemplos aplicados en Argentina.

Finalmente se concluye el capítulo con la presentación del marco teórico que da sustento a la investigación.

1. Perspectivas post-positivistas en Economía y la sinergia entre Microeconomía y Econometría.

El paradigma post-positivista en la ciencia representa una evolución crítica del paradigma positivista que dominó la investigación científica durante gran parte del siglo XX. Su surgimiento se encuentra en la crítica y reflexión sobre las limitaciones del positivismo, que proponía un enfoque científico basado en la observación objetiva y la verificación empírica de las teorías. A mediados del siglo XX, filósofos de la ciencia como Thomas Kuhn y Paul Feyerabend comenzaron a cuestionar la objetividad completa y la neutralidad del observador, subrayando la influencia de los contextos sociales y culturales en la construcción del

conocimiento científico. La historia del paradigma post-positivista está marcada por la transición de una visión determinista y reduccionista de la realidad a una comprensión más contextual y situacional, abrazando la complejidad inherente a los fenómenos estudiados.

Haciendo referencia a la importancia de la contextualización social, Kuhn (1968) sostiene que “si se toman al pie de la letra las restricciones positivistas sobre el alcance de la aplicabilidad legítima de una teoría, el mecanismo que indica a la comunidad científica qué problemas pueden conducir a un cambio fundamental debe dejar de funcionar” (p. 113).

Como ejemplo estrella, se encuentra el del entendimiento acerca del movimiento de los cuerpos: las dinámicas Newtonianas brillaban en el siglo XIX hasta la llegada de la teoría de la relatividad de Einstein; cuyo aporte terminó arrebatándole a las leyes de Newton la posición del estado actual del núcleo paradigmático. Este cambio, sin embargo, no podría haberse dado si no hubiese existido influencia alguna de las ideas de Newton en las ideas de Einstein, y tampoco si Einstein hubiese tomado a las leyes de Newton como universales.

Acerca de este mismo fenómeno, Sigman (2023) expone que “esto representa una revolución que, sin embargo, no modifica un ápice la forma en que los cuerpos se mueven. En cambio, el liberalismo, el marxismo, las teorías sobre la inflación o el supremacismo racial pueden alterar de manera decisiva los universos económicos y sociales que pretenden describir” (p.36).

Entrando en el ámbito económico, la adopción del paradigma post-positivista ha influido en la forma en que los economistas abordan el estudio de los fenómenos sociales y económicos. Se reconoce que los actores económicos no siempre se comportan de manera puramente racional y que las decisiones económicas están influidas por una variedad de factores, incluidos los valores culturales y las percepciones subjetivas. Este cambio de perspectiva ha llevado a un enfoque más amplio y multidisciplinario en la economía, donde se considera la interacción entre factores económicos, sociales y culturales en la formulación de teorías y políticas económicas.

Es importante recalcar que dentro de la perspectiva post-positivista en la economía se encuentra la economía positiva que es un enfoque que se centra en describir y explicar los fenómenos económicos tal como son, sin emitir juicios de valor. Su objetivo primordial es el análisis objetivo y la formulación de teorías basadas en observaciones y evidencias empíricas, entendiéndolas dentro de un contexto social. Los economistas positivos se esfuerzan por comprender y prever el comportamiento económico a través de modelos y teorías que se apoyan en datos concretos, desvinculando sus investigaciones de juicios éticos o normativos; los cuales sí son aplicados dentro de los enfoques de economía normativa.

Dentro de la economía positiva surge la Microeconomía como disciplina autónoma, que emergió a mediados del siglo XIX, como respuesta a la necesidad de comprender los comportamientos económicos a

nivel individual y de las firmas. En su esencia, busca analizar cómo toman decisiones individualmente los agentes económicos frente a la escasez de recursos.

Hasta ese entonces, la economía política clásica se enfocaba mayormente en cuestiones macroeconómicas, dejando de lado el estudio detallado de las decisiones y transacciones en el ámbito microeconómico.

La obra de Alfred Marshall, especialmente su influyente libro *Principles of Economics* publicado en 1890, es considerada un hito en el desarrollo de la Microeconomía al recopilar conocimientos previos y formalizar conceptos fundamentales como la oferta y la demanda, la utilidad y la teoría del consumidor, sentando las bases para el análisis de los mercados desde una perspectiva microscópica.

La teoría del consumidor, un pilar de la Microeconomía explora cómo los individuos maximizan la utilidad sujeta a restricciones presupuestarias, mientras que la teoría de la firma, por otro lado, se centra en cómo las empresas optimizan su producción y beneficios. Estos conceptos básicos forman la base teórica para entender las fuerzas que impulsan a la oferta y la demanda, así como la formación de precios en los mercados.

La teoría económica formula principalmente afirmaciones o hipótesis de naturaleza cuantitativa. Un ejemplo de esto es la afirmación de la teoría microeconómica que sostiene que, manteniendo constantes los demás factores, una disminución en el precio de un bien debería resultar en un aumento en la cantidad demandada de ese bien. En este sentido, la teoría establece una relación negativa o inversa entre el precio y la cantidad demandada de un bien. Sin embargo, la teoría en sí misma no proporciona valores numéricos que cuantifiquen esta relación; es decir, no especifica cuánto aumentará o disminuirá la cantidad demandada como consecuencia de un cambio particular en el precio del producto.

Como una respuesta a la necesidad de incorporar métodos cuantitativos y estadísticos en el estudio de los fenómenos económicos surge la Econometría en la primera mitad del siglo XX; siendo Ragnar Frisch y Jan Tinbergen pioneros en este campo. Su aparición se relaciona estrechamente con el desarrollo de la Estadística y la aplicación de técnicas matemáticas en la economía.

La Econometría busca identificar, modelar y cuantificar las relaciones causales entre variables económicas mediante el uso de modelos matemáticos y estadísticos. Agregando a lo anterior, “la Econometría da contenido empírico a la mayor parte de la teoría económica” (Gujarati y Porter, 2008, p. 1)

Su crecimiento se aceleró con el avance de la computación y la disponibilidad de grandes conjuntos de datos, permitiendo análisis más sofisticados y la formulación de modelos más precisos. Esta se ha convertido en una herramienta crucial para la validación empírica de teorías económicas y para la toma de decisiones basada en datos en diversos campos, desde la política económica hasta la investigación académica.

Por añadidura, la intersección entre la Microeconomía y la Econometría ha generado una sinergia valiosa, ya que la Microeconomía proporciona las bases teóricas para los modelos, mientras que la Econometría ofrece las herramientas para estimar y validar empíricamente dichos modelos. Esta colaboración ha llevado a avances significativos en nuestra comprensión de los fenómenos económicos a nivel micro y ha fortalecido la capacidad de la disciplina para abordar preguntas económicas de manera más precisa y rigurosa.

2. Introducción a la valuación de precios del sector inmobiliario

Existe un consenso en que los primeros conceptos relacionados a la idea del valor de la tierra y de los bienes inmuebles, tienen su raíz en la teoría del pensador clásico inglés David Ricardo. Ricardo desarrolló la teoría del valor de la tierra, también conocida como la teoría de la renta ricardiana.

En su libro *Principios de la Economía Política y Tributación* publicado en 1817, Ricardo postula que el valor de la tierra estaba determinado por su productividad diferencial. Argumentaba que a medida que la población crecía y se requería más tierra para la agricultura, se cultivaban tierras menos fértiles. La teoría propone que la renta diferencial surgía de la diferencia de productividad entre las tierras más fértiles y las menos fértiles. Esta teoría se centraba en la idea de que el valor de la tierra no estaba relacionado directamente con el trabajo invertido en ella como sostenían pensadores como John Locke, sino más bien con su productividad relativa.

Otro ejemplo surge a partir del alemán Johann Heinrich von Thünen, quién desarrolló un modelo para explicar cómo los precios agrícolas variaban con la distancia desde el mercado. Aunque se enfocó en el sector rural, sus ideas sentaron las bases para entender la variación de precios inmobiliarios en áreas urbanas.

El pensador presenta su modelo teórico de los “anillos concéntricos” de uso del suelo agrícola alrededor de una ciudad aislada. La teoría de von Thünen (1826) parte de suponer una ciudad aislada rodeada por tierras agrícolas homogéneas. Los agricultores buscan maximizar sus ganancias eligiendo los cultivos más rentables para la tierra que ocupan. Como el transporte hasta el mercado urbano tiene un costo, esto afectará qué cultivos son más rentables según la distancia. Así, las tierras más cercanas al mercado se dedicarán a productos perecederos y de alto valor como lácteos, verduras y flores. Conforme uno se aleja del mercado, aparecerán granos, madera y finalmente usos ganaderos extensivos. De esta forma, la renta de la tierra y su precio disminuirá con la distancia al centro urbano.

Tanto Ricardo como von Thünen convergen en la idea central de explicar el valor de la tierra a partir de su productividad sea por la fertilidad intrínseca de la tierra (Ricardo) o por su ubicación con respecto a los mercados (von Thünen). Estas variables eran cruciales para determinar la rentabilidad de esta, y por tanto en la formación del precio de esta en un contexto de paradigma clásico (donde los precios se formaban a partir del costo y la oferta, independientemente de la valoración de la sociedad).

La llegada de la Revolución Marginalista de la mano de los economistas Stanley Jevons, Carl Menger y León Walras representó un hito significativo en la teoría económica al introducir el concepto de utilidad marginal decreciente. En pocas palabras, el concepto de utilidad marginal decreciente busca explicar que el sentimiento de satisfacción que reporta cada unidad adicional consumida de un bien irá decreciendo a medida que un individuo consuma unidades adicionales del bien.

Adicionalmente, entendían que el precio de un bien se determina a partir de la interrelación de la oferta (que tiene en cuenta costos y productividades al estilo clásico) y la demanda (que toma la utilidad percibida por los consumidores).

Aunque ninguno de los tres economistas desarrolló su teoría en torno a la valoración de precios de tierras o viviendas, esta perspectiva transformó radicalmente la manera de entender la determinación de precios.

Las teorías de valuación de precios en el sector inmobiliario han evolucionado considerablemente en los últimos dos siglos. Los primeros enfoques surgieron a principios del siglo XX con autores como Richard M. Hurd, quien elaboró acerca del desarrollo de principios y métodos formales para la tasación de bienes raíces en Estados Unidos. Hurd publicó en 1904 *Principles of City Land Values*, donde expuso técnicas para la valuación masiva de áreas urbanas. Hirst (1924) muestra como Hurd propuso dividir ciudades en zonas homogéneas por usos e intensidad de desarrollo, para luego aplicar factores y tasas por metro cuadrado que permitieran estimar el valor total de cada zona.

Otro aporte significativo surgió de la mano del economista estadounidense Homer Hoyt, quien se convertiría en pionero en segmentación del mercado de bienes raíces para comprender la formación de precios.

Hoyt (1934) analizó el mercado inmobiliario en Chicago y descubrió que no existía un mercado inmobiliario único en la ciudad, sino múltiples sectores con sus propias dinámicas de oferta y demanda; es decir submercados dentro de Chicago. Así, los precios variaban significativamente entre áreas residenciales, comerciales e industriales. Dentro de cada segmento, los precios se movían en ciclos que no siempre coincidían. Por ejemplo, cuando subían los precios residenciales, no necesariamente ocurría lo mismo en locales comerciales. Cada sector tenía comportamientos diferenciados.

La gran contribución de Hoyt fue demostrar que no se podía modelar correctamente el mercado inmobiliario como uno solo, sino reconociendo la existencia de submercados con lógicas propias.

Sumergido en la sinergia de la Microeconomía en auge y las apariciones de la Econometría como ciencia, Frederick A. Babcock (1927) presenta una serie de métodos y procedimientos de tasación. Toma al ingreso como base de valor, y aborda temas como la tasación de diversos tipos de propiedades como residenciales y problemas de depreciación.

Presenta el enfoque de capitalización de rentas para estimar el valor de una propiedad como el flujo de ingresos futuros descontados a una tasa de capitalización, el cual sigue siendo uno de los métodos más utilizados en la actualidad.

La evolución de la teoría de valuación de precios de inmuebles constituye un proceso intrincado que se ha extendido a lo largo de varias décadas, involucrando a un conjunto significativo de académicos visionarios. Desde los fundamentos establecidos del contexto de economía clásica, los aportes microeconómicos y econométricos que sirvieron como piedra angular para la investigación contemporánea en valuación inmobiliaria, hasta las refinadas contribuciones teóricas a principios del siglo XXI, la disciplina ha experimentado una metamorfosis continua.

Se destaca la aparición de las obras de William Kinnard Jr. (1970) y sus aportes a los modelos econométricos para predecir rentas e ingresos.

Años más tarde, se destaca la publicación de la obra de Goldberg y Mark (1988); cuyo objetivo era converger la valuación de inmuebles con técnicas econométricas avanzadas. Los autores examinaron valuaciones masivas de viviendas unifamiliares. Sus análisis se centran en la determinación de las especificaciones de los modelos, abordando temas como la interacción de variables, no linealidad, estratificación y multicolinealidad.

A raíz de esta gran variedad de metodologías para la valuación de inmuebles y los grandes avances computacionales, el intento de integrar los tres enfoques existentes hasta ese entonces en un único modelo (de costos, de mercado y de ingresos²), se volvió más atractivo y plausible.

Los análisis de regresión hedónica, popularizada por Lancaster (1966) y desarrollada en profundidad por Rosen (1974a), han tomado relevancia ante la formulación de un nuevo enfoque para la teoría del consumidor. Este tipo de análisis busca descomponer el valor de un bien en base a sus características específicas. Dicho enfoque comenzó a ser aplicado con el objetivo de valorar activos en el sector inmobiliario.

En las últimas décadas, han proliferado nuevos modelos y conceptos como los modelos de equilibrio espacial impulsados por John Clapp (1980) que analizan la distribución de actividades económicas en un área geográfica, considerando factores como la localización de la producción, migración y comercio entre regiones, utilizando teoría de juegos y equilibrio general. Busca entender cómo las decisiones en una región impactan a otras, explicando la distribución óptima de recursos y actividades económicas en función de las interacciones geográficas.

Un ejemplo dentro del mundo de las finanzas es la teoría de cartera aplicada a la inversión inmobiliaria explicadas por Richard Brueggeman y Jeffrey Fisher (2022). En el contexto de inversiones inmobiliarias, la

² Se presenta al final de esta sección, un resumen con los enfoques utilizados en la valuación de precios inmobiliarios, como una categorización de los aportes históricos nombrados hasta el momento.

teoría de cartera implica la diversificación de la cartera de bienes raíces para maximizar los rendimientos y reducir el riesgo global de inversores inmobiliarios.

Adicionalmente, han cobrado fuerza los modelos automáticos de valuación masiva basados en aprendizaje automático por redes neuronales artificiales.

El trayecto histórico de la teoría de valuación sienta las bases para el análisis contemporáneo de la valuación de precios de inmuebles, donde convergen distintos enfoques metodológicos para abordar la complejidad de este campo. En la actualidad, se emplean varios enfoques en la práctica de la valuación inmobiliaria, destacándose el enfoque comparativo o de mercado, el enfoque de ingresos y el enfoque del costo, los modelos de precios hedónicos y utilización de redes neuronales artificiales. Cada uno aporta perspectivas innovadoras y desafíos metodológicos que enriquecen la comprensión contemporánea de la valuación de precios de inmuebles. Este amplio espectro de enfoques refleja la complejidad inherente al mercado inmobiliario y la necesidad de adaptarse a un entorno dinámico en constante cambio.

A modo de conclusión, se presenta un resumen categórico de los enfoques nombrados a continuación.

A. Resumen de enfoques en la valuación de bienes inmuebles

a. Enfoque comparativo o de mercado

El enfoque comparativo o de mercado, se fundamenta en la premisa de que el valor de un bien inmueble puede ser determinado mediante la comparación con transacciones similares en el mercado.

Como ejemplos se encuentran los aportes de Hurd (1904) y Hoyt (1934), quienes establecieron los cimientos al enfocarse en la importancia de la comparación de propiedades similares para derivar estimaciones precisas de valor.

Si bien este enfoque refleja transacciones reales del mercado, depende de la disponibilidad de comparables adecuados. Además, no está exenta de ajustes subjetivos como la influencia de factores psicológicos en las transacciones inmobiliarias. Shiller (2016) afirma que esta puede generar que los precios se desacoplen de los valores fundamentales y que sean impulsados por el comportamiento irracional y las expectativas excesivamente optimistas de los participantes del mercado, culminando, por ejemplo, en una crisis al estilo del de las hipotecas “subprime” de 2008 en Estados Unidos.

b. Enfoque de ingresos

El enfoque de ingresos se basa en la noción de que el valor de un inmueble puede ser determinado por los ingresos que puede generar, ya sea por su venta, por su renta o explotación agrícola; descontados a una tasa para determinar el valor en el presente.

Como se mencionó anteriormente, Ricardo y Von Thünen ya enfatizaban en la capacidad productiva de la tierra como factor determinante de su valor, en consonancia con el enfoque de capitalización de ingresos

dentro de la valuación inmobiliaria moderna. Ejemplos más contemporáneos son las contribuciones de Babcock (1927) y Kinnard (2003) que han consolidado el enfoque de ingresos como una herramienta para el valuador profesional.

Si bien valuar un inmueble desde el enfoque de ingresos permite capturar el potencial de rentabilidad, requiere proyecciones que pueden ser subjetivas e inciertas y es sensible a la tasa de capitalización elegida.

c. Enfoque del costo

El enfoque del costo se fundamenta en la premisa de que el valor de un inmueble está intrínsecamente relacionado con el costo de reemplazo o reproducción de la estructura y las mejoras asociadas.

Existen en la actualidad aportes destacados que han contribuido significativamente a este enfoque. Dentro de ellos se encuentra el de Marshall & Swift, cuyos estudios orientan hacia el uso de costos reales totales de construcción. Han sido fundamentales para establecer bases metodológicas contemporáneas del enfoque del costo en *Residential Cost Handbook* donde han proporcionado directrices y herramientas para evaluar el valor de los componentes físicos de un inmueble. Otro ejemplo a nivel nacional es el de suplemento de arquitectura del diario Clarín "ARQ", que compila datos de costos de cada una de las variables utilizadas para la construcción de inmuebles. De esta manera consiguen calcular los costos por metro cuadrado de distintos modelos de viviendas.

El enfoque de costo tiene como primera limitación que basa el valor de una propiedad únicamente en su costo de construcción o reemplazo, y la precisión de la estimación del valor dependerá intrínsecamente la calidad de las estimaciones de los costos de construcción.

d. Modelos de precios hedónicos³

Los modelos de precios hedónicos representan una perspectiva innovadora en la valuación de precios de inmuebles, centrándose en descomponer el valor en características específicas. Como se nombró anteriormente, Lancaster y Rosen han contribuido en esta área demostrando cómo atributos específicos de una propiedad, como su ubicación o características estructurales, afectan de manera diferenciada su valor en el mercado.

Este enfoque ha evolucionado con el tiempo, incorporando avances estadísticos y econométricos para modelar de manera más precisa la relación entre las características de una propiedad y su precio en el mercado. Si bien permiten cuantificar la contribución de cada característica, requieren gran cantidad de datos y conocimientos econométricos avanzados como los presentados anteriormente por Mark y Goldberg (1988).

³ Esta teoría es exhaustivamente desarrollada en la sección de marco teórico, dado que es aplicada para la consecución de los objetivos de este trabajo.

e. Redes neuronales artificiales

Las redes neuronales artificiales (“RNA”) son sistemas computacionales inspirados en las redes biológicas del cerebro y las neuronas que constituyen el sistema nervioso. Estas redes están diseñadas para realizar tareas mediante el aprendizaje a partir de ejemplos dados, sin necesidad de seguir reglas específicas o tener conocimiento previo de las tareas. Nkolika et al (2020) explican que las RNA generan resultados a partir de los ejemplos que procesan, y son miembros de los métodos computacionales evolutivos.

Las RNA aprenden de grandes conjuntos de datos para reconocer patrones ocultos que luego aplican para predecir valores de nuevos inmuebles. Si bien superan limitaciones de modelos tradicionales y se adaptan a mercados dinámicos, los parámetros que se obtienen son más difíciles de interpretar.

Es importante recalcar que los distintos enfoques de valuación inmobiliaria no son mutuamente excluyentes, sino que pueden complementarse entre sí para obtener una estimación más precisa del valor de una propiedad. Utilizados de manera aislada, cada enfoque tiene limitaciones importantes, tal como se ha discutido previamente. Sin embargo, al combinarlos se puede compensar las debilidades de cada uno y también confirmar la validez de la valuación final mediante la convergencia de diferentes métodos.

Por ejemplo, el valor obtenido mediante el enfoque de ingresos puede contrastarse con evidencia de transacciones comparables recientes para el enfoque de mercado. Asimismo, el costo de reposición brinda un límite inferior o piso para el valor. Inclusive los modelos estadísticos avanzados requieren después calibrarse con datos empíricos del mercado.

3. Introducción a la valuación de “non-market values”

En paralelo, la Microeconomía evolucionó para abordar una variedad de fenómenos, desde la teoría de juegos, la economía del comportamiento, las externalidades y los bienes públicos; muchos de los cuales implican fallos del mercado originados en la institución orientadora de la asignación de recursos.

Históricamente, el paradigma predominante se centraba exclusivamente en la medición del valor mediante transacciones comerciales, descuidando aspectos cruciales de la vida humana que no estaban directamente vinculados al intercambio comercial. No obstante, en el siglo XX, surgieron preocupaciones ambientales, sociales y culturales que llevaron a economistas y expertos en políticas a reconocer la necesidad de integrar los “non-market values” o valores no de mercado en la toma de decisiones económicas y políticas.

Los “non-market values” son aquellos aspectos de la vida y del entorno que no están directamente asociados con transacciones comerciales y que no son fácilmente cuantificables en términos monetarios. Estos valores abarcan una amplia gama de dimensiones, incluyendo aspectos ambientales, culturales y sociales que contribuyen al bienestar de la sociedad.

Es de las motivaciones nombradas anteriormente que surge la valuación económica de valores de no mercado. Marciano & Ramello (2020) la entienden como un conjunto de técnicas que busca reflejar el valor económico de los cambios, en la calidad y cantidad, de bienes y servicios que el mercado no alcanza a exponer como efectivos determinantes de sus señales de precios.

La literatura proporciona numerosos casos que ilustran cómo se han valuado estos aspectos no comerciales en diversos contextos. Se presenta una sección con los ejemplos relevantes a continuación.

A. Ejemplos de valuaciones de non-market values

a. Valuaciones a la belleza escénica y recreación en espacios naturales.

Krutilla y Fisher (1975) introdujeron el método del costo de viaje para valorar espacios naturales con fines recreativos. Específicamente, presentaron un marco conceptual y empírico para estimar los beneficios recreativos de espacios naturales como parques, reservas y áreas silvestres. Su aproximación se basa en utilizar la información sobre los costos en que incurrir las personas para visitar un sitio natural (por ejemplo, costos de viaje, cuotas de entrada, tiempo) como un medio para inferir el valor económico que le asignan al acceso a dicho sitio.

A su vez, Bateman y Langford aplicaron la valoración contingente a parques nacionales con fines recreativos. Específicamente, aplicaron encuestas para estimar la disposición a pagar de distintos individuos por la conservación del parque nacional de Lake District en Reino Unido. Bateman y Langford (1997) señalan: "se pidió a los encuestados que imaginaran un escenario hipotético donde el parque dejaría de existir y luego se les preguntó su disposición a pagar ("DAP") para evitar dicha pérdida" (p. 574). Mediante un cuidadoso diseño de encuesta y riguroso análisis econométrico, los autores encontraron una DAP positiva y estadísticamente significativa incluso para no usuarios del parque, resaltando la importancia de los valores de no uso. Este estudio sentó un precedente metodológico importante para la valoración contingente.

b. Valuaciones de la calidad del aire y del agua.

Carson y Mitchell (1993) valoraron mejoras en la calidad del agua. Los autores llevaron a cabo una estimación de los beneficios totales de cumplir con los objetivos de la Ley del Agua Limpia en Estados Unidos, utilizando una función de valoración que considera la DAP en función de la calidad del agua, ingresos y otras variables. Las implicaciones clave del estudio indican que los beneficios y costos de los esfuerzos para controlar la contaminación del agua son aproximadamente iguales, y muchas nuevas acciones políticas necesarias para mejorar la calidad del agua al menos un nivel de calidad adecuado para nadar, pueden no generar beneficios netos positivos.

Por su lado, Alberini y Krupnick (2000) estimaron los beneficios de una mejor calidad del aire. Los autores emplearon tanto el método de costo de enfermedad ("CDE") como la valuación contingente para estimar los beneficios asociados a la reducción de la morbilidad infantil causada por la contaminación

atmosférica en Taiwán. A través del método de CDE, calcularon los costos directos e indirectos vinculados al aumento de casos de bronquitis y asma infantil atribuibles a la contaminación. Luego, mediante una encuesta de valoración contingente, determinaron la DAP de los padres por disminuir la incidencia de estas enfermedades. Como resultado, la DAP era mayor que las estimaciones de CDE, superándolas de 1.61 a 2.26 veces, dependiendo de los niveles de contaminación.

c. Valuaciones de riesgos a la salud

En salud pública, existen trabajos, como el de Chasin, Tolley, Kenkel et al (1995), que proponen el método de capital humano para valorar vidas. Los autores presentan en detalle la fundamentación teórica del método de capital humano, el cual estima el valor de la vida a partir del valor presente de los ingresos futuros que se perderían debido a una muerte prematura. Por ejemplo, ajustar las estimaciones basadas en ingresos promedio según diferencias individuales.

Asimismo, discuten críticamente los supuestos y limitaciones del método, como la exclusión de valores intangibles de la vida, y estrategias para su aplicación en evaluaciones de políticas de salud ambiental y ocupacional.

d. Valuaciones del patrimonio cultural e histórico

Existen trabajos que abordan los beneficios de la conservación del patrimonio cultural. Pollicino y Maddison (2001) presentan un estudio de caso aplicando el método de valoración contingente a la Catedral de Lincoln, estimando la DAP por su limpieza y restauración. Navrud y Ready (2003) analizan en profundidad la aplicabilidad de métodos como valoración contingente y costo de viaje para valorar sitios históricos y artefactos.

Mientras Pollicino y Maddison se enfocan en un caso de estudio específico, Navrud y Ready presentan una discusión integral sobre técnicas de valuación para el patrimonio cultural, incluyendo distintos tipos de sitios y objetos históricos. En sus conclusiones, ambos trabajos destacan la utilidad de la valoración económica rigurosa para fundamentar decisiones sobre conservación del patrimonio cultural tangible, evidenciando su viabilidad y relevancia a través de sus respectivos enfoques teórico-metodológico y empírico.

e. Valuaciones de servicios ecosistémicos.

En cuanto a los servicios ecosistémicos, Constanza et al. (1997) estimaron el valor global de los mismos. Utilizando una variedad de fuentes de datos y métodos de valoración, realizaron la primera estimación comprehensiva del valor anual de los servicios generados por los ecosistemas del planeta. Estos incluyen servicios como polinización, control de erosión, hábitat, recursos genéticos y control climático, entre otros. Los autores estimaron que el valor total agregado de los servicios ecosistémicos a nivel global era de alrededor de 33 billones de dólares anuales. Esta cifra equivalía aproximadamente al Producto Interno Bruto ("PIB") mundial en ese entonces.

En definitiva, se ha transitado desde visiones meramente especulativas a principios del siglo XX a un corpus teórico sólido sobre la valoración no de mercado.

Rose (1999) publica el informe *Non-Market Valuation Techniques: The State of the Art* donde proporciona una visión general de las principales técnicas de valoración no monetaria utilizadas para estimar valores económicos de bienes ambientales intangibles. Las técnicas discutidas incluyen la valoración contingente, los experimentos de elección, el método de costos de viaje y el de precios hedónicos.

A modo de conclusión de la sección, se presentan a continuación descripciones de cada una de las técnicas nombradas.

B. Resumen de enfoques en la valuación de non-market values

a. Valoración contingente

Esta técnica se basa en encuestas acerca de cuánto estarían dispuestos a pagar los encuestados por un bien ambiental intangible, o cuánto estarían dispuestos a aceptar para renunciar a él. La valoración contingente es útil para estimar valores de bienes ambientales que no tienen un mercado establecido, pero puede ser susceptible a problemas de sesgo de respuesta y puede ser difícil de aplicar en países en desarrollo.

Si bien ha tomado mayor participación en la valuación de non-market values, este enfoque no está exento de limitaciones. Bateman y Langford (1997) realizan una crítica constructiva al método de valoración contingente, la cual se centra en los siguientes puntos: sesgos hipotéticos, información imperfecta, sesgo estratégico, formatos de preguntas, exclusión de no uso y de agregación.

b. Experimentos de elección

Esta técnica implica presentar a los encuestados diferentes escenarios y pedirles que elijan su opción preferida. Los experimentos de elección son útiles para estimar valores de bienes ambientales intangibles y pueden proporcionar información sobre las preferencias de los encuestados, pero pueden ser costosos y difíciles de diseñar.

c. Método de costos de viaje

Esta técnica se basa en la observación de los costos de viaje de los visitantes a un sitio ambiental y la inferencia de su valoración del sitio. Como se nombró anteriormente, Krutilla y Fisher (1975) utilizan el método de costo de viaje para estimar el valor económico de áreas naturales con fines recreativos, a partir de los gastos en que incurren las personas para visitar dichos sitios.

El método de costos de viaje es útil para estimar valores de bienes ambientales que tienen un mercado establecido, pero puede ser difícil de aplicar en áreas remotas o donde los visitantes son pocos.

d. Modelos de precios hedónicos

Esta técnica se basa en la observación de los precios de mercado de bienes inmuebles y la inferencia de la valoración de los atributos ambientales del área. El precio hedónico es útil para estimar valores de bienes ambientales que tienen un mercado establecido, pero puede ser sensible a factores no ambientales que afectan los precios de mercado.

Si bien no son desarrollados en el informe de Rose (1999), en la literatura también se encuentran enfoques como el de costos de enfermedad aplicados por Alberini y Krupnick (2000) y la utilización de más de un enfoque para la valuación de un mismo non-market value.

4. Evolución de la teoría de precios hedónicos

La estimación de valores implícitos a partir de modelos de precios hedónicos tiene sus orígenes a finales de la década de 1930, cuando se empezaron a estudiar técnicas econométricas para analizar los determinantes del precio de automóviles usados en función de sus atributos. Desde entonces, se ha desarrollado una amplia literatura que cubre desde los aspectos teóricos más fundamentales hasta aplicaciones prácticas en una gran variedad de mercados, con especial énfasis en el sector inmobiliario.

El primer trabajo formal sobre la teoría de precios hedónicos fue el artículo "Hedonic Price Indexes with Automotive Examples" de A.T. Court, publicado en el libro *The Dynamics of Automobile Demand* del fabricante General Motors. En él, se presentan desarrollos formales sobre la metodología de índices de precios hedónicos y su aplicación para estimar valores implícitos de las características de automóviles.

Court (1939) argumenta que, al comparar los precios de automóviles en diferentes períodos de tiempo, es necesario ajustar dichas comparaciones a partir de los cambios en las cualidades y atributos de los modelos. Propone construir índices de precios hedónicos que descompongan el precio total observado en el mercado en los valores marginales de elementos como potencia del motor, tamaño, sistema de frenos, compartimentos internos, entre otras.

Años posteriores y dentro del mismo rubro, Griliches construyó índices de precios hedónicos para automóviles nuevos en los Estados Unidos entre 1950 y 1960. Griliches (1961) aplicó regresiones para estimar los valores marginales implícitos de atributos como peso, potencia, dimensiones y características mecánicas. Luego usó estos coeficientes para derivar índices de precio puro ajustados por la calidad y composición de los modelos. Sus resultados mostraban que los índices oficiales sobreestimaban significativamente el crecimiento real de los precios al no tener en cuenta mejoras en la calidad⁴.

Posteriormente, en la década de 1960 Kelvin Lancaster desarrolló el denominado "Enfoque de las Características". Lancaster (1966) argumentó que los bienes pueden ser descritos en términos de sus

⁴ La crítica a la falta de consenso y regularidad en la elección de variables para el sector automovilístico en Estados Unidos también fue hecha por Court (1939).

características, y que la utilidad se deriva de la posesión de esos atributos. Este enfoque, también conocido como la teoría de la calidad, sugiere que los consumidores valoran las características específicas de un producto y que su elección se basa en la combinación óptima de estas características, en lugar de simplemente la cantidad de bienes consumidos. Este concepto fue fundamental para el desarrollo posterior de la teoría de precios hedónicos, ya que estableció la base para entender cómo las características de un bien influyen en su precio y cómo los consumidores valoran estas características al tomar decisiones de compra.

Ya adentrada la teoría en el ámbito inmobiliario, Ridker y Henning (1967) aplicaron análisis de precios hedónicos para estimar el impacto de la contaminación del aire en los valores de propiedades residenciales en St. Louis, Missouri. Utilizaron datos de precios de venta, características de las viviendas y niveles de contaminación por SO₂ y partículas suspendidas en distintas zonas. Estimaron modelos econométricos para derivar el descuento implícito en el precio asociado a mayores niveles de polución. Sus resultados mostraban una relación negativa estadísticamente significativa entre contaminación del aire y precios de propiedades. Esto implicaba que los compradores estaban dispuestos a pagar una prima en el precio para residir en áreas con mejor calidad del aire.

Uno de los trabajos seminales en la teoría de precios hedónicos fue el artículo de Sherwin Rosen en 1974, *Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition*, publicado en el *Journal of Political Economy*. Rosen (1974) presentó un modelo teórico formal para estimar equilibrios de mercado y precios hedónicos implícitos. Introdujo el concepto de "precios implícitos de equilibrio" y derivó condiciones de optimalidad.

Joseph Gyourko y Joseph Tracy (1991) aplicaron funciones de precios hedónicos para estimar la calidad de vida en diferentes ciudades de Estados Unidos. Utilizando datos de precios de vivienda, características estructurales, ingresos y variables fiscales locales, estimaron modelos hedónicos para derivar valores implícitos de atributos del clima como temperatura, humedad, viento y precipitación.

Sus resultados indicaban que las ciudades con mejores condiciones climáticas tenían sustanciales primas implícitas reflejadas en mayores precios de vivienda. Por ejemplo, estimaron que los hogares estaban dispuestos a pagar cientos de dólares más por evitar días extremadamente fríos o calurosos.

Otros autores destacados que han hecho contribuciones a la metodología de precios hedónicos son:

Can y Megbolugbe (1997) analizan el impacto de la dependencia espacial en la construcción de índices de precios de vivienda mediante regresiones hedónicas. Los autores argumentan que las propiedades inmobiliarias próximas tienden a compartir características no observables que se reflejan en una correlación positiva de sus precios, incluso cuando se mantienen constantes otros atributos estructurales y de localización. Este fenómeno de autocorrelación espacial puede sesgar los coeficientes y errores estándar en modelos hedónicos tradicionales. Para abordar este problema, Can y Megbolugbe incorporan pesos

espaciales en las variables independientes usando una matriz de contigüidad espacial. Los resultados empíricos con datos de Singapur muestran que el modelo con autocorrelación espacial logra un mejor ajuste y modifica sustancialmente los coeficientes de las características.

Por otro lado, se encuentra Malpezzi que aplicó modelos hedónicos para estudiar el impacto de subsidios gubernamentales en el mercado de vivienda de bajos ingresos en Estados Unidos (2008). En sus estudios, Malpezzi estimó elasticidades precio de la demanda y simuló los efectos de cambios en los programas habitacionales. Sus contribuciones enfatizan la utilidad de los modelos hedónicos para el análisis de política pública.

Ekeland, Heckman y Nesheim (2004), quienes derivaron un nuevo estimador hedónico basado en emparejamiento que no requiere especificar a priori ninguna forma funcional. El estimador es semiparamétrico y utiliza información sobre los precios de las transacciones y las características de las propiedades para construir contrafactuales, sin depender de supuestos restrictivos.

Bajari y Benkard (2005), aplicaron métodos Bayesianos semiparamétricos para modelar precios hedónicos de automóviles. Argumentaron que los enfoques tradicionales parametrizan de forma rígida los atributos y sus interacciones. En cambio, demostraron que los modelos Bayesianos permiten mayor flexibilidad al inferir la forma funcional a partir de los datos.

Bishop y Timmins (2019), usaron regresiones geográficamente ponderadas para capturar variaciones espaciales en los precios hedónicos de propiedades y sus atributos. Este enfoque permite estimar variaciones locales en las primas marginales implícitas, relajando el supuesto de coeficientes globales constantes. Sus resultados apoyan la presencia de heterogeneidad espacial en los equilibrios de precios hedónicos.

En décadas recientes, los modelos econométricos para estimar precios hedónicos se han vuelto crecientemente sofisticados. Se han propuesto desde especificaciones paramétricas tradicionales hasta técnicas no paramétricas, semiparamétricas, geoestadísticas y basadas en machine learning. Sin duda, esta evolución se ha visto facilitada por los avances en capacidad computacional.

5. Antecedentes en Argentina

En la revisión de literatura realizada para esta investigación, se ha encontrado una notable escasez de estudios enfocados en el análisis del mercado inmobiliario y la construcción de índices de precios para C.A.B.A. A pesar de la importancia del sector inmobiliario en la economía local, no se han identificado antecedentes o trabajos previos que hayan abordado de manera específica este objetivo.

Si bien existen algunos estudios que han explorado aspectos relacionados con el mercado inmobiliario en CABA, como los realizados por Cruces et al. (2008), Ham González (2009) y Zoloa (2015), ninguno de ellos ha tenido como objetivo principal la elaboración de un análisis hedónico en esta región. Domínguez Prost (2019) señala que, a diferencia de otros países donde se han desarrollado metodologías e índices de

precios inmobiliarios a nivel nacional, en Argentina existe una falta de herramientas para monitorear y hacer proyecciones sobre la dinámica de precios en el mercado de bienes raíces, tanto a nivel local como nacional.

Prost utiliza técnicas de precios hedónicos y su aporte principal se centra en desarrollar herramientas para el seguimiento de la evolución de precios en el mercado inmobiliario de CABA, más que en analizar el impacto individual de variables explicativas. Aprovechando la información geográfica disponible para cada departamento, Prost construye variables de distancia a puntos de interés y estima índices de precios corrigiendo por dependencia espacial. Su estudio no considera factores ambientales o de criminalidad que también podrían afectar los precios de las propiedades.

Siguiendo varios aspectos de la metodología utilizada por Prost, se buscó desarrollar un modelo de precios hedónicos que, a diferencia del objetivo perseguido por el autor, intentó desvelar los atributos determinantes de los precios de oferta, incluyendo los factores mencionados en el párrafo anterior, para un momento específico en el tiempo

6. Marco teórico

El marco expuesto a continuación delinea los principales constructos teóricos y enfoques metodológicos que guiaron el análisis del problema planteado en la presente investigación. En este sentido, se abordaron perspectivas fundamentales provenientes de campos tan diversos como la economía, la microeconomía, la econometría y la valuación de precios de mercado y no mercado; estableciendo vínculos y convergencias entre estas áreas del conocimiento para obtener una comprensión integral del fenómeno en estudio.

La palabra "hedónico" proviene del griego "hedonikos", que significa placer. En el contexto económico, este concepto se refiere a la utilidad o satisfacción derivada del consumo de un bien o servicio. De acuerdo con Rosen (1974), los precios hedónicos buscan estimar la contribución de cada característica al precio total observado en el mercado.

El supuesto central en los modelos hedónicos es que los bienes son valorados por sus características y no como entidades homogéneas. Como se mencionó anteriormente, es posible descomponer el precio de un bien para obtener valores implícitos de sus distintos atributos.

Comprendiendo la esencia conceptual de los precios hedónicos y su propósito teórico de descomponer el valor de un bien en sus características individuales, el supuesto implica que existe una función que permite modelar y estimar la contribución de cada atributo específico al precio total observado en el mercado. Se presenta a continuación, la ecuación genérica de un modelo de precios hedónicos.

Ecuación 1: Forma funcional genérica de un modelo de precios hedónicos

$$P = f (X_1; X_2; \dots; X_n)$$

Donde el precio de un bien (P) se modela como una función (f) de las n distintas características relevantes de dicho bien ($X_1; X_2; \dots; X_n$).

Los modelos de precios hedónicos tienen una amplia aplicación en la valuación de bienes para los cuales no existen mercados establecidos. Según Freeman (1979) son particularmente útiles para estimar el valor de características que no se comercializan de forma independiente, como por ejemplo la calidad ambiental o el prestigio de un vecindario.

Esta flexibilidad para incorporar una gran variedad de atributos, tanto cuantitativos como cualitativos, hace que los modelos de precios hedónicos sean muy valiosos para analizar mercados tan heterogéneos como el de bienes raíces. Permiten capturar en el precio final los diversos factores que los consumidores valoran al tomar sus decisiones de compra.

En las siguientes subsecciones se aborda con más detalle la evolución de la teoría de precios hedónicos, las distintas formas funcionales propuestas, y los tipos de variables tanto de mercado como de no mercado que se suelen incluir en estos modelos econométricos.

a. Formas funcionales

La elección de la forma funcional apropiada es una decisión crucial en la estimación de modelos de precios hedónicos, ya que determina la manera en que se relacionan matemáticamente los atributos con el precio. A lo largo de las décadas de investigación en valuaciones hedónicas, se han propuesto y aplicado una gran variedad de formas funcionales, tanto paramétricas como no paramétricas.

- Formas funcionales lineales

Las formas funcionales lineales son de las más simples y han sido ampliamente utilizadas desde los primeros trabajos en precios hedónicos debido a su fácil interpretación. La Ecuación 2 presenta la forma funcional lineal.

Ecuación 2: Forma funcional lineal

$$P = \beta_1 * X_1 + \beta_2 * X_2 + \dots + \beta_n * X_n$$

Donde los parámetros β_i representan las contribuciones marginales de cada variable explicativa al precio del inmueble (P).

Aunque su simplicidad facilita la estimación e interpretación, lo que la ha llevado a ser ampliamente utilizada en aplicaciones sencillas de modelos hedónicos, conllevan supuestos restrictivos de linealidad y aditividad que pocas veces se cumplen en la realidad. En particular, asume que los atributos contribuyen al precio de forma independiente y con primas constantes a lo largo de todo su rango de valores, además, no permite interacciones entre variables ni efectos no lineales.

- Formas funcionales logarítmicas y semi-logarítmicas

Las formas logarítmicas y semilogarítmicas se popularizaron porque permiten interpretar a los coeficientes como elasticidades.

A continuación se desglosan breves explicaciones de las ecuaciones de las formas funcionales logarítmicas y semilogarítmicas.

Ecuación 3: Forma funcional logarítmica (log-log)

$$\log(P) = \beta_1 * \log(X_1) + \beta_2 * \log(X_2) + \dots + \beta_n * \log(X_n)$$

Tanto el precio del inmueble (P) como los atributos explicativos se encuentran en forma logarítmica. Esto permite interpretar los coeficientes estimados β_i como elasticidades, lo que significa que un cambio porcentual en una variable explicativa (por ejemplo, un aumento del 1% en el área del terreno) se asocia con un cambio porcentual en el precio del inmueble igual al valor del coeficiente respectivo. La forma log-log ha sido ampliamente usada por su flexibilidad en estudios como el de Can y Megbolugbe (1997).

Ecuación 4: Forma funcional semilogarítmica (lin-log)

$$P = \beta_1 * \log(X_1) + \beta_2 * \log(X_2) + \dots + \beta_n * \log(X_n)$$

En esta especificación, el precio del inmueble (P) se encuentra en su escala original, mientras que los atributos explicativos están en forma logarítmica. Los coeficientes estimados β_i se interpretan como semielasticidades, lo que significa que un cambio porcentual en una variable explicativa (por ejemplo, un aumento del 1% en el área del terreno) se asocia con un cambio en el precio del inmueble igual al valor del coeficiente respectivo multiplicado por el precio promedio.

Ecuación 5: Forma funcional semilogarítmica (log-lin)

$$\log(P) = \beta_1 * X_1 + \beta_2 * X_2 + \dots + \beta_n * X_n$$

En este caso, el precio del inmueble (P) se encuentra en forma logarítmica, mientras que los atributos explicativos β_i se mantienen en su escala original. Los coeficientes estimados β_i se interpretan como semielasticidades, lo que significa que un cambio unitario en una variable explicativa (por ejemplo, un aumento de una habitación) se asocia con un cambio porcentual en el precio del inmueble igual al valor del coeficiente respectivo multiplicado por 100.

La elección entre estas formas funcionales dependerá de la naturaleza de los datos, los supuestos sobre la relación entre las variables y los objetivos específicos del análisis.

- Formas funcionales “spline”

Las formas funcionales “spline” son más flexibles que las anteriores, ya que ajustan polinomios por tramos para capturar relaciones no lineales complejas. El spline ajusta diferentes polinomios en distintos

rangos de los predictores, lo que permite representar relaciones como convexidad, umbrales e interacciones. Se definen nodos donde ocurren los quiebres o cambios en la relación. Las ventajas incluyen gran adaptabilidad y evitar sobreestimación. La desventaja es que se debe determinar apropiadamente la ubicación de los nodos con base en conocimiento del mercado.

- **Modelos funcionales geoestadísticos**

Los modelos geoestadísticos como el modelo de semivariograma usado por Bishop y Timmins (2019) permiten incorporar dependencia y autocorrelación espacial. Por ejemplo, estiman los precios hedónicos mediante una combinación lineal de los valores observados en ubicaciones cercanas. Capturan el hecho de que propiedades próximas tienden a tener características y precios correlacionados espacialmente. Estas técnicas son útiles cuando se espera una variación sistemática y suave a lo largo del espacio en las variables y sus interrelaciones. Requieren definir una función de covarianza espacial y un vecindario apropiado. Una desventaja es el aumento en la complejidad computacional.

- **Otras formas funcionales**

Otras formas funcionales avanzadas son las basadas en series de Fourier, wavelets, núcleos (kernels), modelos aditivos generalizados, y métodos no paramétricos como redes neuronales y Random Forest. Buscan mayor flexibilidad para adaptarse a relaciones complejas y no lineales, sin necesidad de imponer una forma funcional paramétrica limitada. Por ejemplo, Bajari y Benkard (2005) aplicaron métodos bayesianos no paramétricos para el mercado de automóviles. La desventaja es que pueden ser más difíciles de interpretar y requerir mucho entrenamiento y ajuste de parámetros.

En resumen, la elección de la forma funcional apropiada debe basarse en el conocimiento del mercado en estudio y de la naturaleza de las variables disponibles. Las formas flexibles permiten capturar mejor las complejidades de la formación de precios, pero deben balancearse con la interpretabilidad. Por lo tanto, en la presente investigación, se tuvo en cuenta formas funcionales lineales, logarítmicas y se analizó cuál fue la más apropiada para la consecución de los objetivos.

En el siguiente apartado se profundizará en los distintos tipos de variables que típicamente se incluyen como determinantes en los modelos de precios hedónicos aplicados al sector inmobiliario.

b. Variables relevantes para valuaciones del mercado inmobiliario

Los modelos de precios hedónicos buscan explicar el precio de los bienes inmuebles en función de sus múltiples características o atributos. Décadas de investigación han identificado y medido una amplia gama de variables que capturan valores implícitos significativos para los compradores.

Las variables de mercado son aquellas que se transan de forma observable en transacciones inmobiliarias. Representan atributos tangibles que tienen un impacto directo en los precios.

Dentro de las variables estructurales se incluyen características físicas como superficie, número de habitaciones y baños, materiales de construcción, antigüedad, estado de conservación, entre otras.

Las variables de localización buscan medir las características derivadas de la ubicación de la propiedad, como cercanía a centros de empleo, acceso al transporte público, distancia a comercios y servicios; como lo planteado por Can y Megbolugbe (1997).

Finalmente, las variables de vecindario capturan características del entorno inmediato como nivel socioeconómico, tasa de criminalidad, calidad de los servicios públicos, acceso a áreas verdes e instituciones educativas. Autores como Bishop y Timmins (2019) han estimado sus primas implícitas.

Las variables de no mercado representan atributos intangibles que no se transan directamente pero que los compradores valoran y se reflejan en los precios.

Dentro de las características ambientales se encuentran variables como calidad del aire, niveles de ruido, vista panorámica y cercanía a áreas naturales. Estudios pioneros como el de Ridker y Henning (1967) buscaron capturar el valor de la contaminación ambiental mediante precios hedónicos.

Las externalidades sociales se refieren a variables como composición étnica, densidad poblacional, segregación residencial, inequidad económica y otros factores sociodemográficos que afectan las preferencias de ubicación. Kiel y Zabel (1996) analizaron su impacto en el mercado inmobiliario de EE. UU.

Finalmente, los atributos psicológicos buscan representar factores subjetivos como estatus social, preferencias individuales, seguridad percibida y aversión al riesgo. La importancia de incorporar estas variables menos tangibles ha sido estudiada por autores como Farber y Yeates (2006).

En síntesis, los modelos hedónicos contemporáneos buscan capturar la más amplia gama de determinantes que subyacen a la formación de preferencias y precios en el complejo mercado inmobiliario.

III. ESTIMACIÓN DE PRECIOS DE OFERTA EN EL SECTOR INMOBILIARIO

El trabajo se realizó siguiendo un enfoque cuantitativo. Hernández Sampieri (2014) define al enfoque cuantitativo como un conjunto de procesos secuenciales y probatorios, en el que cada etapa precede a la siguiente. Siguiendo este enfoque, las hipótesis mencionadas anteriormente implican variables desarrolladas en determinado contexto. El objetivo final fue establecer conclusiones mediante la prueba de dichas hipótesis, a través de la medición y el tratamiento estadístico de las variables que involucran.

Este estudio tiene un alcance correlacional ya que su objetivo general plantea determinar la relación o asociación entre el precio de los inmuebles en venta (variable dependiente) y sus distintos atributos (variables explicativas). Las preguntas de investigación también se orientaron a medir el grado de asociación entre dichas variables, utilizando términos como "influencia" y "vinculación". Asimismo, el énfasis del estudio se centra en cuantificar los efectos de ciertos predictores (atributos estructurales, de accesibilidad, etc.) sobre el precio, lo cual es propio de un abordaje cuantitativo correlacional.

La unidad de análisis fueron casas y departamentos en venta publicados en Mercado Libre en C.A.B.A. en diciembre de 2023. Es un estudio de carácter transversal, ya que captura las condiciones del mercado inmobiliario en un punto específico en el tiempo, sin analizar su evolución.

En el presente capítulo se expone el proceso integral del desarrollo de la investigación.

En primer lugar, se explica cómo se seleccionó y obtuvo la muestra de inmuebles en venta en C.A.B.A. Adicionalmente, se expone una breve descripción de esta, así como el proceso de limpieza de datos necesario para la obtención del resto de las variables.

En segunda instancia, se aborda la extracción de datos a partir de diversas fuentes, incluyendo portales inmobiliarios, institutos de estadística gubernamentales y bases de datos geográficas. Se describen los procedimientos de limpieza y filtrado implementados para garantizar la calidad e integridad de la información, así como la creación de las variables explicativas relevantes según los lineamientos teóricos y las hipótesis planteadas.

En paralelo, se lleva a cabo un análisis exploratorio de los datos procesados, examinando características estadísticas fundamentales, patrones de correlación entre variables y diferencias por localidad geográfica. Esto brinda una comprensión inicial de las relaciones subyacentes en el mercado inmobiliario de C.A.B.A.

A continuación, se desarrolla la especificación del modelo econométrico de precios hedónicos, el cual busca cuantificar los efectos de los distintos atributos estructurales, de accesibilidad, ambientales y socioeconómicos sobre el precio de venta de las propiedades. Partiendo de una forma funcional lineal básica, se evalúan los supuestos del modelo y, de ser necesario, se exploran transformaciones y especificaciones alternativas con el fin de obtener el mejor ajuste posible a los datos observados.

Finalmente, se presenta el modelo hedónico final seleccionado, detallando su forma funcional, variables incluidas y principales resultados e implicancias. Este modelo brinda una cuantificación empírica de la importancia relativa que los distintos factores tienen en la determinación de los precios inmobiliarios en el área de estudio al cierre de 2023.

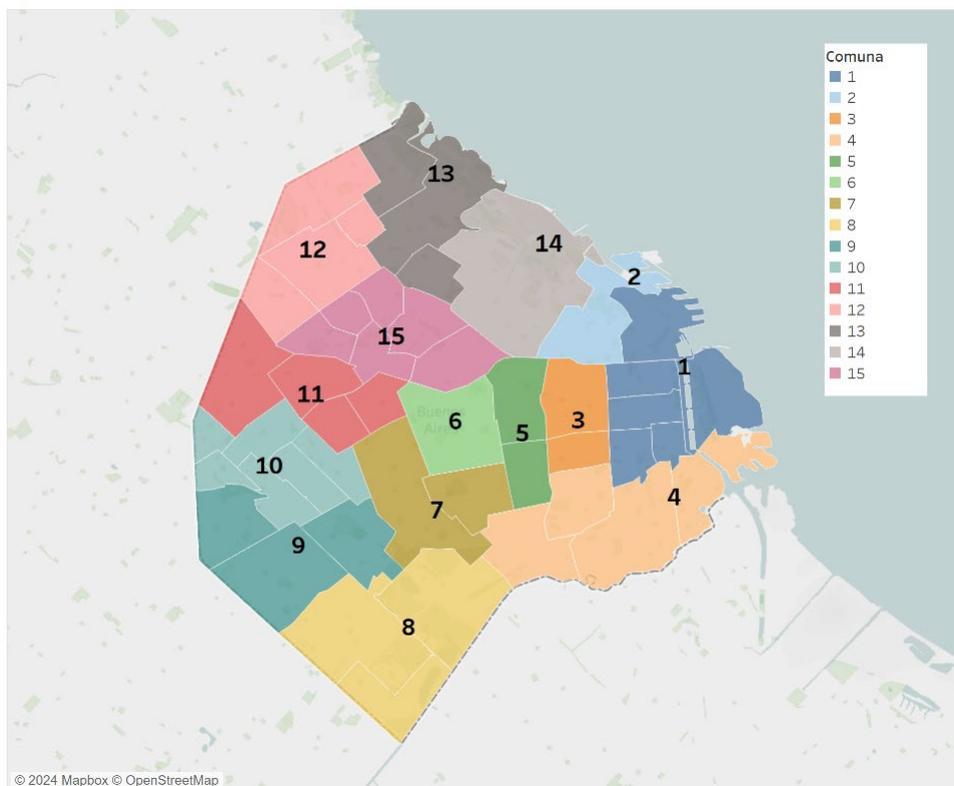
Adicionalmente, se presenta como agregado la aplicación de la metodología del rezago espacial y el uso de modelos de Random Forest, que permiten una mejor predicción del precio de las propiedades en base a sus atributos.

1. Muestra

Antes de adentrarse en la obtención de la muestra utilizada en esta investigación, es importante presentar el contexto geográfico en el que se enmarca el estudio. La Ciudad Autónoma de Buenos Aires se encuentra dividida administrativa y políticamente en comunas y barrios, cada uno con sus características distintivas y dinámicas propias.

Para visualizar de manera clara la distribución espacial de las comunas (desde la Comuna 1 hasta la Comuna 15) de C.A.B.A., se presenta el mapa correspondiente en la Figura 1.

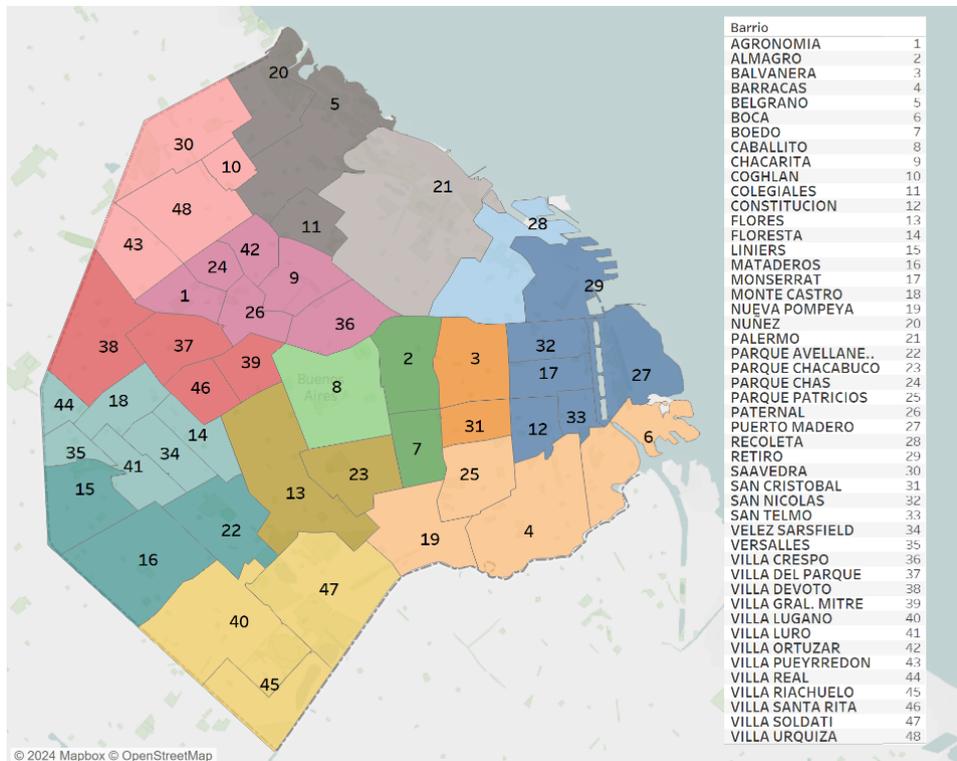
Figura 1: Comunas de C.A.B.A



Fuente: Elaboración propia

Para visualizar de manera clara la distribución espacial de los barrios de C.A.B.A., se presenta el mapa correspondiente en la Figura 2.

Figura 2 Barrios de C.A.B.A



Fuente: Elaboración propia.

Estos mapas permiten apreciar la delimitación geográfica de cada comuna y barrio, así como su ubicación relativa dentro de la ciudad. Contar con esta información visual facilita la comprensión de los patrones espaciales y la interpretación de los resultados del modelo de precios hedónicos que se desarrolló más adelante en este estudio.

A. Elección de la base de datos principal

En primer lugar, se realizó la elección de la base de datos de las propiedades inmobiliarias. Esta es la base principal del análisis, ya que de ella se desprende la obtención de los demás datos.

Los principales portales inmobiliarios disponibles en Argentina son al momento del análisis ArgenProp, Zona Prop, Mercado Libre e Inmoclick. La Tabla 1 muestra una comparación de la cantidad de inmuebles disponibles en venta en cada portal inmobiliario para diciembre de 2023 en C.A.B.A como así también si poseen una interfaz de programación de aplicaciones (“API”) para acceder fácilmente a los datos de los inmuebles.

Tabla 1: Comparación de principales portales inmobiliarios en Argentina

	Departamentos	Casas	Inmueble	API
ArgenProp	69.845	6.115	75.960	No
ZonaProp	70.187	5.247	75.434	No
Mercado Libre	55.971	4.457	60.428	Si
Inmoclick	12.793	1.229	14.022	No

Fuente: Elaboración propia en base a datos disponibles en cada portal inmobiliario (revisión diciembre 2023).

En base a la tabla anterior, se eligió la base de datos de publicaciones de Mercado Libre, y los datos fueron obtenidos a través de su API pública. La elección de esta fuente se justificó en primer lugar por la facilidad que brinda la API pública para la obtención de datos en un formato útil, y por la cantidad de datos disponibles; que, sin ser la fuente de mayor disponibilidad de publicaciones, proporciona una cantidad elevada de observaciones.

En esta primera instancia, se logró extraer un total de 38.414 observaciones⁵, de las cuales 33.962 corresponden a departamentos y 4.452 a casas. Las observaciones están compuestas por las propiedades inmobiliarias con su precio en dólares, características estructurales y su ubicación.

Se pudo observar que la muestra mantiene proporciones similares, en términos de cantidad de inmuebles por comuna, con respecto a la cantidad total de viviendas (total de la población de viviendas) proveniente del Censo Nacional de 2022. Estas proporciones se pueden observar en la Tabla 14 en el Apéndice H.

B. Limpieza de la muestra inicial

Algunas de las observaciones presentaban faltantes en alguna de sus características relevantes, lo que hizo necesaria su eliminación. Por otro lado, varias observaciones presentaban valores atípicos (“*outliers*”) para ciertos atributos. Algunos ejemplos de esto son propiedades que presentaban un área total menor a su área cubierta, propiedades con precios atípicos como \$111.111 o propiedades con valores de antigüedad negativos. Para la eliminación de *outliers* en las variables estructurales no dicotómicas se utilizó el Criterio del Rango Intercuartil (“IQR”)⁶, cuya conveniencia radica en su simplicidad y su capacidad para identificar valores atípicos de manera robusta. Una vez realizada las limpiezas mencionadas anteriormente la base resultante consistía en 16.102 observaciones, de las cuales 14.971 eran departamentos y 1.131 eran casas.

⁵ La API de Mercado Libre posee un límite de llamadas (o “requests”). Es por ello por lo que la cantidad de observaciones extraídas no coincide exactamente con la cantidad de observaciones disponibles en el portal.

⁶ El rango intercuartil se define como la diferencia entre el tercer cuartil (Q3) y el primer cuartil (Q1). Es decir, $IQR = Q3 - Q1$. Este rango cubre el 50% central de los datos y proporciona una medida de la dispersión de los valores. Cualquier valor que se encuentre por debajo del límite inferior ($Q1 - 1.5 \times IQR$) o por encima del límite superior ($Q3 + 1.5 \times IQR$) es considerado un outlier y, por lo tanto, puede ser eliminado del conjunto de datos.

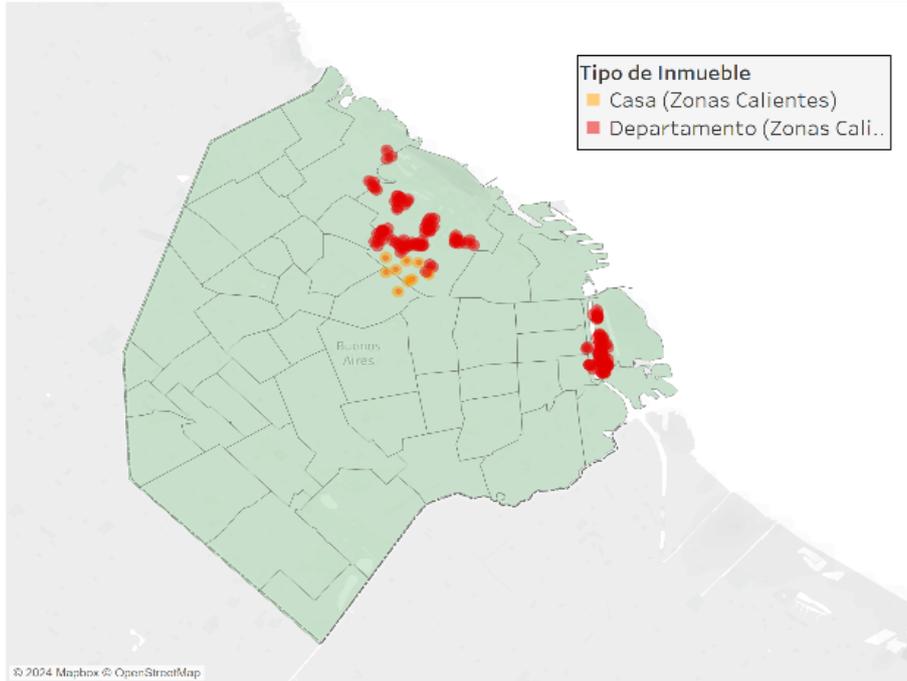
En segundo lugar, para la eliminación de *outliers* en la variable precios, la utilización del IQR no resultó apropiada ya que no se estaría considerando el efecto espacial que podría influir sobre dicha variable. Es decir, algunas propiedades pueden presentar precios de venta atípicos ya que están ubicados en una zona en la que las propiedades suelen ser caras, por lo que si se elimina este tipo de propiedades se estaría perdiendo información importante. Para esta variable se siguió el criterio utilizado por Domínguez Prost (2019), que consiste en la definición de “Zonas Calientes”. Bajo este criterio, primero se detectan los *outliers* en la variable precio por metro cuadrado siguiendo el criterio del IQR. Una vez definidos los inmuebles que cumplen con dicha condición, se procede a detectar la cantidad de propiedades que también son consideradas *outliers* en un rango de 300 metros alrededor de cada propiedad. Las zonas de 300 metros que cuentan con al menos 5 propiedades que serían consideradas *outliers* siguiendo el criterio de IQR son consideradas “Zonas Calientes”, y todas las propiedades ubicadas en dicho radio se conservan en la muestra. Las demás propiedades consideradas *outliers*, siguiendo el criterio del IQR, que no se encuentran en una zona caliente se eliminan de la muestra.

Para la utilización de este criterio se requirió la ubicación precisa de cada propiedad, para lo que fue necesaria la utilización de la API de Google Geocoding que permitió obtener la latitud y longitud de cada inmueble a partir de su domicilio. Siguiendo este criterio, por un lado, se eliminaron 2.054 departamentos y 34 casas domiciliados fuera de C.A.B.A., y por otro lado 154 departamentos y 20 casas considerados *outliers* por su precio.

En total, se encontraron 9 casas ubicadas en “Zonas Calientes” y dichas zonas corresponden a los barrios de Palermo y Villa Crespo. Análogamente, se encontraron 316 departamentos ubicadas en “Zonas Calientes” que corresponden principalmente a los barrios de Puerto Madero, Palermo y Belgrano.

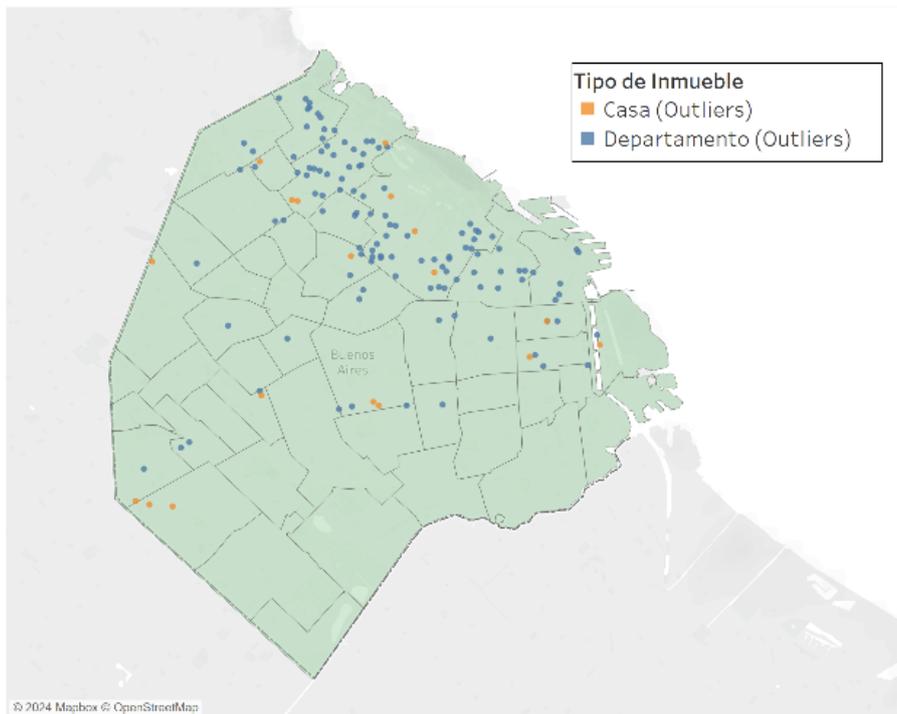
La visualización de las zonas calientes e inmuebles removidos de la muestra se presentan a continuación en la Figura 3 y Figura 4 respectivamente:

Figura 3: Zonas calientes en C.A.B.A.



Fuente: Elaboración propia

Figura 4: Inmuebles outliers en precio por metro cuadrado en C.A.B.A.



Fuente: Elaboración propia

La Tabla 2 presenta un resumen de la cantidad de observaciones luego de cada limpieza:

Tabla 2: Resumen de limpieza de datos

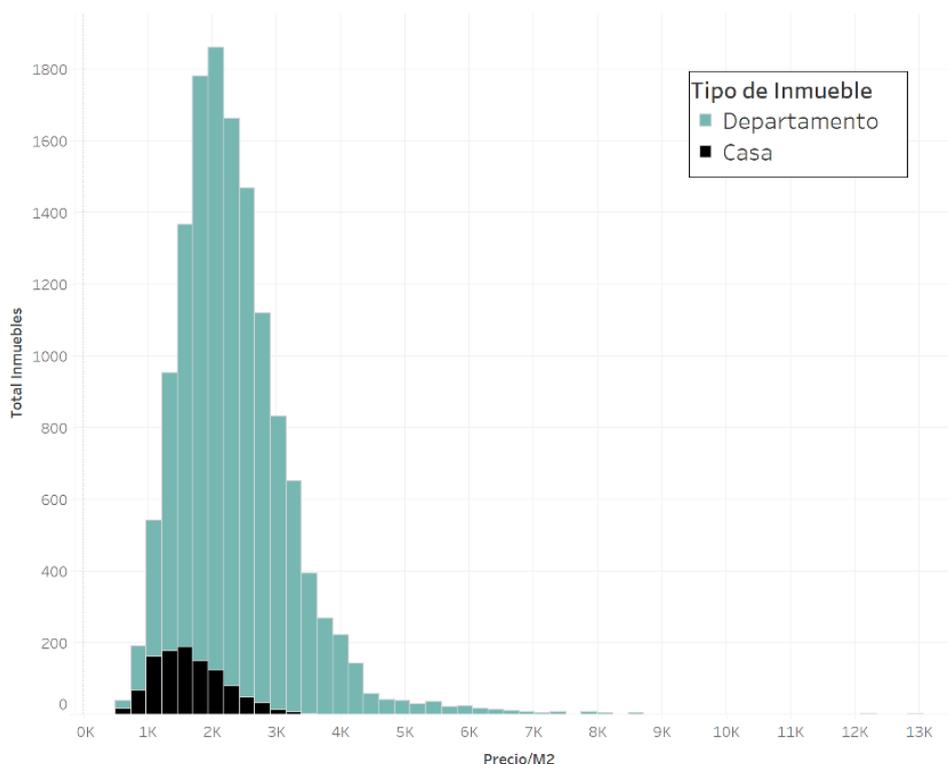
	Casas	Departamentos	Totales
Muestra inicial	4.452	33.962	38.414
Muestra luego de limpieza por valores atípicos	1.131	14.971	16.102
Muestra luego de limpieza por viviendas fuera de C.A.B.A.	1.111	12.943	14.054
Muestra luego de limpieza por zonas calientes	1.077	12.763	13.840

Fuente: Elaboración propia

Luego de la limpieza, la base resultante constó de un total de 13.840 observaciones, de las cuales 12.763 correspondían a departamentos y 1.077 a casas. Para cada uno de los inmuebles se contó con el precio en dólares; variable que se buscó descomponer mediante la aplicación de la teoría de precios hedónicos, relacionándola con sus diferentes atributos.

Para comprender mejor la distribución de los precios de los inmuebles en la muestra obtenida, se presenta a continuación el Gráfico 1, que muestra un histograma de los precios en dólares por metro cuadrado.

Gráfico 1: Distribución de precios por metro cuadrado (USD/m²) de inmuebles en C.A.B.A.

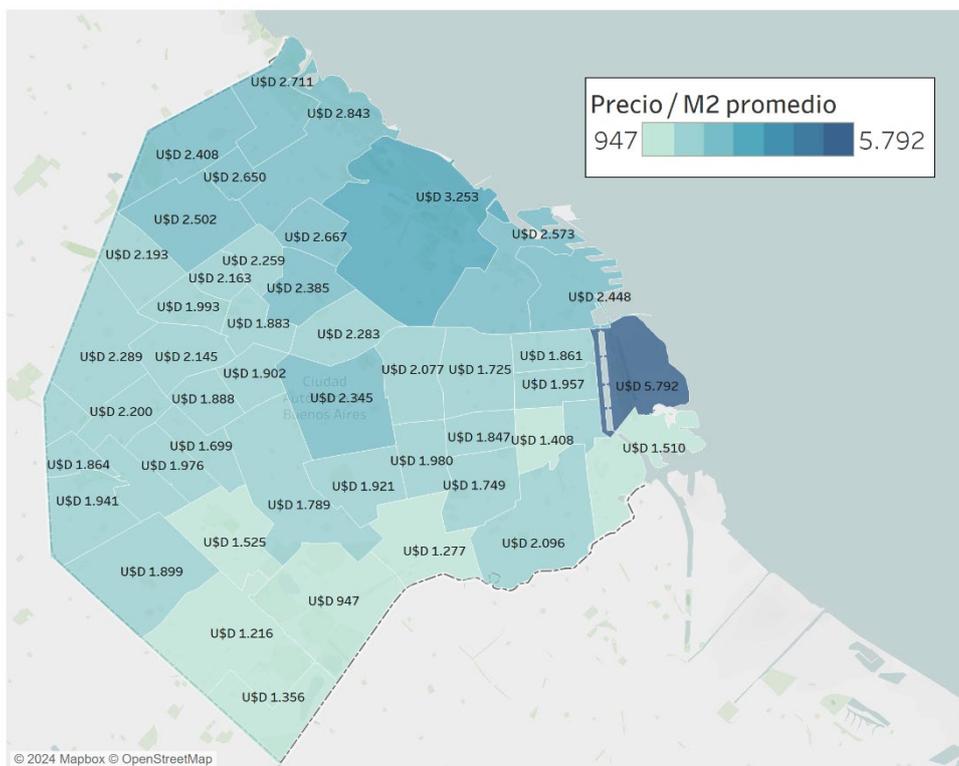


Fuente: Elaboración propia

Este gráfico permite visualizar el total de casas y departamentos en diferentes rangos de precios por metro cuadrado, brindando una idea general de la dispersión y la concentración de los valores en el mercado inmobiliario de C.A.B.A. Se observa que ambos tipos de inmuebles presentan una distribución a simple vista simétrica en sus precios por unidad de superficie y que las casas tienden a tener valores de precios por metro cuadrado menores en comparación a los departamentos.

Además, para analizar la distribución espacial de los precios por metro cuadrado de los inmuebles en la ciudad, se presenta a continuación la Figura 5, que consiste en un mapa de C.A.B.A. que muestra el precio por metro cuadrado promedio por barrio.

Figura 5: Mapa de precio promedio por metro cuadrado (USD/m2) por barrio en C.A.B.A.



Fuente: Elaboración propia

Esta visualización permite identificar de manera intuitiva los barrios con precios más altos y precios más bajos, así como posibles patrones de agrupamiento o dispersión espacial de los valores inmobiliarios. La interpretación de este mapa contribuye a comprender la heterogeneidad de los precios en diferentes áreas de la ciudad y su posible relación con las características de cada barrio. Para los inmuebles disponibles, el barrio más caro es Puerto Madero con un precio por metro cuadrado en dólares (USD/m2) de 5.792 y el barrio más barato es Villa Soldati con un precio por metro cuadrado de 947.

2. Extracción de datos y creación de variables

Como se adelantó anteriormente en el Capítulo II, la teoría de precios hedónicos busca demostrar cómo algunos atributos específicos de una propiedad afectan de manera diferenciada su valor en el mercado.

En la literatura se han tenido en cuenta diferentes y amplias categorías de atributos que tienen influencia sobre el precio de inmuebles. Para la presente investigación se tuvieron en cuenta las siguientes categorías, resumidas en la Figura 6.

Figura 6: Resumen de categorías de variables utilizadas



A continuación, se detalla el proceso integral de obtención de datos para la creación de indicadores de cada una de las categorías de variables, criterios de limpieza en caso de haber sido necesario, como así también un breve análisis exploratorio de los mismos.

A. Variables estructurales

Las variables estructurales son aquellas características físicas intrínsecas de los inmuebles que influyen significativamente en su valor de mercado. Estos atributos están directamente relacionados con la construcción, diseño y espacios de la propiedad, y abarcan aspectos como la superficie total y cubierta, la cantidad y distribución de habitaciones, baños y espacios funcionales, así como la presencia de espacios adicionales y la disponibilidad de estacionamiento propio. Cada uno de estos atributos contribuye de manera diferencial al valor percibido por los compradores y vendedores y, en consecuencia, al precio de mercado del inmueble.

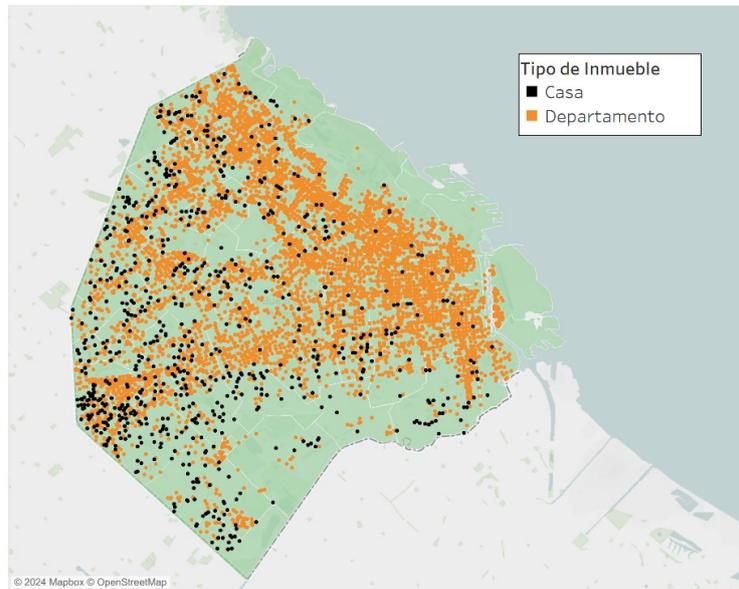
Los datos correspondientes a las variables estructurales se obtuvieron a través de la API pública de Mercado Libre. Esta plataforma de comercio electrónico cuenta con una sección dedicada a la publicación de inmuebles en venta, donde los vendedores proporcionan información detallada sobre las características de sus propiedades. La API permite acceder de manera estructurada a estos datos, facilitando su extracción y procesamiento para el análisis.

A partir de la información obtenida de la API de Mercado Libre, se crearon los siguientes indicadores para capturar las principales variables estructurales de los inmuebles:

- **Antigüedad:** Esta variable representa la antigüedad del inmueble desde su construcción medida en años.
- **Área cubierta (m²):** Esta variable representa la superficie total construida del inmueble, excluyendo espacios descubiertos como patios o terrazas.
- **Área total (m²):** Este indicador se refiere a la superficie total del inmueble, incluyendo tanto el área cubierta como los espacios descubiertos.
- **Número de ambientes:** Este indicador se refiere a la cantidad total de ambientes o habitaciones que componen el inmueble, incluyendo dormitorios, living, comedor, entre otros.
- **Número de dormitorios:** Esta variable indica la cantidad de habitaciones destinadas a dormitorios en el inmueble.
- **Número de baños:** Este indicador representa la cantidad total de baños disponibles en la propiedad.
- **Número de espacios para estacionamiento:** Esta variable captura la cantidad de espacios destinados a estacionamiento de vehículos, ya sea en garaje o cochera.
- **Es casa o departamento:** Esta variable dicotómica indica si el inmueble es una casa (valor 1) o un departamento (valor 0).
- **Tiene piscina o no:** Este indicador binario señala la presencia (valor 1) o ausencia (valor 0) de piscina en la propiedad.
- **Tiene jardín o no:** Esta variable dicotómica indica si el inmueble cuenta con jardín (valor 1) o no (valor 0).

Tras haber detallado los indicadores estructurales obtenidos para cada propiedad, resulta relevante visualizar la distribución geográfica de las casas y departamentos dentro de C.A.B.A., lo cual se presenta en la Figura 7, donde se aprecia la ubicación de ambos tipos de inmuebles en el mapa de la ciudad

Figura 7: Distribución geográfica de casas y departamentos en C.A.B.A.

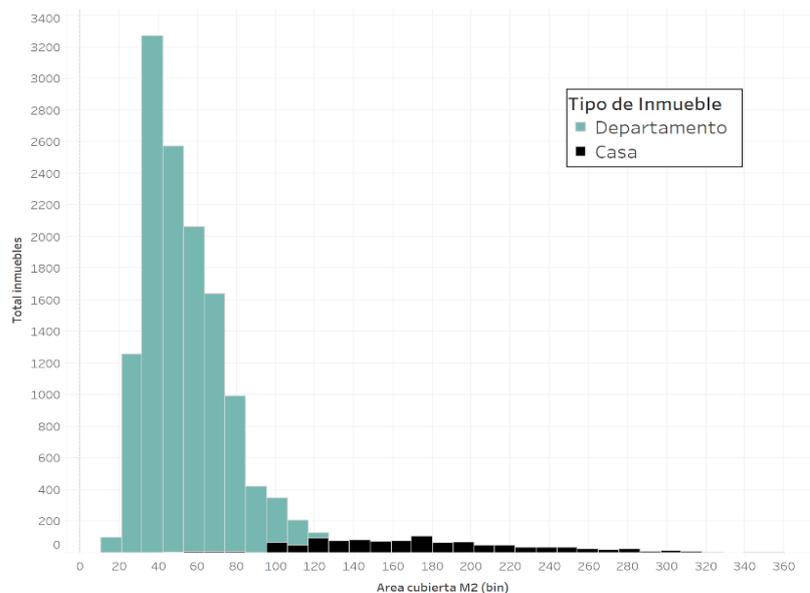


Fuente: Elaboración propia

Podemos observar como las casas están ubicadas mayormente en las afueras de C.A.B.A al Oeste. Los departamentos se distribuyen a lo largo y ancho de todo C.A.B.A. observando una mayor concentración hacia el Este y más cerca del Río de la Plata.

. A su vez se presenta en el Gráfico 2, un histograma que permite visualizar la distribución de los inmuebles en base a su tamaño medido por metro cuadrado.

Gráfico 2: Distribución del tamaño de los inmuebles de la muestra medido en metros cuadrados



Fuente: Elaboración propia

Es posible observar como la mayor cantidad de observaciones presentan un área cubierta menor a 80 metros cuadrados. A partir de esa superficie hay una mayor escasez de propiedades, con propiedades que van desde los 80 metros cuadrados a 360 metros cuadrados en esta muestra.

B. Variables ambientales

Las variables ambientales son aquellos atributos relacionados con el entorno natural y la calidad del medio ambiente que rodea a un inmueble. Estos factores pueden tener un impacto significativo en la valoración de una propiedad, ya que influyen en la calidad de vida de los residentes y en la percepción general del área en donde residen.

Los datos utilizados para la creación de las variables ambientales provienen de tres fuentes principales. En primer lugar, se utilizó la base de datos del Instituto Geográfico Nacional ("IGN"), que proporciona información geográfica y cartográfica de alta calidad para el territorio argentino. En segundo lugar, se recurrió a la base de datos del Gobierno de la Ciudad de Buenos Aires, que contiene información detallada sobre diversos aspectos de la ciudad, incluyendo datos ambientales y de infraestructura.

Además, se utilizó la API de Google Pollution para obtener datos sobre la calidad del aire en cada ubicación específica de los inmuebles. Esta API permite acceder al Índice Universal de Calidad de Aire ("UAQI")⁷, una medida estandarizada que refleja el nivel de contaminación atmosférica en una escala numérica.

Para calcular las distancias entre cada inmueble y los puntos de interés, se utilizaron las coordenadas geográficas de las propiedades y se aplicó la fórmula Haversine⁸. Esta fórmula permite determinar la distancia más corta entre dos puntos en la superficie de una esfera, lo que la hace especialmente adecuada para calcular distancias entre ubicaciones geográficas en la Tierra.

A partir de los datos obtenidos, se crearon varios indicadores para capturar las variables ambientales relevantes. Estos indicadores se dividen en dos categorías principales: atributos escénicos y calidad del aire.

A) Atributos escénicos:

- **Distancia a parques:** Este indicador mide la distancia en metros desde cada inmueble hasta el parque más cercano. Los parques se caracterizan por su paisaje escénico y la presencia de vegetación, y se erigen como espacios públicos que cautivan con su belleza natural. Su cercanía suele ser considerada un atributo apreciado por los potenciales compradores.

⁷ El índice universal de calidad de aire ("UAQI") mide la calidad del aire en un rango de 0 (calidad de aire mala) a 100 (calidad de aire excelente), detectando contaminantes del aire como "CO", "NO₂", "O₃", "PM₁₀", "PM_{2.5}" y "SO₂". Una explicación detallada del índice en cuestión se presenta en el Apéndice A.

⁸ El cálculo de la distancia entre dos puntos con latitudes φ_1 y φ_2 y longitudes λ_1 y λ_2 , se realiza mediante la siguiente fórmula: $d = 2r \cdot \arcsen(\sqrt{hav(\varphi_2 - \varphi_1) + \cos(\varphi_1) \cdot \cos(\varphi_2) \cdot hav(\lambda_2 - \lambda_1)})$ donde r es el radio de la esfera (de alrededor 6.371 mil metros para la Tierra) y $hav(X) = \sin^2(\frac{X}{2})$.

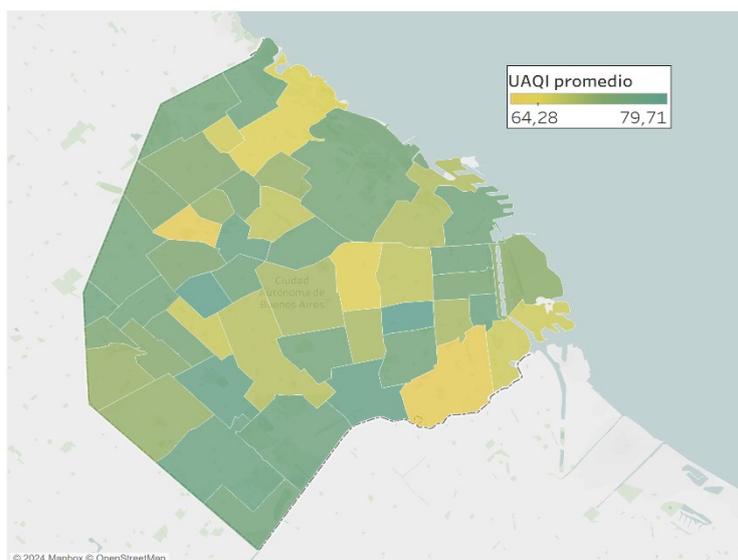
- **Distancia al Río de la Plata:** Este indicador mide la distancia en metros desde cada propiedad hasta la costa del Río de la Plata. La cercanía al río puede ser atractiva por razones escénicas y recreativas.
- **Distancia al Río Matanza-Riachuelo:** Este indicador mide la distancia de cada inmueble hasta el Río Matanza-Riachuelo. Contrario al caso anterior la cercanía a este río podría influir negativamente en el precio por ser un río muy contaminado y que recibe gran cantidad de desperdicios constantemente.
- **Distancia al Puerto de Buenos Aires:** Este indicador mide la distancia en metros desde cada inmueble hasta el Puerto de Buenos Aires. La proximidad al puerto puede tener implicaciones en términos de accesibilidad y vistas.
- **Distancia a plantas de tratamiento de residuos:** Este indicador calcula la distancia en metros desde cada propiedad hasta la planta de tratamiento de residuos más cercana. La cercanía a este tipo de instalaciones puede tener un impacto negativo en la valoración de los inmuebles debido a posibles molestias ambientales.

B) Calidad del aire:

- **UAQI:** Este indicador representa el Índice Universal de Calidad de Aire para cada ubicación de los inmuebles. Un valor más alto del UAQI indica una mayor calidad del aire, lo que puede afectar positivamente la valoración de las propiedades.

A modo exploratorio, resulta interesante visualizar la distribución espacial de la calidad del aire en C.A.B.A., lo cual se muestra en la Figura 8, donde se aprecia el valor del UAQI para cada barrio de la ciudad, representado mediante una escala de colores.

Figura 8: Distribución del UAQI por barrio en C.A.B.A.



Fuente: Elaboración propia

Analizando la escala de valores del UAQI representada en el mapa, se observa que los índices oscilan entre 64 y 80 aproximadamente para los diferentes barrios de C.A.B.A. Según la clasificación del UAQI⁹, valores entre 60 y 80 se consideran "Calidad del aire buena", lo que indica que la calidad del aire en todos los barrios de la ciudad se encuentra dentro de este rango. Si bien no se presentan niveles de contaminación atmosférica alarmantes, es importante destacar que existen diferencias entre los barrios, con algunas zonas mostrando valores más cercanos al límite inferior del rango. Estas variaciones espaciales en la calidad del aire pueden influir en las preferencias de los compradores de viviendas y, por lo tanto, en los precios de los inmuebles en diferentes áreas de la ciudad.

C. Variables comunitarias

Las variables comunitarias son aquellas características relacionadas con el entorno social, económico y cultural en el que se encuentra una propiedad. Estos atributos pueden influir significativamente en la valoración de un inmueble, ya que reflejan la calidad de vida, la seguridad y las dinámicas sociales de la zona.

Los datos utilizados para la creación de las variables comunitarias provienen de diversas fuentes.

En primer lugar, la ubicación precisa de cada propiedad, obtenida a través de la API de Google, permitió determinar la comuna y el barrio correspondiente. Esta información es crucial para la imputación de las características comunitarias a nivel de comuna y barrio.

Los datos sobre las características socioeconómicas, específicamente el porcentaje de hogares con necesidades básicas insatisfechas ("NBI") por comuna, se obtuvieron de los resultados del Censo Nacional de 2010 (INDEC, 2010).

Para capturar el aspecto de la seguridad, se utilizaron los datos de la tasa de robos y hurtos consumados cada 100.000 habitantes por barrio en el año 2022. Estos datos provienen del informe estadístico "Mapa del Delito" del Gobierno de la Ciudad de Buenos Aires (Ministerio de Justicia y Seguridad, 2022).

Además, se buscó capturar las diferencias culturales entre las comunas a través de la variable "cuidado consumo eléctrico". Esta variable se construyó a partir de los datos sobre el porcentaje total de usuarios residenciales con consumo promedio de 12 meses menor a 50 kwh por comuna, obtenidos de las bases de datos del Gobierno de Buenos Aires.

Por último, para capturar el efecto de la densidad poblacional sobre el precio, se utilizó la variable "población por vivienda". Esta variable se calculó dividiendo el total de población por el total de viviendas habitacionales por comuna, utilizando los datos del Censo Nacional de 2022, (INDEC, 2022).

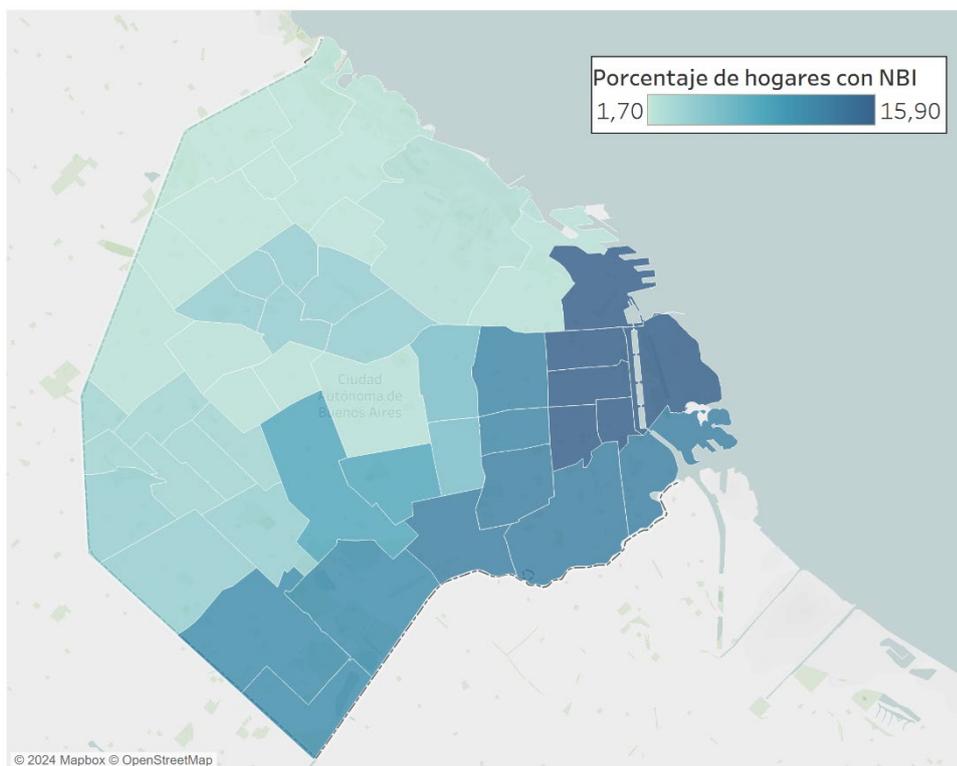
⁹ Se presenta una explicación en detalle del índice de UAQI junto con la clasificación según la API de Google Air Quality el Apéndice A.

A partir de los datos obtenidos, se crearon los siguientes indicadores para representar las variables comunitarias:

- **Porcentaje de viviendas con necesidades básicas insatisfechas por comuna:** Este indicador refleja el nivel de pobreza y carencias en cada comuna. Se espera que un mayor porcentaje de viviendas con necesidades básicas insatisfechas pueda influir negativamente en el precio de las propiedades.
- **Tasa de robos y hurtos consumados cada 100.000 habitantes por barrio:** Este indicador captura el nivel de seguridad en cada barrio. La hipótesis es que una mayor tasa de delitos podría reducir el atractivo y, por lo tanto, el precio de las propiedades en esa zona.
- **Porcentaje total de usuarios residenciales con consumo promedio de 12 meses menor a 50 kwh por comuna:** Este indicador busca capturar diferencias culturales y socioeconómicas entre las comunas. Se plantea la hipótesis de que un menor consumo eléctrico podría estar asociado a una mayor conciencia ambiental, uso de electrodomésticos de bajo consumo y una composición familiar particular (etaria o por cantidad de residentes), factores que podrían influir positivamente en el precio de las propiedades.
- **Población por total de viviendas habitacionales por comuna:** Este indicador refleja la escasez habitacional en cada comuna. La alta densidad de población puede ejercer una presión al alza sobre los precios de oferta de los inmuebles debido a la mayor demanda y la escasez de propiedades, pero también puede afectar negativamente los precios en caso de niveles excesivos de ruido, contaminación y hacinamiento.

Tras haber presentado los indicadores comunitarios considerados en este estudio, resulta relevante visualizar la distribución espacial de dos variables destacadas: el porcentaje de necesidades básicas insatisfechas y el cuidado del consumo eléctrico en los hogares. La Figura 9 muestra el porcentaje de necesidades básicas insatisfechas por comuna en C.A.B.A., mientras que la Figura 10 presenta el porcentaje de usuarios residenciales con consumo eléctrico promedio menor a 50 kwh por comuna, ambos representados mediante una escala de colores.

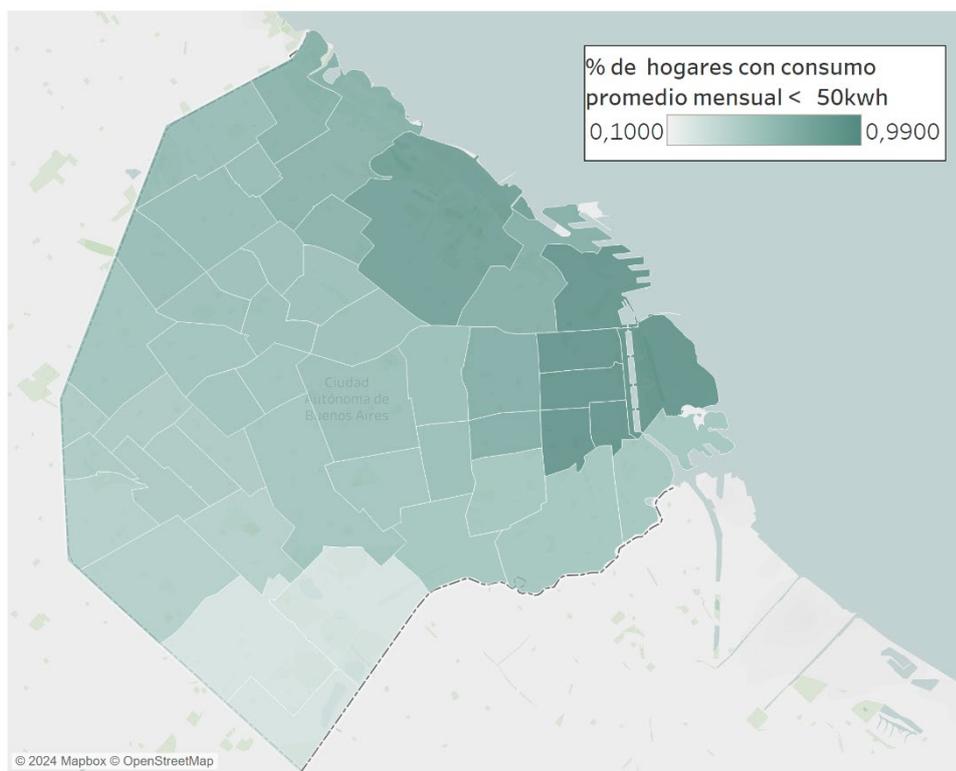
Figura 9: Porcentaje de viviendas con necesidades básicas insatisfechas por comuna en C.A.B.A.



Fuente: Elaboración propia

El mapa de necesidades básicas insatisfechas muestra una clara diferenciación entre las comunas del norte y del sur de la ciudad, más cercanas al Río Matanza-Riachuelo (límite natural de C.A.B.A. con la Provincia de Buenos Aires). Las comunas del sur presentan porcentajes más altos de hogares con necesidades básicas insatisfechas, mientras que las comunas del norte exhiben porcentajes más bajos. Esta distribución espacial sugiere una disparidad socioeconómica entre diferentes zonas de C.A.B.A., lo que puede influir en los precios de los inmuebles.

Figura 10: Porcentaje de usuarios residenciales con consumo eléctrico promedio menor a 50 kwh por comuna en C.A.B.A



Fuente: Elaboración propia

El mapa de cuidado del consumo eléctrico revela un patrón interesante, con las comunas más cercanas al Río de la Plata mostrando menor consumo eléctrico. Esto sugiere que en estas comunas puede haber una mayor conciencia ambiental o una menor restricción presupuestaria con uso de electrodomésticos eficientes, los cuales suelen ser más caros, y una composición demográfica particular. Por otro lado, las comunas con porcentajes más bajos están ubicadas al Sur y al Este de C.A.B.A. Estas variaciones espaciales pueden influir en los precios de las propiedades en diferentes zonas de la ciudad.

D. Variables de accesibilidad

Las variables de accesibilidad son aquellas que miden la facilidad y conveniencia con la que los residentes de un inmueble pueden acceder a servicios de transporte público, sitios de interés y vías de circulación. Estos factores pueden influir significativamente en la valoración de una propiedad, ya que afectan la calidad de vida y la practicidad de vivir en una ubicación determinada.

Los datos utilizados para la creación de las variables de accesibilidad provienen de diversas fuentes. En primer lugar, se utilizaron las coordenadas geográficas de cada inmueble, obtenidas a través de la API de Google Geocoding, para calcular las distancias a los puntos de interés más cercanos como paradas de transporte público, escuelas, comisarías, hospitales, centros culturales o religiosos, centros de

esparcimiento, rutas, autopistas y avenidas. La ubicación precisa de estos puntos de interés se obtuvo a partir de las bases de datos geográficas del Instituto Geográfico Nacional y del Gobierno de la Ciudad de Buenos Aires.

A partir de los datos obtenidos, se crearon los siguientes indicadores para representar las variables de accesibilidad:

A) De acceso a transporte público

- **Total de paradas de bus en un radio de 130 metros**
- **Total de paradas de subte en un radio de 400 metros**
- **Total de paradas de tren en un radio de 1000 metros**

Para crear estos indicadores, se siguió la metodología de Domínguez Prost (2019), justificando el uso de zonas de influencia o "*buffers*" debido a la cantidad de observaciones disponibles y la distribución no contigua de los puntos geográficos. Se calcularon las distancias de cada inmueble al punto de interés más cercano y se realizaron cortes en los percentiles 25, 50 y 75. Luego, se generaron *buffers* con estos radios y se calculó la cantidad de puntos dentro de cada *buffer* para cada inmueble. Se comparó el ajuste de cada *buffer* en el modelo mediante el criterio de Akaike ("AIC") y se seleccionó el radio óptimo para cada variable¹⁰.

Un radio mayor implica que los vendedores del inmueble se ven beneficiados por la proximidad al transporte incluso si está un poco más lejos de las estaciones. Esto les permitiría obtener un mayor precio de venta al poder resaltar la accesibilidad como una ventaja competitiva.

Por el contrario, si el radio óptimo resultase ser menor, los vendedores de inmuebles ubicados a distancias más lejanas de las estaciones no podrían capitalizar tanto la cercanía al transporte público como un factor que aumente el valor de su propiedad.

B) Cercanía a sitios de interés (en metros)

- **Total de sitios de interés en un radio de 500 metros**
- **Distancia a escuela más cercana**
- **Distancia a comisaría más cercana**
- **Distancia a hospital más cercano**
- **Distancia a centro cultural o religioso más cercano**
- **Distancia a centro de esparcimiento más cercano**

¹⁰ Se presenta el proceso integral de cálculo del radio de los *buffers* en el Apéndice B.

Estos indicadores miden la distancia en metros desde cada inmueble hasta el sitio de interés más cercano. Se espera que una menor distancia a estos sitios pueda influir positivamente en el precio de las propiedades.

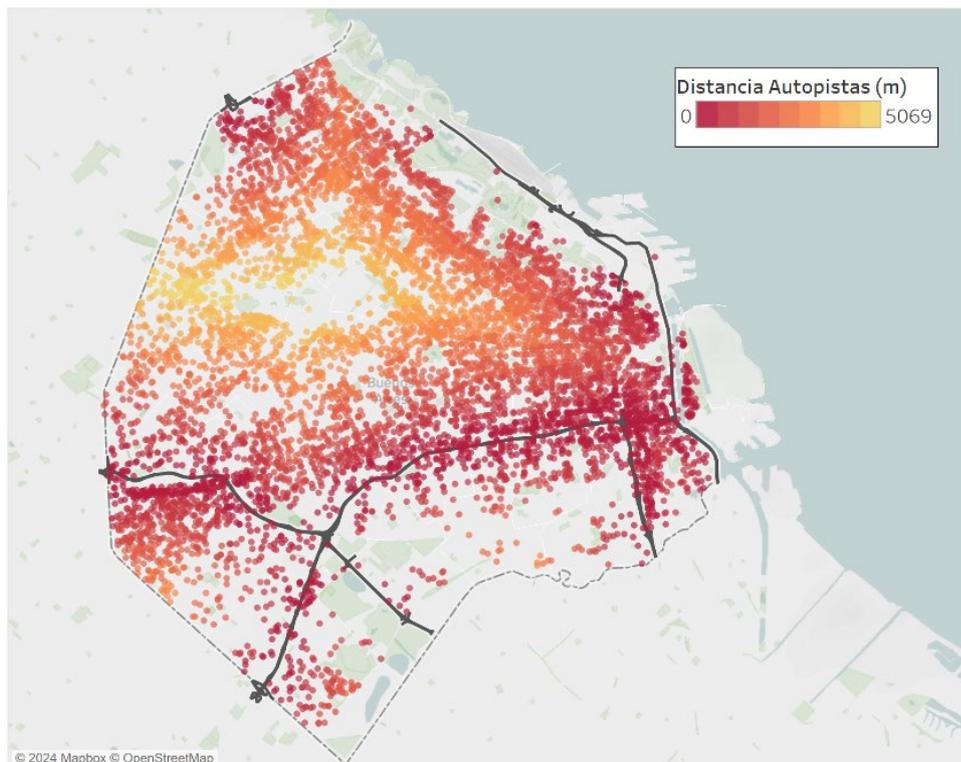
C) Cercanía a vías de circulación (en metros)

- **Distancia a ruta más cercana**
- **Distancia a autopista más cercana**
- **Distancia a avenida más cercana**

Estos indicadores miden la distancia en metros desde cada inmueble hasta la vía de circulación más cercana.

Tras haber detallado los indicadores de accesibilidad obtenidos para cada propiedad, resulta interesante visualizar la distribución espacial de la distancia de los inmuebles a las autopistas en C.A.B.A., lo cual se presenta en la Figura 11, donde se aprecia la ubicación de las propiedades y su proximidad a las principales vías de circulación rápida en el mapa de la ciudad.

Figura 11: Distribución de la distancia de inmuebles a autopistas en C.A.B.A.



Fuente: *Elaboración propia*

El mapa de distancia a autopistas revela un patrón interesante, con inmuebles ubicados en las cercanías de las principales vías de circulación rápida mostrando colores más oscuros, lo que indica una menor

distancia. Por otro lado, las propiedades más alejadas de las autopistas presentan colores más claros. La proximidad a autopistas puede ser un factor relevante para algunos compradores, ya que facilita la movilidad y el acceso a diferentes puntos de la ciudad. Sin embargo, también es importante considerar que la cercanía excesiva a las autopistas podría ser percibida como una desventaja por algunos compradores debido a posibles problemas de ruido o disturbio. Estas variaciones en la distancia a las autopistas pueden influir en los precios de las propiedades en diferentes zonas de C.A.B.A.

En resumen, para la especificación del modelo de precios hedónicos se consideraron diversas categorías de variables explicativas que capturan distintos aspectos relevantes en la determinación del precio de los inmuebles en C.A.B.A. Estas categorías incluyen variables estructurales, que describen las características físicas de las propiedades; variables ambientales, que reflejan la calidad del entorno y la proximidad a espacios verdes; variables comunitarias, que representan aspectos socioeconómicos y culturales de las comunas; y variables de accesibilidad, que miden la cercanía a servicios, transporte y vías de circulación.

La Figura 12 presenta un resumen detallado de todas las variables consideradas en el modelo incluyendo categorías y nombre de indicadores. Esta tabla proporciona una visión completa de los atributos seleccionados para explicar la variabilidad de los precios de los inmuebles en C.A.B.A.

Figura 12: Resumen de variables explicativas para el precio de inmuebles

Variables explicativas del precio de inmuebles	
ESTRUCTURALES	
1. Área total en metros cuadrados	6. Número de espacios para estacionamiento
2. Área cubierta en metros cuadrados	7. Es casa (1) o departamento (0)
3. Número de ambientes	8. Tiene piscina (1) o no (0)
4. Número de dormitorios	9. Tiene jardín (1) o no (0)
5. Número de baños	
AMBIENTALES	
Atributos escénicos	
1. Distancia a parques	4. Distancia al Puerto de Buenos Aires
2. Distancia al Río de la Plata	5. Distancia a plantas de tratamiento de residuos
3. Distancia al Río de la Matanza	
Calidad del aire	
1. UAQI	
COMUNITARIAS	
1. Porcentaje de viviendas con necesidades básicas insatisfechas por comuna	3. Porcentaje total de usuarios residenciales con consumo promedio de 12 meses menor a 50 kwh por comuna
2. Tasa de robos y hurtos consumados cada 100.000 habitantes por barrio	4. Población por total de viviendas habitacionales por comuna
DE ACCESIBILIDAD	
Acceso a transporte público	
1. Total de paradas de bus en un radio de 130m	3. Total de paradas de tren en un radio de 1000m
2. Total de paradas de subte en un radio de 400m	
Cercanía a puntos de interés	
1. Total de sitios de interés en un radio de 500m	4. Distancia a hospital más cercano
2. Distancia a escuela más cercana	5. Distancia a centro cultural o religioso más cercano
3. Distancia a comisaría más cercana	6. Distancia a centro de esparcimiento más cercano
Cercanía a vías de circulación	
1. Distancia a ruta más cercana	3. Distancia a avenida más cercana
2. Distancia a autopista más cercana	

3. Especificación del modelo de precios hedónicos

Para lograr los objetivos planteados en esta investigación y determinar la relación entre el precio de los inmuebles en venta y sus diversos atributos, es necesario especificar un modelo de precios hedónicos. Este modelo econométrico permite cuantificar la influencia de cada variable explicativa sobre la variable dependiente, en este caso, el precio en dólares de las propiedades en venta. Mediante el enfoque denominado “Enfoque de las características” popularizado por Lancaster (1966), se buscó describir el precio de los inmuebles en base a sus características, argumentando que la utilidad del inmueble se deriva de la posesión de esos atributos.

En función de esto se decidió utilizar dos tipos de formas funcionales (lin-lin y log-log), cuyo proceso de especificación se desarrolla en esta sección. Además, se presentan enfoques complementarios de especificación incluyendo un modelo con rezago espacial y un modelo de Random Forest en el Apéndice E.

A. Modelo lin-lin

En el marco teórico se presentaron diversas formas funcionales utilizadas en la literatura de precios hedónicos, desde especificaciones lineales simples hasta modelos más complejos como los semilogarítmicos o log-log. En investigaciones previas, como las de Ridker y Henning (1967), Freeman (1979), Rosen (1974) y Malpezzi (2008), se han empleado diferentes formas funcionales según las características de los datos y los objetivos del estudio.

Inicialmente, se exploró una especificación lineal para el ajuste de los datos utilizando el lenguaje de programación Python. El modelo de precios hedónicos puede representarse mediante la siguiente ecuación de regresión:

Ecuación 6: Primera especificación del modelo de precios hedónicos

$$P_i = \beta_0 + \beta_E * E + \beta_A * A + \beta_C * C + \beta_B * B + \varepsilon_i$$

Donde:

- P es el vector de la variable dependiente que representa el precio en dólares de cada inmueble.
- β_0 es el término de intercepto o constante del modelo.
- E es la matriz de variables explicativas que incluye los atributos estructurales.
- A es la matriz de variables explicativas que incluye los atributos ambientales.
- C es la matriz de variables explicativas que incluye los atributos comunitarios.
- B es la matriz de variables explicativas que incluye los atributos de accesibilidad.
- $\beta_E, \beta_A, \beta_C$ y β_B son los vectores de coeficientes de regresión correspondientes a cada matriz de atributos.
- ε es el vector de términos de error estocástico o perturbaciones aleatorias.

Los resultados de la regresión inicial se presentan en la Tabla 8, en el Apéndice D .

a. Verificación de supuestos del modelo

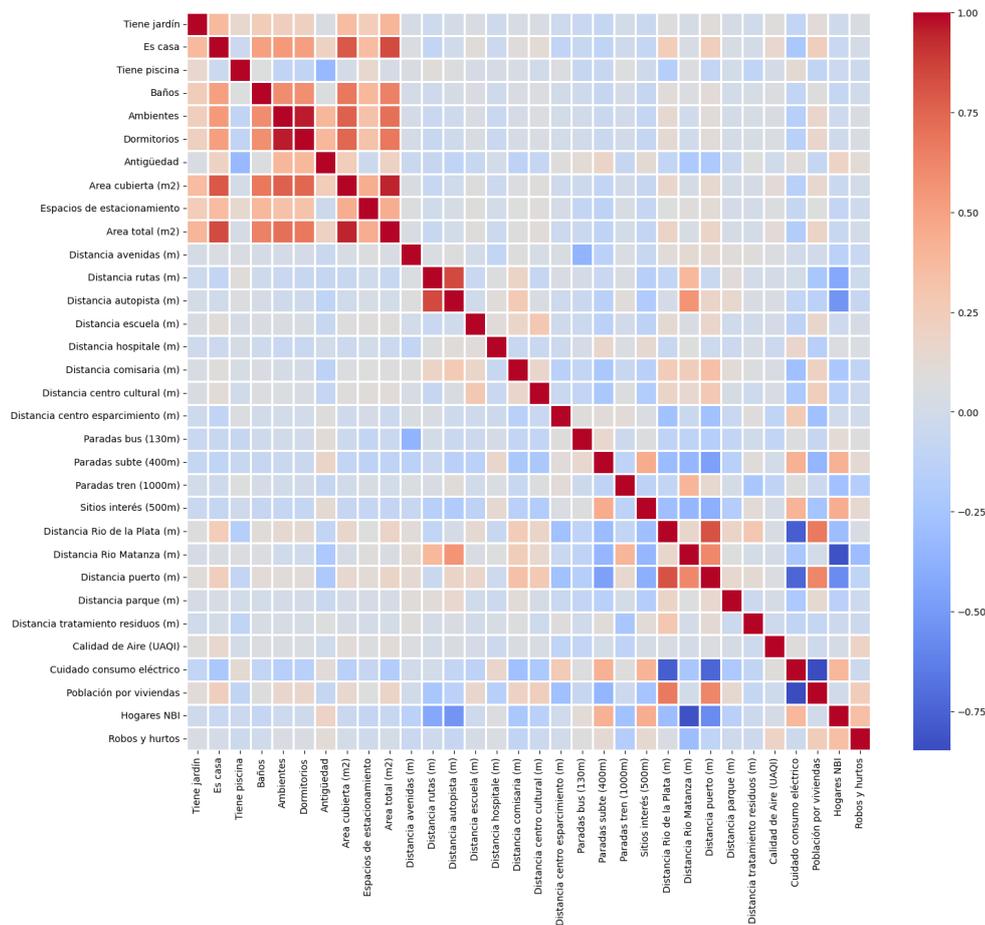
Para que las estimaciones de los coeficientes de regresión sean confiables y consistentes, el modelo clásico de regresión lineal se basa en una serie de supuestos¹¹. Se realizó un análisis detallado para verificar el cumplimiento de estos supuestos en el modelo inicialmente especificado.

¹¹ Esta serie de supuestos se presentan en detalle en el Apéndice C Apéndice A.

El supuesto de multicolinealidad del modelo fue corroborado a través del uso del Factor Inflacionario de la Varianza Generalizado (GVIF), que se puede aplicar también a variables dicotómica. Mediante el criterio del GVIF¹² se corrobora la ausencia de multicolinealidad cuando el valor de este no supera el límite de 5 (Buteikis, 2020). Los resultados, presentados en la Tabla 11 en el Apéndice F demuestran que no hay serios problemas de multicolinealidad en el modelo.

Para complementar visualmente este fenómeno, se examinó la matriz de correlaciones de primer orden de las variables independientes, la cual se muestra a continuación como Figura 13:

Figura 13: Matriz de correlación de primer orden de variables explicativas



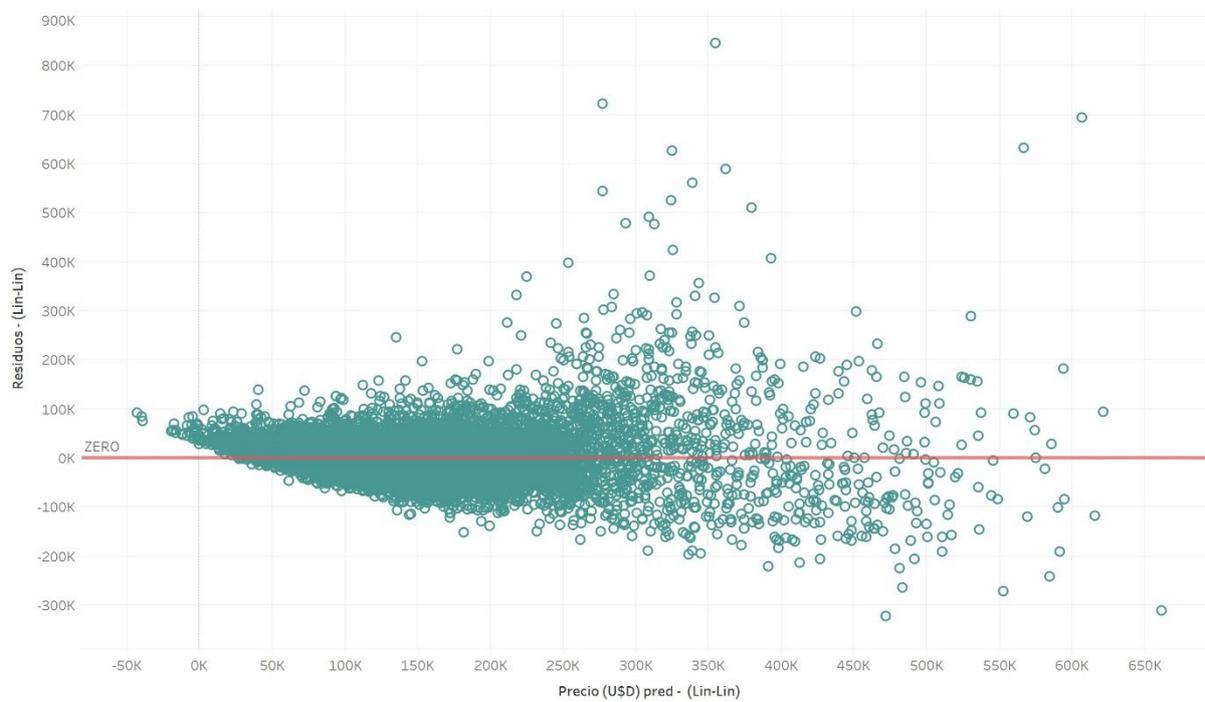
Esta matriz muestra los coeficientes de correlación de primer orden entre cada par de variables explicativas. A simple vista pueden apreciarse valores altos en términos absolutos (cercanos a 1 o -1) para algunas variables lo que indicaría una posible correlación de primer orden entre las mismas.

Por otro lado, surgió la sospecha de la violación del supuesto de homoscedasticidad de los residuos al analizar gráficamente los residuos al cuadrado contra los valores predichos. Se observó un patrón que

¹² El GVIF utilizado para la corroboración de multicolinealidad es el GVIF corregido por los grados de libertad del modelo $GVIF\left(\frac{1}{2 \cdot Df}\right)$. El cálculo de GVIF para cada variable explicativa se presenta en Apéndice E.

sugiere que, a medida que aumenta el precio de los inmuebles, la variabilidad de los residuos también se incrementa. Se presenta dicho fenómeno en el Gráfico 3.

Gráfico 3: Dispersión de residuos al cuadrado contra valores predichos



Para confirmar esta observación, se realizaron la prueba de Breusch-Pagan¹³. La prueba arrojó un estadístico con valor p infinitesimal lo cual respalda la presencia de heteroscedasticidad en el modelo inicial.

La detección de heteroscedasticidad en el modelo de precios hedónicos plantea desafíos importantes para la estimación y la inferencia. Estos problemas pueden afectar la precisión de los coeficientes estimados, la significancia estadística de las variables y la capacidad predictiva del modelo. Por lo tanto, fue necesario abordar este problema mediante técnicas econométricas adecuadas antes de proceder con la interpretación de los resultados.

b. Transformaciones al modelo lin-lin

Aun cuando se constató la ausencia de multicolinealidad mediante el criterio del GVIF, entendiendo que la multicolinealidad es un problema de grado y no de existencia, se procedió a simplificar el modelo con proceso de reducción de variables.

En primer lugar, se identificaron aquellas variables que presentaban sospecha de alta correlación de primer orden con otras variables. Este es el caso de la variable estructural ambientes con las variables

¹³ Los resultados de la prueba de heteroscedasticidad de residuos para el primer modelo estimado se encuentran en el Apéndice D, junto con los resultados de la regresión.

estructurales cantidad de baños y cantidad de dormitorios, por lo que se procedió a eliminar del modelo estas últimas dos. Asimismo, la variable área total también fue descartada al presentar una alta correlación con el área cubierta. Finalmente, se decidió eliminar la variable cuidado del consumo eléctrico, la cual estaba altamente correlacionada con las variables población por vivienda y distancia al Río de la Plata.

Tras abordar el proceso anterior, se procedió a solventar los problemas de heteroscedasticidad detectados en el modelo. Para enfrentar la heteroscedasticidad, se optó por aplicar una transformación logarítmica tanto a la variable dependiente (precio) como a las variables explicativas continuas. Aplicar una transformación logarítmica puede ayudar a estabilizar la varianza y cumplir con el supuesto de homoscedasticidad.

Es importante destacar que la transformación logarítmica no se aplicó a las variables dicotómicas, ya que estas representan categorías y no tienen una interpretación cuantitativa. En su lugar, las variables dicotómicas se incluyeron en el modelo transformado sin aplicarles logaritmos, manteniendo su interpretación original. Esta forma funcional conocida como log-log fue presentada en el marco teórico y se desarrolla en detalle a continuación.

Ante esta situación de heteroscedasticidad, es recomendable utilizar el método de Mínimos Cuadrados Generalizados ("MCG") en lugar de MCO. El estimador de MCG pondera las observaciones de manera inversamente proporcional a la varianza de los errores, asignando mayor peso a las observaciones con menor varianza del error y menor peso a las observaciones con mayor varianza. De esta forma, MCG logra ser un estimador eficiente, alcanzando la menor varianza posible entre todos los estimadores lineales insesgados (Gujarati y Porter, 2008). Esto se traduce en estimaciones más precisas de los coeficientes, con menores errores estándar, intervalos de confianza más estrechos y pruebas de hipótesis más acertadas en comparación con MCO.

B. Modelo log-log

La aplicación de logaritmos tiende a comprimir las escalas de las variables, reduciendo la dispersión de los datos y estabilizando la varianza de los errores. Esto ayuda a mitigar los problemas de heteroscedasticidad y mejora la calidad de las estimaciones.

Además, la interpretación de los coeficientes en un modelo log-log es intuitiva, ya que representan elasticidades. Un coeficiente β_j puede interpretarse como el cambio porcentual en el precio ante un cambio porcentual en la variable explicativa X_j , manteniendo constantes las demás variables.

Habiéndose tratado la simplificación del modelo mediante la reducción del número de variables explicativas, y con el fin de tratar los problemas de heteroscedasticidad se procedió a la especificación de un segundo modelo de forma funcional log-log, esta vez utilizando el método de estimación de mínimos cuadrados generalizados.

La forma funcional del mismo se puede expresar de la siguiente manera:

Ecuación 7: Segunda especificación del modelo de precios hedónicos

$$\log(P_i) = \beta_0 + \beta_E * \log(E) + \beta_A * \log(A) + \beta_C * \log(C) + \beta_B * \log(B) + \beta_D * D + \mu_i$$

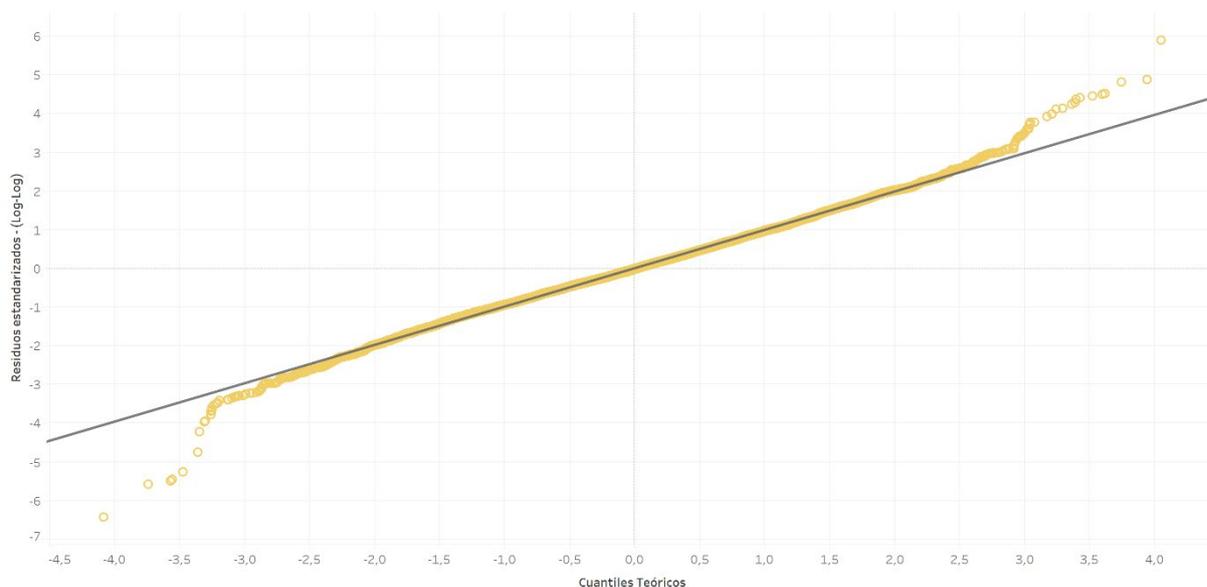
Donde:

- P_i es el vector de la variable dependiente que representa el precio en dólares de cada inmueble.
- β_0 es el término de intercepto o constante del modelo.
- E es la matriz de variables explicativas no dicotómicas que incluye los atributos estructurales.
- A es la matriz de variables explicativas que incluye los atributos ambientales.
- C es la matriz de variables explicativas que incluye los atributos comunitarios.
- B es la matriz de variables explicativas que incluye los atributos de accesibilidad.
- D es la matriz de variables explicativas dicotómicas que incluyen atributos estructurales.
- $\beta_E, \beta_A, \beta_C$ y β_B, β_D son los vectores de coeficientes de regresión correspondientes a cada matriz de atributos.
- μ_i es el vector de términos de error estocástico o perturbaciones aleatorias.

Luego de aplicar la transformación log-log y realizar las pruebas correspondientes, se constató que los problemas de heteroscedasticidad no desaparecieron. Esto pudo corroborarse mediante la prueba de Bausch-Pagan¹⁴, la cual evidencia la existencia del problema. Asimismo, esto puede visualizarse en el Gráfico 4 de tipo cuantil-cuantil ("QQ-plot") en donde puede observarse colas pesadas, lo que indica que la distribución de los residuos no es de tipo gaussiano.

¹⁴ Los resultados de la prueba de heteroscedasticidad de residuos para el segundo modelo estimado se encuentran en el Apéndice D, junto con los resultados de la regresión.

Gráfico 4: Cuantil-cuantil de residuos del modelo log-log



Ambos modelos trabajados en la parte central de la investigación (lin-lin y log-log) violan el supuesto de homocedasticidad residual. La violación de este supuesto, como advierte Potrawa (2020), es una de las principales críticas del modelo de precios hedónicos ya que este no logra capturar relaciones no lineales de las variables.

Es importante aclarar que, pese a la presencia de heteroscedasticidad residual, la propiedad de insesgamiento y consistencia de los estimadores de mínimos cuadrados generalizados no se pierden. Puede demostrarse además, que cuando el número de observaciones tiende a infinito, los estimadores de la regresión se distribuyen de forma normal. Además la varianza sigue siendo la mínima con respecto a los demás estimadores lineales, aunque esta no necesariamente sea baja.

Se encuentran en la literatura de precios hedónicos, autores que también han lidiado y explicado los orígenes de este fenómeno en los modelos de este estilo:

Según Goodman, A., Thibodeau, T. (1995) una causa de la heteroscedasticidad es la antigüedad del inmueble, que influye en el precio de mercado a través de la depreciación económica y el efecto "*vintage*". Viviendas más antiguas tienen mayor probabilidad de haber sido renovadas o mejoradas significativamente, lo cual típicamente no está registrado en los datos disponibles. Esto introduce heteroscedasticidad ya que el error de predicción del precio probablemente aumenta con la edad. La interacción entre el mantenimiento y el efecto *vintage* también contribuye. Incluso asumiendo una sola renovación durante la vida de la casa, éstas pueden diferir según la época en que se hicieron, aumentando la varianza a medida que la edad aumenta. Estos efectos no lineales no han sido captados por los modelos propuestos.

En la práctica los modelos no paramétricos ofrecen una alternativa atractiva para el análisis de precios hedónicos debido a su menor dependencia de supuestos restrictivos sobre la forma funcional y la distribución de los errores. Estos modelos, como la regresión local (“*lowess*”), los modelos aditivos generalizados (“*GAM*”), las redes neuronales artificiales y los modelos de Random Forest, son menos sensibles a valores atípicos y pueden capturar relaciones complejas y no lineales entre las características de los inmuebles.

Si bien el alcance de esta tesis no se extiende a la aplicación exhaustiva de modelos no paramétricos para el análisis de precios hedónicos, se presentan dos modelos adicionales que tienen el potencial de mejorar los resultados obtenidos, en el Apéndice E: “Extensiones al modelo de precios hedónicos”.

Se presentan a continuación los resultados de la regresión log-log con el fin de la verificación de las hipótesis y la consecución de los objetivos de la tesis. Una versión reducida de los resultados de la regresión final se presenta a continuación en la Tabla 3 y Tabla 4. Los resultados extendidos se presentan la Tabla 9 en el Apéndice D.

Tabla 3: Resumen reducido de coeficientes de la regresión log-log (MCG)

Nombre de la variable	Coficiente	Error estándar	z	P> z	[0.025	0.975]
Constante	8,9120	0,150	59,420	0,000	8,618	9,206
Área cubierta (m ²)	0,8765	0,012	74,845	0,000	0,854	0,899
Espacio de estacionamiento	0,2650	0,008	31,927	0,000	0,249	0,281
Tiene piscina	0,1577	0,006	25,652	0,000	0,146	0,170
Es casa	-0,1076	0,013	-7,987	0,000	-0,134	-0,081
Ambientes	0,0833	0,015	5,645	0,000	0,054	0,112
Antigüedad	-0,0668	0,002	-44,343	0,000	-0,070	-0,064
Tiene jardín	0,0111	0,011	1,033	0,302	-0,010	0,032
Distancia Río de la Plata (m)	-0,1296	0,005	-25,817	0,000	-0,139	-0,120
Distancia puerto (m)	-0,0720	0,013	-5,460	0,000	-0,098	-0,046
Distancia Río Matanza (m)	0,0572	0,008	6,878	0,000	0,041	0,073
Distancia tratamiento residuos (m)	0,0370	0,004	8,476	0,000	0,028	0,046
Calidad del aire (UAQI)	0,0342	0,028	1,235	0,217	-0,020	0,089
Distancia parque (m)	-0,0029	0,002	-1,328	0,184	-0,007	0,001
Población por viviendas	-0,3398	0,037	-9,146	0,000	-0,413	-0,267
Hogares NBI	-0,0952	0,006	-15,630	0,000	-0,107	-0,083
Robos y hurtos	-0,0030	0,005	-0,573	0,567	-0,013	0,007
Paradas subte (400 m)	-0,0758	0,007	-10,793	0,000	-0,090	-0,062
Distancia centro esparcimiento (m)	0,0303	0,004	8,640	0,000	0,023	0,037
Distancia ruta (m)	0,0282	0,006	5,123	0,000	0,017	0,039
Paradas bus (130 m)	-0,0264	0,003	-7,630	0,000	-0,033	-0,020
Sitios de interés (500 m)	-0,0235	0,004	-6,176	0,000	-0,031	-0,016
Distancia autopista (m)	-0,0232	0,005	-4,234	0,000	-0,034	-0,012
Distancia hospital (m)	0,0175	0,003	5,515	0,000	0,011	0,024
Distancia centro cultural (m)	0,0113	0,003	3,530	0,000	0,005	0,018
Distancia comisaría (m)	0,0074	0,003	2,273	0,023	0,001	0,014
Distancia escuela (m)	0,0054	0,003	1,608	0,108	-0,001	0,012
Paradas tren (1000 m)	0,0043	0,007	0,658	0,511	-0,008	0,017
Distancia avenida (m)	-0,0015	0,002	-0,943	0,346	-0,005	0,002

La mayoría de las variables incluidas en el modelo son estadísticamente significativas al nivel del 1%, lo que sugiere que tienen una influencia relevante en la determinación de los precios inmobiliarios.

Tabla 4: Resumen reducido de resultados de la regresión log-log (MCG)

R2	0,797	Jarque-Bera (JB)	579,642
R2 ajustado	0,796	Prob(JB)	1,36 E-126
F-statistic	1778	Breusch-Pagan test statistic	684,83
Prob (F-statistic)	0	Breusch-Pagan test p-value	2,90 E-126

El modelo muestra un buen ajuste general, con un R-cuadrado ajustado de 0,796 lo que indica que las variables incluidas explican aproximadamente el 79,6% de la variabilidad en los precios de las propiedades. A su vez, la prueba de Breusch-Pagan demuestra la aún existencia de heteroscedasticidad, aunque como se dijo anteriormente, esto no desempeña rol alguno con la propiedad de insesgamiento e inconsistencia de los estimadores.

C. Interpretación de los resultados

El modelo de precios hedónicos log-log aplicado a los precios de oferta de inmuebles en C.A.B.A. en diciembre de 2023 proporciona información valiosa sobre cómo diversos atributos influyen en las expectativas de precios de los vendedores. Los resultados revelan patrones interesantes y brindan una perspectiva interesante sobre la dinámica del mercado inmobiliario desde el punto de vista de la oferta.

a. Variables estructurales

En cuanto a los atributos estructurales, el modelo destaca la importancia del área cubierta en la determinación de los precios de oferta. Un aumento del 1% en el área cubierta está asociado con un incremento promedio del 0,88% en el precio solicitado por los vendedores, manteniendo todo lo demás constante. Esto sugiere que los vendedores valoran significativamente el espacio habitable al establecer sus expectativas de precios. Además, el modelo indica que las casas tienen en promedio precios de oferta un 10,76% más bajos que los departamentos, controlando por otros factores, lo que podría reflejar una percepción de mayor demanda o preferencia por departamentos en el mercado de C.A.B.A.

La presencia de piscina también se destaca como un atributo que los vendedores consideran valioso, ya que añade una prima considerable a los precios de oferta de las propiedades que cuentan con esta característica. La variable antigüedad en años es significativa y tiene un coeficiente negativo, lo que indica que un aumento en la antigüedad de un inmueble está asociado con una disminución promedio en su precio de oferta.

La variable de espacio de estacionamiento es significativa y tiene un coeficiente positivo, lo que sugiere que los vendedores valoran la disponibilidad de espacios de estacionamiento al establecer sus expectativas de precios. En una ciudad con una alta densidad de tráfico y limitaciones de estacionamiento en la vía

pública, contar con un espacio de estacionamiento propio puede ser un atributo atractivo para los compradores y, por lo tanto, los vendedores pueden asignar una prima en el precio de oferta a las propiedades que cuentan con esta característica.

La variable "Tiene jardín" no muestra un efecto estadísticamente significativo en los precios de oferta, lo que sugiere que la presencia de este atributo no es un factor determinante en las expectativas de precios de los vendedores. De manera similar, la variable "Ambientes" tampoco resulta significativa. Esto podría deberse a que los vendedores valoran el tamaño de dichos ambientes que ya es capturada por la variable de área cubierta por sobre la cantidad de estos.

b. Variables ambientales

En relación con los atributos ambientales, el modelo revela diferencias interesantes en cómo los vendedores perciben la ubicación de las propiedades con respecto a los principales ríos de la región. Los resultados indican que los vendedores tienden a establecer precios de oferta más altos para propiedades más cercanas al Río de la Plata, lo que podría reflejar una valoración positiva de las vistas al río o de la proximidad a zonas costeras. Por otro lado, los vendedores parecen asignar precios de oferta más altos a medida que aumenta la distancia al Río Matanza, lo que sugiere una percepción menos favorable de las ubicaciones cercanas a este río.

La variable "Distancia puerto (m)" es significativa y tiene un coeficiente negativo, lo que indica que un aumento del 1% en la distancia al puerto está asociado con una disminución promedio del 0,07% en el precio de oferta de los inmuebles en C.A.B.A. Es importante destacar que la ubicación del puerto de Buenos Aires, situado estratégicamente entre el Aeroparque Jorge Newbery y el barrio de Puerto Madero, puede contribuir al atractivo de las propiedades cercanas y, por lo tanto, influir en las expectativas de precios de los vendedores. La combinación de exclusividad y vistas escénicas en esta ubicación puede justificar los precios de oferta más altos para las propiedades más cercanas al puerto.

Los resultados del modelo muestran que la variable "Calidad del aire (UAQI)" no tiene un efecto estadísticamente significativo en los precios de oferta de las propiedades en C.A.B.A. Esto sugiere que, según los datos analizados, la calidad del aire no parece ser un factor determinante en las expectativas de precios de los vendedores. Este hallazgo puede parecer sorprendente, aunque es importante tener en cuenta que la calidad del aire de las propiedades en C.A.B.A. se encuentra dentro de un rango considerablemente bueno y sin mucha diferencia entre las mismas. Sin embargo, se observa que otros atributos ambientales como lo es la cercanía a centros de tratamiento de residuos parecería tener un efectivo negativo significativo sobre los precios de venta lo que podría estar capturando la percepción del entorno de las propiedades.

c. Variables comunitarias

Los atributos comunitarios también influyen en las expectativas de precios de los vendedores. El modelo muestra que los vendedores tienden a establecer precios de oferta más bajos en áreas con mayor densidad

poblacional. Este hallazgo puede estar relacionado con varios factores. En primer lugar, una mayor densidad poblacional puede estar asociada con una menor calidad de vida percibida, ya que puede implicar una mayor congestión, ruido, tráfico y presión sobre los servicios públicos y la infraestructura. En segundo lugar, las áreas con mayor densidad poblacional pueden tener una mayor proporción de viviendas más pequeñas o de menor calidad, lo que puede influir en los precios de oferta.

Por otro lado existe una relación negativa entre el precio de venta y la proporción de hogares con necesidades básicas insatisfechas ("NBI"). Estos resultados sugieren que los vendedores son conscientes del impacto de las condiciones socioeconómicas del entorno en la demanda y la valoración de las propiedades.

Además, la tasa de robos y hurtos tiene un efecto negativo, aunque pequeño, en los precios de oferta, lo que indica que los vendedores tienen en cuenta la seguridad al determinar sus expectativas de precios.

d. Variables de accesibilidad

Al analizar las variables relacionadas con las distancias a diferentes puntos de interés, se observan patrones interesantes en las percepciones de los vendedores y sus expectativas de precios de oferta. En general, los resultados sugieren que los vendedores perciben la lejanía de centros de esparcimiento, hospitales, rutas y centros culturales como atributos ligeramente positivos. Esto puede indicar que los vendedores valoran más la tranquilidad y la privacidad que ofrecen las ubicaciones más alejadas de estos puntos de interés, y están dispuestos a asignar precios un poco más altos a las propiedades que cumplen con estas características.

En cuanto a las distancias a las vías de circulación, se observa un patrón diferente para las autopistas en comparación con las avenidas y rutas. Los vendedores parecen percibir la cercanía a las autopistas como un atributo ligeramente positivo, ya que un aumento del 1% en la distancia a estas vías está asociado con una disminución promedio del 0,02% en los precios de oferta. Esto sugiere que los vendedores valoran la accesibilidad y la conectividad que ofrecen las autopistas, y están dispuestos a asignar precios un poco más altos a las propiedades más cercanas a estas vías rápidas. Por otro lado, la distancia a las avenidas no parece tener un impacto significativo en los precios de oferta según este modelo. Por el contrario, la cercanía a las rutas parece tener un impacto significativo negativo sobre el precio que perciben los vendedores.

Al examinar las variables relacionadas con las paradas de transporte público, se observa que tanto la presencia de paradas de subte en un radio de 400m como la presencia de paradas de bus en un radio de 130m están asociadas con disminuciones promedio en los precios de oferta. Estos resultados sugieren que los vendedores perciben la alta densidad de paradas de transporte público en las cercanías de una propiedad como un atributo negativo. Teniendo en cuenta que la mayoría de los inmuebles se encuentran relativamente cerca a estos puntos de interés, es posible que los vendedores asocien una mayor densidad de paradas con una mayor congestión, ruido y tráfico peatonal, lo que podría influir en su disposición a

asignar precios más bajos a las propiedades ubicadas en áreas con mayor concentración de paradas de subte y bus. En contraste, la cantidad de paradas de tren en un radio de 1000m no parece tener un efecto significativo en los precios de oferta según este modelo.

Es importante tener en cuenta que este modelo no captura posibles no linealidades en la relación entre las distancias y los precios de oferta. Por ejemplo, podría haber umbrales de distancia en los que la relación cambia de positiva a negativa o viceversa. Los vendedores podrían valorar la proximidad a ciertos atributos hasta cierto punto, pero percibir distancias demasiado cercanas como menos deseables debido a factores como el ruido o la congestión. Del mismo modo, distancias muy grandes podrían hacer que el atributo sea menos relevante para los precios de oferta. Estas no linealidades no se capturan en este modelo log-log, que asume una relación constante entre los cambios porcentuales en las distancias y los cambios porcentuales en los precios de oferta.

Con el objetivo de obtener una perspectiva adicional sobre la relación entre las variables explicativas y el precio de los inmuebles, sobre todo de aquellas que pudiesen ser no lineales, se decidió explorar dos enfoques alternativos, ambos presentados en el Apéndice E.

Por un lado, se optó por la utilización de rezagos espaciales, con el objetivo de capturar la relación espacial del precio causada por la cercanía de los inmuebles. El modelo logró un buen ajuste, con un R-cuadrado de 0.841, lo que representa una mejora significativa con respecto al modelo log-log (0,796). Por su parte el rezago espacial es altamente significativo y tiene un gran impacto positivo. Esto indica una fuerte dependencia espacial entre el precio de una propiedad y el precio de sus diez vecinos más cercanos. A su vez, esto genera una reducción de la significancia de varias de las variables relacionadas con la ubicación del inmueble, ya que parte de su efecto sobre el precio es absorbido por el rezago espacial.

Finalmente, se ajustó un modelo aplicando el algoritmo de Random Forest para capturar las no linealidades de los parámetros. Mediante esta aplicación se logró una notable mejora en la capacidad predictiva en comparación con el modelo de regresión lineal básico, alcanzando un R2 ajustado de 0.86. Este modelo consiguió capturar relaciones complejas entre las variables, pudiendo incluir la latitud y la longitud como variables explicativas. Si bien su interpretación es más compleja que para los modelos lineales, el modelo de Random Forest presentó un mejor ajuste de los residuos, especialmente en cuanto las colas pesadas, como se muestra en el Gráfico 5 del Apéndice F. A pesar de su utilidad predictiva superior, puede tener menos aplicabilidad cuando se busca comprender la influencia individual de cada variable en el precio.

4. Verificación de las hipótesis de la investigación.

Realizando una comparación entre las hipótesis planteadas en el Capítulo I y las interpretaciones de los resultados del modelo de precios hedónicos, podemos evaluar si estas hipótesis se cumplen o no en el contexto del mercado inmobiliario de C.A.B.A. en diciembre de 2023.

La Hipótesis Sustantiva planteaba que los atributos estructurales, de accesibilidad, socioambientales y de comuna tienen un efecto significativo y diferenciado en el precio de venta de los inmuebles. Los resultados del modelo respaldan esta hipótesis, ya que se observan efectos significativos y variados de estos atributos en los precios de oferta. Esto confirma que los diferentes atributos considerados influyen de manera distinta en las expectativas de precios de los vendedores.

En cuanto a la Hipótesis Específica 1, que planteaba una relación positiva entre varios atributos estructurales y el precio de venta, y una relación negativa para la antigüedad, los resultados del modelo son mixtos. Si bien se confirma la relación positiva para el área cubierta, la presencia de piscina y el espacio de estacionamiento, así como la relación negativa para la antigüedad, no se encuentra un efecto significativo para las variables "Tiene jardín" y "Ambientes". Por lo tanto, esta hipótesis se cumple parcialmente.

La Hipótesis Específica 2, que proponía una relación positiva entre la cercanía a amenidades y el precio de los inmuebles, no se cumple en su totalidad según los resultados del modelo. Contrariamente a lo esperado, los vendedores parecen percibir la lejanía de centros de esparcimiento, hospitales y centros culturales como atributos ligeramente positivos. Sin embargo, la cercanía a autopistas sí se asocia con precios más altos, lo que respalda parcialmente esta hipótesis.

Respecto a la Hipótesis Específica 3, que planteaba un efecto positivo de la seguridad, calidad de vida y nivel cultural en el precio de las propiedades, los resultados son consistentes. La proporción de hogares con necesidades básicas insatisfechas y la tasa de robos y hurtos tienen un efecto negativo en los precios de oferta, lo que sugiere que los vendedores valoran la seguridad y la calidad de vida al establecer sus expectativas de precios. Por lo tanto, esta hipótesis se cumple en gran medida.

Finalmente, la Hipótesis Específica 4, que proponía una relación significativa entre los aspectos ambientales y escénicos y el precio de venta, se cumple parcialmente. Si bien la cercanía al Río de la Plata y la lejanía del Río Matanza tienen efectos opuestos en los precios de oferta, la calidad del aire no muestra un efecto significativo. Además, la cercanía a centros de tratamiento de residuos tiene un efecto negativo, lo que respalda la importancia de los aspectos ambientales en las expectativas de precios de los vendedores.

En resumen, las hipótesis planteadas en el Capítulo I se cumplen en diferentes grados según los resultados del modelo de precios hedónicos. Mientras que la Hipótesis Sustantiva 1 y la Hipótesis Específica 3 encuentran un fuerte respaldo, las Hipótesis Específicas 1, 2 y 4 se cumplen parcialmente, con algunos resultados contrarios a lo esperado. Estos hallazgos resaltan la complejidad del mercado inmobiliario y la necesidad de considerar múltiples factores al analizar las expectativas de precios de los vendedores en C.A.B.A.

IV. CONCLUSIONES

La presente investigación se propuso analizar en profundidad los factores que determinan los precios de oferta de los inmuebles en la Ciudad Autónoma de Buenos Aires (C.A.B.A.) en diciembre de 2023. A través de la aplicación de la metodología de precios hedónicos y la incorporación de "valores no de mercado" como variables explicativas, se buscó determinar la relación entre el precio de oferta y una amplia gama de atributos estructurales, ambientales, comunitarios y de accesibilidad.

Los resultados obtenidos respaldan en gran medida las hipótesis planteadas. Se encontró que los atributos estructurales, ambientales, comunitarios y de accesibilidad tienen un efecto significativo y diferenciado sobre los precios de oferta de los inmuebles en C.A.B.A. En particular, se destacó la influencia de variables estructurales como el área cubierta, la presencia de piscina y si el inmueble es una casa o departamento. En cuanto a los atributos ambientales, la cercanía al Río de la Plata y al Río Matanza mostraron un impacto significativo en los precios. Asimismo, variables comunitarias como la población por viviendas y el porcentaje de hogares con necesidades básicas insatisfechas resultaron ser factores relevantes.

Un análisis comparativo de los distintos tipos de atributos sugiere que los estructurales tienen una mayor influencia sobre el precio de venta, seguidos por los comunitarios y los escénicos. Por su parte, las variables de accesibilidad parecen tener una influencia menor según el modelo utilizado.

No obstante, es importante reconocer las limitaciones del estudio. Futuras investigaciones podrían considerar las siguientes sugerencias:

- Siguiendo a Hoyt (1934), enfocar el análisis en unidades geográficas más pequeñas, como barrios o comunas, para capturar de manera más precisa las dinámicas locales del mercado inmobiliario.
- Utilizar modelizaciones no paramétricas, especialmente para las variables de accesibilidad, con el fin de captar posibles efectos no lineales entre las variables explicativas y los precios de oferta.
- Incluir variables estructurales que reflejen la calidad de los materiales de construcción, ya que esta característica puede influir significativamente en el valor de los inmuebles.
- Extender el análisis a otros períodos de tiempo y regiones geográficas, con el objetivo de evaluar la estabilidad de los resultados y la posible influencia de factores coyunturales.

Los hallazgos de esta investigación tienen implicancias relevantes tanto para el ámbito académico como para los actores del mercado inmobiliario. Los modelos desarrollados pueden servir como herramientas de apoyo para la toma de decisiones de inversión y la valuación de propiedades.

Finalmente, los autores desean expresar su más sincero agradecimiento a todas las personas que hicieron posible esta investigación. Al director de tesis, por su guía, su paciencia y apoyo constante. A los familiares, amigos y compañeros, por su comprensión y aliento a lo largo de este proceso. Y a los lectores, por su interés y disposición a sumergirse en las páginas de este estudio. Esperamos que los resultados y reflexiones aquí presentados contribuyan al avance del conocimiento en el campo de la economía espacial y la valuación de bienes raíces, y sirvan de inspiración para futuras investigaciones.

V. BIBLIOGRAFÍA

- Alberini, A., & Krupnick, A. (2000). Cost-of-illness and willingness-to-pay estimates of the benefits of improved air quality: Evidence from Taiwan. *Land Economics*, 76(1).
- Bajari, P., & Benkard, C. L. (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. *Journal of Political Economy*, 113(6).
- Bateman, I. J., & Langford, I. H. (1997). Non-users' willingness to pay for a National Park: An application and critique of the contingent valuation method. *Regional Studies*, 31(6).
- Bishop, K. C., & Timmins, C. (2019). Estimating the marginal willingness to pay function without instrumental variables. *Journal of Urban Economics*, 109.
- Brueggeman, W., & Fisher, J. D. (2022). Real Estate Finance & Investments. In *Real Estate Finance & Investments* (Vol. 1).
- Buteikis, A. (2020). Practical econometrics and data science. Vilnius University, 06–12.
- Can, A., & Megbolugbe, I. (1997). Spatial Dependence and House Price Index Construction. *Journal of Real Estate Finance and Economics*, 14(1–2).
- Carson, R. T., & Mitchell, R. C. (1993). The Value of clean water: The public's willingness to pay for boatable, fishable, and swimmable quality water. *Water Resources Research*, 29(7).
- Chasin, S. H., Tolley, G., Kenkel, D., & Fabian, R. (1995). Valuing Health for Policy: An Economic Approach. *Southern Economic Journal*, 62(2).
- Clapp, J. M. (1980). The Intrametropolitan Location of Office Activities. *Journal of Regional Science*, 20(3).
- Costanza, R., D'Arge, R., De Groot, R., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'Neill, R. V., Paruelo, J., Raskin, R. G., Sutton, P., & Van Den Belt, M. (1997). The value of the world's ecosystem services and natural capital. *Nature*, 387(6630).
- Court, A. T. (1939). Hedonic Price Indexes with Automotive Examples. *The Dynamics of Automobile Demand*.
- Cruces, G., Ham, A., & Tetaz, M. (2008). Quality of Life in Buenos Aires Neighborhoods: Hedonic Price Regressions and the Life Satisfaction Approach.
- Domínguez Prost, E. R. (2019). Construcción de un índice de precios inmobiliario para la Ciudad Autónoma de Buenos Aires a partir de datos espaciales. www.permanenttsbgroup.ie
- Ekeland, I., Heckman, J. J., & Nesheim, L. (2004). Identification and estimation of hedonic models. *Journal of Political Economy*, 112(1).

- Freeman, A. M. (1979). Hedonic Prices, Property Values and Measuring Environmental Benefits: A Survey of the Issues. *The Scandinavian Journal of Economics*, 81(2).
- Griliches. (2014). Hedonic Price Indexes for Automobiles: An Econometric Analysis of Quality Change. In *Price Indexes and Quality Change*.
- Gujarati, D., & Porter, D. (2008). *Basic Econometrics* (fifth edition). McGraw-hill.
- Gyourko, J., & Tracy, J. (1991). The Structure of Local Public Finance and the Quality of Life. *Journal of Political Economy*, 99(4).
- Ham González, A. (2009). La calidad de vida en los barrios de buenos aires: estimaciones hedónicas de la valuación de los amenities urbanos y su distribución espacial. Facultad de Ciencias Económicas San Andrés.
- Hernández Sampieri, R. (2014). *Metodología de la Investigación*. Editorial McGraw Hill.
- Hirst, F. W., & Hurd, R. M. (1904). Principles of City Land Values. *The Economic Journal*, 14(55).
- Hoyt, H. (1934). One Hundred Years of Land Values in Chicago. *Social Service Review*, 8(4).
- Instituto de Estadísticas y Censos de Argentina. (2010). Censo 2010 Argentina: Resultados. In INDEC. Censo Nacional de Población, Hogares y Viviendas 2010. Instituto Geográfico Nacional (IGN) e INDEC, datos de superficie.
- Instituto de Estadísticas y Censos de Argentina. (2022). Censo 2022 Argentina: Resultados. In INDEC. Censo Nacional de Población, Hogares y Viviendas 2022. Instituto Geográfico Nacional (IGN) e INDEC, datos de superficie.
- Kiel, K. A., & Zabel, J. E. (1996). House price differentials in U.S. Cities: Household and neighborhood racial effects. *Journal of Housing Economics*, 5(2).
- Kinnard, W. N. (2003). New Thinking in Appraisal Theory. In C. F. Sirmans & E. M. Worzala (Eds.), *Essays in Honor of William N. Kinnard, Jr* (pp. 147–159). Springer US.
- Krutilla, J. V., & Fisher, A. C. (1975). The Economics of Natural Environments: Studies in the Valuation of Commodity and Amenity Resources. In *The Economics of Natural Environments: Studies in the Valuation of Commodity and Amenity Resources*.
- Kuhn, T. S. (Thomas S. (1968). The structure of scientific revolutions.
- Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, 74(2).
- Malpezzi, S. (2008). Hedonic Pricing Models: A Selective and Applied Review. In *Housing Economics and Public Policy*. <https://doi.org/10.1002/9780470690680.ch5>

- Mankiw, G. (1997). *Principles of Economics* (2a ed.). Editorial McGraw Hill.
- Marciano, A., & Ramello, G. (2020). *Non-market Valuation*. *Encyclopedia of Law and Economics*. Springer, New York, NY.
- Mark, J., & Goldberg, M. A. (1988). Multiple regression analysis and mass assessment: A review of the issues. *Appraisal Journal*, 56.
- Marshall, A. (1890). *Principles of Economics*. In *Principles of Economics*.
- Ministerio de Justicia y Seguridad, S. de S. C. e I. C. (2022). Informe complementario al Mapa del Delito 2022.
- Navrud, S., & Ready, R. C. (2003). Valuing Cultural Heritage: Applying Environmental Valuation Techniques to Historic Buildings, Monuments and Artifacts, 2002. *Journal of Cultural Economics*, 27.
- Nkolika, J. P., Okagbue, H. I., Obas, E. C. M., & Akinola, A. O. (2020). Review on the Application of Artificial Neural Networks in Real Estate Valuation. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(3).
- Pollicino, M., & Maddison, D. (2018). Valuing the benefits of cleaning Lincoln Cathedral. In *The Stated Preference Approach to Environmental Valuation: Volume III: Applications: Benefit-Cost Analysis and Natural Resource Damage Assessment*.
- Potrawa, T. (2020). Hedonic Pricing Model for Rotterdam Housing Market Empowered by Image Recognition and Text Analysis Methods Student name.
- Ricardo, D. (1817). *The Principles of Political Economy and Taxation*.
- Ridker, R. G., & Henning, J. A. (1967). The Determinants of Residential Property Values with Special Reference to Air Pollution. *The Review of Economics and Statistics*, 49(2).
- Rose, S. K. (1999). Non-Market Valuation Techniques: The State of the Art. In *Environmental and Resource Economics*.
- Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Tetrahedron Letters*, 27(50).
- Shiller, R. J. (2016). *Irrational Exuberance*. Princeton University Press.
- Sigman, M. (2023). *El poder de las palabras*. Debate.
- The University of Chicago Press. (1927). The Appraisal of Real Estate. Frederick M. Babcock. *Journal of Political Economy*, 35(2), 319–320.

Thünen, J. H. von. (1826). Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie. In
Der isolierte Staat in Beziehung auf Landwirtschaft und Nationalökonomie.

Zoloa, J. I. (2015). Tópicos en el mercado inmobiliario. Facultad de Ciencias Universidad de San Andrés.

APÉNDICES

APÉNDICE A. EXPLICACIÓN DEL ÍNDICE UAQI	68
APÉNDICE B. EJERCICIO DE CÁLCULO DE RADIO DE BUFFERS PARA INDICADORES DE ACCESIBILIDAD.....	69
APÉNDICE C. SUPUESTOS DEL MODELO CLÁSICO DE REGRESIÓN LINEAL.....	70
APÉNDICE D. RESUMEN DE RESULTADOS DE MODELOS ESTIMADOS.....	71
APÉNDICE E. EXTENSIONES AL MODELO DE PRECIOS HEDÓNICOS	74
APÉNDICE F. RESULTADOS DE PRUEBAS ACCESORIAS AL ANÁLISIS	77
APÉNDICE G. RESUMEN COMPARATIVO DE LOS RESULTADOS DE LOS MODELOS	81
APÉNDICE H. COMPARACIÓN DE PARTICIPACIÓN DE INMUEBLES POR COMUNA	82

Apéndice A. Explicación del índice UAQI

La API de Google Air Quality permite solicitar datos sobre la calidad del aire para una ubicación específica, incluidos más de 70 índices de calidad del aire ("AQI"), contaminantes y recomendaciones de salud. Abarca más de 100 países con una resolución de 500 por 500 metros.

En base al índice UAQI, la API genera los siguientes rangos categóricos, que fueron utilizados para la inclusión de fenómenos ambientales en el modelo:

Tabla 5: Rango de valores y categorías de calidad del aire según API de Google Air Quality

Rango de valores	Categoría según API en base a UAQI
80 a 100	Calidad del aire excelente
60 a 79	Calidad del aire buena
40 a 59	Calidad del aire moderada
20 a 39	Calidad del aire baja
1 a 19	Calidad del aire mala

Fuente: Google Air Quality API (diciembre 2023)

Para una mayor comprensión el índice UAQI tiene en cuenta los siguientes contaminantes¹⁵:

- **Partículas en suspensión (PM):** Las PM son un indicador común para la contaminación del aire. Existe una sólida evidencia sobre los impactos negativos en la salud asociados con la exposición a este contaminante. Los principales componentes de las PM son sulfatos, nitratos, amoníaco, cloruro de sodio, carbono negro, polvo mineral y agua. A su vez, el índice de UAQI considera a las PM₁₀ son partículas gruesas que pueden penetrar en las vías respiratorias superiores y los pulmones, mientras que las PM_{2.5} son partículas finas que pueden penetrar profundamente en los pulmones y el sistema circulatorio, lo que las hace más perjudiciales para la salud.
- **Monóxido de carbono (CO):** El monóxido de carbono es un gas tóxico incoloro, inodoro e insípido producido por la combustión incompleta de combustibles carbonosos como madera, gasolina, carbón vegetal, gas natural y queroseno.
- **Ozono (O₃):** El ozono a nivel del suelo – que no debe confundirse con la capa de ozono en la atmósfera superior – es uno de los principales componentes del smog fotoquímico y se forma a través de la reacción con gases en presencia de luz solar.
- **Dióxido de nitrógeno (NO₂):** El dióxido de nitrógeno es un gas que se libera comúnmente por la combustión de combustibles en los sectores del transporte y la industria.
- **Dióxido de azufre (SO₂):** El dióxido de azufre es un gas incoloro con un olor penetrante. Se produce por la quema de combustibles fósiles (carbón y petróleo) y la fundición de minerales que contienen azufre.

¹⁵ Según la Organización Mundial de la Salud en: <https://www.who.int/teams/environment-climate-change-and-health/air-quality-and-health/health-impacts/types-of-pollutants>

Apéndice B. Ejercicio de cálculo de radio de buffers para indicadores de accesibilidad

Como se anteriormente, los indicadores de accesibilidad que tienen en cuenta el acceso a diferentes medios de transporte público se calcularon utilizando *buffers* con radios de r metros basados en percentiles de distancia, seleccionando el óptimo según el criterio de Akaike (AIC).

En esta instancia no se había escogido la forma funcional del modelo por lo que se generó una regresión de modelo logarítmico y con todas las variables disponibles como variables independientes; habiéndose estandarizado aquellas no dicotómicas. Se escogió el radio del buffer en metros para el cual se obtuvo el menor AIC para cada tipo de variable.

A continuación, se presentan los resultados de los radios en cuestión por tipo de medio de transporte y a sitios de interés:

Tabla 6: Elección de buffers por tipo de transporte y a sitios de interés (valores en metros)

Percentil	Colectivo	Tren	Subte	Sitios de interés
25%	36	632	354	456
50%	85	1049	743	997
75%	132	1605	1890	1659
Buffer 1	40	600	400	500
Buffer 2	90	1000	700	1000
Buffer 3	130	1600	1900	1700
AIC (buffer 1)	1689	1713	1607	1693
AIC (buffer 2)	1695	1708	1654	1715
AIV (buffer 3)	1687	1711	1716	1716

A raíz de los resultados, se escogieron los siguientes radios en metros para cada uno de los buffers en cuestión:

Tabla 7: Resumen de radios elegidos para cada buffer

Colectivo	Tren	Subte	Sitios de interés
130 metros	1000 metros	400 metros	500 metros

Apéndice C. Supuestos del modelo clásico de regresión lineal

Gujarati y Porter (2008) exponen los supuestos del modelo clásico de regresión lineal de la siguiente manera:

Supuesto 1: Linealidad en los parámetros

El modelo de regresión es lineal en los parámetros, aunque puede o no ser lineal en las variables.

Supuesto 2: Independencia entre las variables explicativas y el término de error

Los valores de la(s) variable(s) explicativa(s) X pueden considerarse fijos en muestras repetidas (regresora fija) o haber sido muestreados junto con la variable dependiente Y (regresora estocástica). En el segundo caso, se supone que la(s) variable(s) X y el término de error son independientes, es decir: $cov(X_i, \mu_i) = 0$; donde cov significa covarianza.

Supuesto 3: Media condicional cero del término de error:

Dado el valor de X_i , la media o el valor esperado del término de perturbación aleatoria μ_i es cero. Simbólicamente: $E(\mu_i|X_i) = 0$ o si X no es estocástica: $E(\mu_i) = 0$

Supuesto 4: Homoscedasticidad.

La varianza del término de error es constante para todos los valores de X . Es decir:

$var(\mu_i) = E[\mu_i - E(\mu_i|X_i)]^2 = E(\mu_i^2|X_i) = E(\mu_i^2) = \sigma^2$ si X_i son variables no estocásticas; donde var significa varianza.

Supuesto 5: No autocorrelación

Dados dos valores cualesquiera de X , X_i y X_j ($i \neq j$), la correlación entre dos μ_i y μ_j cualesquiera ($i \neq j$) es cero. Es decir, las observaciones se muestrean de manera independiente. Simbólicamente: $cov(\mu_i, \mu_j|X_i, X_j) = 0$ o $cov(\mu_i, \mu_j) = 0$ si X no es estocástica.

Supuesto 6: Muestra suficiente

El número de observaciones n debe ser mayor que el número de parámetros a estimar, es decir, mayor que el número de variables explicativas.

Supuesto 7: Variabilidad en las variables explicativas

No todos los valores X en una muestra determinada deben ser iguales. Técnicamente, $var(X)$ debe ser un número positivo. Además, no puede haber valores atípicos de la variable X , es decir, valores muy grandes en relación con el resto de las observaciones

Apéndice D. Resumen de resultados de modelos estimados

Tabla 8: Resumen de resultados del modelo hedónico inicial (lin-lin)

Nombre de la variable	Coficiente	Error estándar	t	P> t	[0.025	0.975]
Constante	-210.000,00	16.400,00	-12,82	0,000	-242.000,00	-178.000,00
Es casa	-58.760,00	3.042,91	-19,31	0,000	-64.700,00	-52.800,00
Espacio de estacionamiento	30.030,00	1.057,98	28,38	0,000	28.000,00	32.100,00
Tiene piscina	28.360,00	1.205,28	23,53	0,000	26.000,00	30.700,00
Baños	15.260,00	1.082,91	14,09	0,000	13.100,00	17.400,00
Ambientes	9.287,40	1.419,56	6,54	0,000	6.504,87	12.100,00
Tiene jardín	3.385,26	1.962,64	1,73	0,085	-461,79	7.232,31
Área cubierta (m2)	1.220,20	35,84	34,05	0,000	1.149,95	1.290,45
Dormitorios	-724,14	1.462,72	-0,50	0,621	-3.591,28	2.143,00
Antigüedad	-685,45	22,83	-30,03	0,000	-730,19	-640,70
Área total (m2)	284,22	25,38	11,20	0,000	234,47	333,98
Calidad del aire (UAQI)	-81,69	78,87	-1,04	0,300	-236,28	72,91
Distancia tratamiento residuos (m)	9,05	0,52	17,35	0,000	8,02	10,07
Distancia parque (m)	-5,54	2,56	-2,17	0,030	-10,56	-0,53
Distancia puerto (m)	-2,47	0,67	-3,69	0,000	-3,78	-1,16
Distancia Río de la Plata (m)	-1,72	0,48	-3,58	0,000	-2,66	-0,78
Distancia Río Matanza (m)	-1,07	0,56	-1,93	0,053	-2,16	0,02
Cuidado consumo eléctrico	194.000,00	8.574,04	22,62	0,000	177.000,00	211.000,00
Población por viviendas	64.020,00	6.144,82	10,42	0,000	52.000,00	76.100,00
Hogares NBI	-4.099,74	247,54	-16,56	0,000	-4.584,96	-3.614,53
Robos y hurtos	-3,01	1,44	-2,09	0,036	-5,83	-0,19
Paradas subte (400 m)	-10.220,00	747,70	-13,67	0,000	-11.700,00	-8.751,60
Paradas tren (1000 m)	-1.453,49	823,04	-1,77	0,077	-3.066,77	159,78
Paradas bus (130 m)	-1.155,79	192,51	-6,00	0,000	-1.533,14	-778,44
Sitios de interés (500 m)	-835,83	99,39	-8,41	0,000	-1.030,64	-641,02
Distancia escuela (m)	58,43	4,57	12,78	0,000	49,47	67,40
Distancia centro esparcimiento (m)	23,71	1,74	13,63	0,000	20,30	27,12
Distancia centro cultural (m)	7,95	4,26	1,87	0,062	-0,41	16,31
Distancia avenida (m)	7,93	3,88	2,04	0,041	0,32	15,55
Distancia hospital (m)	5,49	0,89	6,17	0,000	3,75	7,23
Distancia comisaría (m)	4,68	0,88	5,34	0,000	2,97	6,40
Distancia ruta (m)	-2,65	0,93	-2,85	0,004	-4,48	-0,83
Distancia autopista (m)	-0,14	0,76	-0,18	0,856	-1,63	1,35

Variable dependiente	Precio
Modelo	OLS
Método	Least Squares
Número de observaciones	13.840
Grados libertad residuos	13.807
Grados libertad modelo	32
Tipo de covarianza	nonrobust
R2	0,722
R2 ajustado	0,721
F-statistic	1.119
Prob (F-statistic)	0
Log-Likelihood	-3,49E+02
AIC	338.400
BIC	338.700
Asimetría	2,82
Curtosis	33,36
Durbin-Watson	1,51
Jarque-Bera (JB)	549.966
Prob(JB)	0
Breusch-Pagan test statistic	1532,20
Breusch-Pagan test p-value	2,78E-302

Tabla 9: Resumen de resultados del segundo modelo hedónico (log-log)

Nombre de la variable	Coficiente	Error estándar	z	P> z	[0.025	0.975]
Constante	8,9120	0,150	59,420	0,000	8,618	9,206
Área cubierta (m2)	0,8765	0,012	74,845	0,000	0,854	0,899
Espacio de estacionamiento	0,2650	0,008	31,927	0,000	0,249	0,281
Tiene piscina	0,1577	0,006	25,652	0,000	0,146	0,170
Es casa	-0,1076	0,013	-7,987	0,000	-0,134	-0,081
Ambientes	0,0833	0,015	5,645	0,000	0,054	0,112
Antigüedad	-0,0668	0,002	-44,343	0,000	-0,070	-0,064
Tiene jardín	0,0111	0,011	1,033	0,302	-0,010	0,032
Distancia Río de la Plata (m)	-0,1296	0,005	-25,817	0,000	-0,139	-0,120
Distancia puerto (m)	-0,0720	0,013	-5,460	0,000	-0,098	-0,046
Distancia Río Matanza (m)	0,0572	0,008	6,878	0,000	0,041	0,073
Distancia tratamiento residuos (m)	0,0370	0,004	8,476	0,000	0,028	0,046
Calidad del aire (UAQI)	0,0342	0,028	1,235	0,217	-0,020	0,089
Distancia parque (m)	-0,0029	0,002	-1,328	0,184	-0,007	0,001
Población por viviendas	-0,3398	0,037	-9,146	0,000	-0,413	-0,267
Hogares NBI	-0,0952	0,006	-15,630	0,000	-0,107	-0,083
Robos y hurtos	-0,0030	0,005	-0,573	0,567	-0,013	0,007
Paradas subte (400 m)	-0,0758	0,007	-10,793	0,000	-0,090	-0,062
Distancia centro esparcimiento (m)	0,0303	0,004	8,640	0,000	0,023	0,037
Distancia ruta (m)	0,0282	0,006	5,123	0,000	0,017	0,039
Paradas bus (130 m)	-0,0264	0,003	-7,630	0,000	-0,033	-0,020
Sitios de interés (500 m)	-0,0235	0,004	-6,176	0,000	-0,031	-0,016
Distancia autopista (m)	-0,0232	0,005	-4,234	0,000	-0,034	-0,012
Distancia hospital (m)	0,0175	0,003	5,515	0,000	0,011	0,024
Distancia centro cultural (m)	0,0113	0,003	3,530	0,000	0,005	0,018
Distancia comisaría (m)	0,0074	0,003	2,273	0,023	0,001	0,014
Distancia escuela (m)	0,0054	0,003	1,608	0,108	-0,001	0,012
Paradas tren (1000 m)	0,0043	0,007	0,658	0,511	-0,008	0,017
Distancia avenida (m)	-0,0015	0,002	-0,943	0,346	-0,005	0,002

Variable dependiente	Precio
Modelo	GLS
Método	Least Squares
Número de observaciones	13840
Grados libertad residuos	13811
Grados libertad modelo	28
Tipo de covarianza	HCO
R	0,797
R ajustado	0,796
F-statistic	1,778
Prob (F-statistic)	0
Log-Likelihood	-237,06
AIC	532,1
BIC	750,6
Asimetría	-0,045
Curtosis	3,998
Durbin-Watson	1,521
Jarque-Bera (JB)	579,6
Prob(JB)	1,36E-126
Breusch-Pagan test statistic	684,83
Breusch-Pagan test p-value	2,90 E-126

Tabla 10: Resumen de resultados del modelo hedónico con variable rezago espacial (log-log) - MCG

Nombre de la variable	Coficiente	Error estándar	z	P> z	[0.025	0.975]
Constante	3,588	0,162	22,177	0,000	3,271	3,905
Espacio de estacionamiento	0,223	0,007	30,280	0,000	0,208	0,237
Área cubierta (m2)	0,160	0,017	9,652	0,000	0,128	0,192
Ambientes	0,136	0,013	10,703	0,000	0,111	0,161
Tiene piscina	0,103	0,005	19,533	0,000	0,093	0,114
Antigüedad	-0,051	0,001	-37,925	0,000	-0,054	-0,049
Es casa	-0,048	0,012	-3,917	0,000	-0,072	-0,024
Tiene jardín	0,031	0,009	3,298	0,001	0,012	0,049
Distancia puerto (m)	-0,042	0,011	-3,854	0,000	-0,063	-0,020
Distancia Río de la Plata (m)	-0,025	0,004	-6,069	0,000	-0,033	-0,017
Distancia tratamiento residuos (m)	0,020	0,004	5,584	0,000	0,013	0,027
Distancia Río Matanza (m)	0,016	0,007	2,309	0,021	0,002	0,029
Calidad del aire (UAQI)	0,005	0,025	0,181	0,857	-0,044	0,053
Distancia parque (m)	-0,001	0,002	-0,529	0,597	-0,005	0,003
Población por viviendas	-0,128	0,031	-4,145	0,000	-0,188	-0,067
Hogares NBI	-0,035	0,005	-6,425	0,000	-0,045	-0,024
Robos y hurtos	-0,002	0,005	-0,445	0,656	-0,011	0,007
Paradas subte (400 m)	-0,033	0,006	-5,277	0,000	-0,045	-0,021
Distancia centro esparcimiento (m)	0,018	0,003	6,003	0,000	0,012	0,024
Distancia centro cultural (m)	0,006	0,003	2,208	0,027	0,001	0,012
Distancia comisaría (m)	-0,005	0,003	-1,560	0,119	-0,010	0,001
Distancia ruta (m)	0,004	0,005	0,863	0,388	-0,005	0,014
Distancia autopista (m)	-0,003	0,005	-0,641	0,521	-0,012	0,006
Paradas tren (1000 m)	0,002	0,006	0,337	0,736	-0,009	0,013
Sitios de interés (500 m)	0,002	0,003	0,569	0,569	-0,004	0,008
Distancia escuela (m)	-0,001	0,003	-0,484	0,629	-0,007	0,004
Distancia hospital (m)	-0,001	0,003	-0,276	0,783	-0,006	0,005
Paradas bus (130 m)	0,000	0,003	-0,136	0,892	-0,007	0,006
Distancia avenida (m)	0,000	0,001	0,028	0,978	-0,003	0,003
Rezago espacial	0,654	0,012	54,949	0,000	0,631	0,678

Variable dependiente	Price
Modelo	GLS
Método	Least Squares
Número de observaciones	13840
Grados libertad residuos	13810
Grados libertad modelo	29
Tipo de covarianza	HC0
R	0,841
R ajustado	0,841
F-statistic	2295
Prob (F-statistic)	0
Log-Likelihood	1464,2
AIC	-2868
BIC	-2642
Asimetría	-0,089
Curtosis	3,975
Durbin-Watson	1,771
Jarque-Bera (JB)	566,709
Prob(JB)	8,72E-124
Breusch-Pagan test statistic	378,79
Breusch-Pagan test p-value	1,44 E-62

Apéndice E. Extensiones al modelo de precios hedónicos

Es importante destacar que la aplicación de las siguientes extensiones al proceso de especificación del modelo en este estudio fue un análisis exploratorio realizado por curiosidad de los autores. Si bien se consideró valioso incluir estos enfoques alternativos para brindar una perspectiva adicional al lector, es fundamental aclarar que no está directamente relacionado con los objetivos principales de la tesis ni con la verificación de las hipótesis planteadas. Los siguientes dos modelos se presentan como un complemento interesante, pero no forman parte integral de la metodología diseñada para alcanzar los objetivos de la investigación.

a. Modelo log-log con rezago espacial

Una vez obtenido el modelo con la forma funcional log-log, se procedió a realizar la prueba de Morán para evaluar la presencia de autocorrelación espacial en los residuos del modelo. Esta prueba fue fundamental para garantizar que se cumplan los supuestos del modelo clásico de regresión lineal y obtener estimadores confiables. La prueba arrojó un estadístico con valor p igual a 0,001 lo cual respalda la presencia de autocorrelación espacial en el modelo. Esto indica que los errores del modelo no son independientes entre sí, sino que están correlacionados en función de la ubicación geográfica de los inmuebles.

En el contexto del mercado inmobiliario, la autocorrelación espacial implica que el precio de un inmueble no solo está influenciado por sus propias características, sino también por los precios de las propiedades vecinas. La dependencia espacial puede surgir por varias razones:

- **Efecto de vecindario:** Las viviendas dentro de un mismo vecindario tienden a ser similares en términos de accesibilidad, servicios públicos y calidad de escuelas, entre otros. Esto hace que los precios de las viviendas en un vecindario tiendan a estar correlacionados.
- **Efecto de difusión espacial:** Los vendedores pueden usar los precios de las viviendas vecinas como referencia para determinar cuánto están dispuestos a recibir por una vivienda en particular. Shiller (2016) menciona que al incorporar este tipo de variables es posible que el precio de las propiedades se desacople de su valor fundamental incorporando un componente de comportamiento irracional de los vendedores (o compradores) que podría desencadenar una tragedia como el de las hipotecas subprime en 2008.
- **Competencia espacial:** Las viviendas cercanas pueden ser sustitutos cercanos y, por lo tanto, sus precios pueden estar relacionados debido a la competencia entre ellas.

Para abordar el problema de la autocorrelación espacial en el modelo, se optó por incluir una variable de rezago espacial ("*spatial lag*"). Esta variable captura la influencia que tienen los precios de las k viviendas más cercanas sobre el precio de cada inmueble en particular. Al incorporar esta variable al modelo, se busca explicar explícitamente la dependencia espacial y obtener estimadores más precisos y confiables.

Para el cálculo de este rezago se forma una matriz de pesos espaciales, una matriz cuadrada en la que cada elemento d_{ij} de la matriz toma el valor uno (1) dependiendo si el inmueble i se considerado vecino (siguiendo un criterio determinado) del inmueble j , o cero (0) si no son considerados vecinos.

Existen varios criterios para la definición de los vecinos, entre los cuales se encuentran el criterio Queen, Rook y KNN. En este caso, se optó por utilizar el algoritmo K-Nearest Neighbors (“KNN”) que encuentra los k vecinos más cercanos, calculando la distancia entre inmuebles en base a la distancia euclidiana. En esta investigación se utilizó un valor de k igual a 10, es decir, se consideraron los 10 vecinos más cercanos a cada propiedad. La elección de este número se basó en la obtención del mayor R^2 ajustado con el modelo que incorporó dicho rezago, en comparación con distintos modelos con distintos valores de k . La Tabla 12 en el Apéndice F resume dicha elección.

Una vez obtenida esa matriz E de coeficientes espaciales, esta se multiplica por la variable precio sobre metro cuadrado. De esta forma se obtiene el promedio del precio sobre metro cuadrado de todos los vecinos definidos para cada propiedad.

Ya calculado dicho promedio para cada propiedad (rezago espacial), se lo multiplica por el área cubierta (en metros cuadrados) de la propiedad. Esto se justifica porque la variable que se intenta explicar es el precio del inmueble, no el precio del metro cuadrado.

El resultante de dicho producto es la variable explicativa que se incorpora en el modelo espacial. La Ecuación 8 presenta la forma funcional del modelo incorporando a la variable en cuestión.

Ecuación 8: Especificación del modelo de precios hedónicos con rezago espacial

$$\log(P_i) = \beta_0 + \beta_E * \log(E) + \beta_A * \log(A) + \beta_C * \log(C) + \beta_B * \log(B) + \beta_D * D + \beta_S * P/mc_{k,i} * mc_i + \mu_i$$

- P_i es el vector de la variable dependiente que representa el precio de cada inmueble en dólares.
- β_0 es el término de intercepto o constante del modelo.
- E es la matriz de variables explicativas estandarizadas no dicotómicas que incluye los atributos estructurales.
- A es la matriz de variables explicativas estandarizadas que incluye los atributos ambientales.
- C es la matriz de variables explicativas estandarizadas que incluye los atributos comunitarios.
- B es la matriz de variables explicativas estandarizadas que incluye los atributos de accesibilidad.
- D es la matriz de variables explicativas dicotómicas que incluyen atributos estructurales.
- $P/mc_{k,i}$ es el precio del metro cuadrado cubiertos de los k inmuebles más cercanos al inmueble i .
- mc_i es el área cubierta del inmueble i medida en metros cuadrados.

- $\beta_E, \beta_A, \beta_C$ y $\beta_B, \beta_D, \beta_S$ son los vectores de coeficientes de regresión correspondientes a cada matriz de atributos y de rezago espacial respectivamente.
- μ_i es el vector de términos de error estocástico o perturbaciones aleatorias.

El modelo que utiliza Mínimos Cuadrados Ordinarios (“MCO”) tiene un buen ajuste, con un R-cuadrado de 0.841, lo que representa una mejora significativa con respecto al modelo log-log (0,80). Por su parte el rezago espacial es altamente significativo y tiene un gran impacto positivo (0.67), lo que indica una fuerte dependencia espacial en los precios de las propiedades. A su vez, esto genera una reducción de la significancia de varias de las variables relacionadas con la ubicación del inmueble, ya que parte de su efecto sobre el precio es absorbido por el rezago espacial. Los resultados se presentaron en la Tabla 10 del Apéndice D.

b. Modelo de Random Forest

Como último complemento al proceso de especificación del modelo de precios hedónicos, se decidió explorar un enfoque alternativo utilizando el algoritmo de Random Forest con el objetivo de obtener una perspectiva adicional sobre la relación entre las variables explicativas y el precio de los inmuebles, sobre todo de aquellas que pudiesen ser no lineales.

Random Forest es un algoritmo de aprendizaje automático que combina múltiples árboles de decisión para realizar predicciones o clasificaciones. En el contexto de la modelación de precios hedónicos, Random Forest construye un conjunto de árboles de decisión utilizando submuestras aleatorias de las observaciones y variables explicativas. Cada árbol se entrena de manera independiente y realiza una predicción del precio basada en las características seleccionadas. La predicción final del modelo se obtiene promediando las predicciones de todos los árboles individuales. Esto tiene la ventaja de capturar relaciones no lineales e interacciones complejas entre las variables, lo que puede ser beneficioso en la modelación de precios inmobiliarios.

Su capacidad predictiva supera ampliamente a los modelos de regresión lineal básicos pero la interpretación de sus coeficientes es muy compleja por lo que suele tener menos utilidad cuando se busca entender como influyen las distintas variables en el precio.

En la presente investigación se ajustó un modelo de Random Forest de 200 árboles obteniendo buenos resultados en términos predictivos y un mayor ajuste de los residuos. Se presenta en el Apéndice F, Tabla 13 que resume la elección del número de árboles utilizado y los resultados obtenidos con este enfoque alternativo.

Este modelo logró capturar relaciones complejas entre las variables e incorporar la latitud y la longitud como variables explicativas. La capacidad predictiva mejora notablemente comparada con el resto de los modelos, logrando un R2 ajustado 0.86. También se puede observar en el Gráfico 5 en el Apéndice F una mejora en las colas pesadas con relación al modelo log-log.

Apéndice F. Resultados de pruebas accesorias al análisis

Tabla 11: Cálculo de GVIF para variables explicativas

Nombre de variable	GVIF
Ambientes	1,071531
Área cubierta (m2)	1,059596
Dormitorios	1,059061
Área total (m2)	1,054927
Baños	1,04094
Es casa	1,022197
Antigüedad	1,016723
Espacio de estacionamiento	1,007881
Tiene piscina	1,006937
Tiene jardín	1,004351
Distancia puerto (m)	1,086445
Calidad del aire (UAQI)	1,083401
Distancia Río Matanza (m)	1,074413
Distancia Río de la Plata (m)	1,063773
Distancia tratamiento residuos (m)	1,033878
Distancia parque (m)	1,019478
Población por viviendas	1,080272
Cuidado consumo eléctrico	1,06682
Hogares NBI	1,044809
Robos y hurtos	1,019091
Distancia ruta (m)	1,046042
Distancia autopista (m)	1,044085
Distancia centro esparcimiento (m)	1,025643
Distancia comisaría (m)	1,024253
Distancia centro cultural (m)	1,024124
Distancia hospital (m)	1,022931
Distancia escuela (m)	1,022383
Distancia avenida (m)	1,013575
Paradas tren (1000 m)	1,013535
Paradas bus (130 m)	1,011758
Paradas subte (400 m)	1,011658
Sitios de interés (500 m)	1,008214

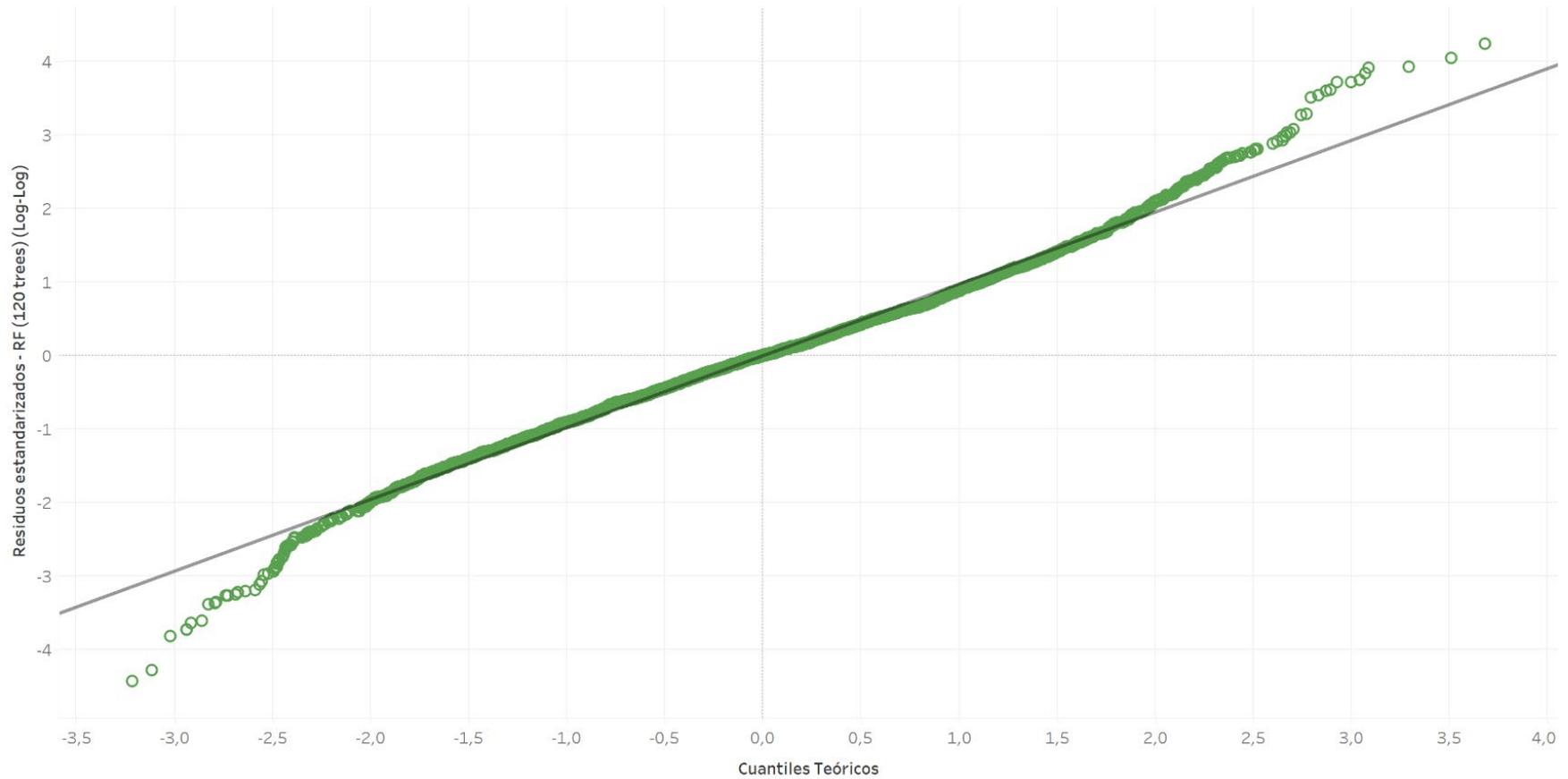
Tabla 12: Resultados de la elección de KNN para el modelo con rezago espacial

N° neighbors (k)	R-squared	Adj. R-squared	F-statistic	Prob (F-statistic)	AIC	Breusch-Pagan statistic	Prob (Breusch-Pagan)
1	0,823	0,822	2209,0	0,0	-1353,8	461,5	2,21E-79
3	0,835	0,835	2417,9	0,0	-2390,3	478,6	7,26E-83
5	0,839	0,839	2488,8	0,0	-2725,7	421,0	4,07E-71
10	0,841	0,841	2519,6	0,0	-2868,3	378,8	1,44E-62
15	0,840	0,840	2501,3	0,0	-2783,5	369,3	1,20E-60
20	0,840	0,840	2508,5	0,0	-2817,2	381,9	3,33E-63
25	0,841	0,840	2513,8	0,0	-2841,8	382,5	2,55E-63
30	0,841	0,840	2510,3	0,0	-2825,5	386,7	3,65E-64
40	0,840	0,840	2504,4	0,0	-2798,2	383,1	1,98E-63
50	0,839	0,839	2487,8	0,0	-2721,0	391,3	4,32E-65
100	0,835	0,835	2417,4	0,0	-2387,9	428,1	1,49E-72
200	0,831	0,831	2346,8	0,0	-2046,2	491,5	1,61E-85

Tabla 13: Elección de número de árboles y resultado de modelo de Random Forest

Trees	R-squared	Adj. R-squared	Mean Squared Error (MSE)	Breusch-Pagan statistic	Prob (Breusch-Pagan)
5	0,830	0,828	0,0519	220,93	1,54E-30
25	0,854	0,853	0,0444	202,10	5,28E-27
50	0,857	0,856	0,0437	219,02	3,54E-30
100	0,857	0,856	0,0434	219,74	2,58E-30
150	0,858	0,857	0,0432	222,52	7,73E-31
200	0,859	0,858	0,0432	221,36	1,28E-30
300	0,858	0,857	0,0431	222,04	9,52E-31

Gráfico 5: Cuantil-cuantil de residuos del modelo de Random Forest



Apéndice G. Resumen comparativo de los resultados de los modelos**Tabla 14: Comparación de bondad de ajuste de los modelos propuestos**

Modelo	R cuadrado ajustado
lin-lin (MCO)	0,72
log-log (MCG)	0,80
log-log con rezago espacial (MCG)	0,84
Random Forest	0,86

Apéndice H. Comparación de participación de inmuebles por comuna

Tabla 15: Comparación de participación de los inmuebles de la muestra vs los inmuebles registrados en el Censo de 2022.

Comuna	Muestra Inicial				Censo			
	Inmuebles comuna/ Inmuebles total	N° inmuebles	N° casas	N° deptos.	Inmuebles comuna/ Inmuebles total	N° inmuebles	N° casas	N° deptos.
COMUNA 1	11%	4.339	70	4.269	8%	104.984	14.131	80.674
COMUNA 2	5%	2.015	22	1.993	6%	83.038	1.745	80.171
COMUNA 3	5%	2.025	66	1.959	7%	91.391	6.803	76.741
COMUNA 4	4%	1.726	271	1.455	6%	90.018	33.894	49.814
COMUNA 5	5%	1.802	146	1.656	7%	92.166	11.954	76.327
COMUNA 6	3%	1.164	166	998	7%	94.566	10.311	81.963
COMUNA 7	4%	1.710	304	1.406	6%	89.905	25.190	60.820
COMUNA 8	6%	2.388	376	2.012	5%	69.286	37.585	27.856
COMUNA 9	7%	2.617	828	1.789	5%	65.918	35.017	29.285
COMUNA 10	5%	1.886	424	1.462	5%	73.305	31.193	40.273
COMUNA 11	6%	2.492	638	1.854	6%	86.950	32.842	52.699
COMUNA 12	7%	2.789	368	2.421	8%	107.094	32.023	73.538
COMUNA 13	11%	4.339	314	4.025	9%	126.307	12.369	112.534
COMUNA 14	11%	4.412	116	4.296	9%	126.532	6.720	117.637
COMUNA 15	7%	2.706	340	2.366	6%	89.798	22.598	64.964
Total	100%	38.410	4.449	33.961	100%	1.391.258	314.375	1.025.296

DECLARACIÓN JURADA RESOLUCIÓN 212/99 CD

Los autores de este trabajo declaran que fue elaborado sin utilizar ningún otro material que no haya dado a conocer en las referencias que nunca fue presentado para su evaluación en carreras universitarias y que no transgrede o afecta los derechos de terceros.

Mendoza, 10 de abril de 2024.



Guillén, Juan Cruz
DNI N° 40.907.531
Legajo N°30.672



Squartini, Santiago
DNI N°42.478.345
Legajo N°31.129