

# METODOLOGÍA PARA LA ESTIMACIÓN DE LAS TASAS DE CAPITALIZACIÓN DE RENTAS DE INMUEBLES RESIDENCIALES, RESULTADOS PARA BOGOTÁ 2017-2020

Unidad Administrativa Especial de Catastro Distrital

Diego Fernando Carrero Barón<sup>\*</sup>      Sergio Enrique Acosta Moreno<sup>\*\*</sup>  
Paula Andrea Mahecha Mahecha<sup>\*\*\*</sup>      Camilo Andrés Avellaneda García<sup>\*\*\*\*</sup>

30 de Junio de 2021

## Índice

<b>1. Introducción</b>	<b>2</b>
<b>2. Objetivos</b>	<b>4</b>
2.1. Objetivo general . . . . .	4
2.2. Objetivos específicos . . . . .	4
<b>3. Marco conceptual</b>	<b>4</b>
3.1. Universo de estudio . . . . .	4
3.2. Población objetivo . . . . .	4
3.3. Conceptos principales . . . . .	5
3.4. Antecedentes . . . . .	7
3.5. Controles de calidad sobre la base . . . . .	7

---

<sup>\*</sup> Jefe Observatorio Técnico Catastral

<sup>\*\*</sup> Coordinador Grupo Estadístico

<sup>\*\*\*</sup> Observatorio Técnico Catastral, Equipo Técnico de Investigación

<sup>\*\*\*\*</sup> Grupo Estadístico, Equipo Técnico de Investigación

<b>4. Modelo aplicado</b>	<b>11</b>
4.1. Apartamentos . . . . .	13
4.2. Casas . . . . .	20
<b>5. Conclusiones</b>	<b>28</b>
<b>6. Trabajo futuro</b>	<b>28</b>
<b>7. Anexo 1</b>	<b>28</b>
7.1. Modelo lineal mixto . . . . .	28

## 1. Introducción

La Unidad Administrativa Especial de Catastro Distrital -UAECD tiene por objeto responder por la recopilación e integración de la información georreferenciada de la propiedad inmueble del Distrito Capital en sus aspectos físico, jurídico y económico. Frente a esto, durante los últimos años la entidad ha enfocado esfuerzos para obtener datos del mercado inmobiliario que le permita apoyar el proceso de actualización catastral, afianzando el uso de métodos indirectos, y de manera general, para aportar conocimiento acerca del mismo.

La UAECD recolecta ofertas de mercado de venta o arriendo, a través de recorridos sectorizados en la ciudad, que constituyen una importante fuente primaria. De manera complementaria y no menos importante, la entidad también obtiene ofertas de fuentes secundarias, gestionadas a través de convenios o acuerdos con entidades del sector público o privado.

Particularmente, en ejercicios realizados al interior de la entidad, tendientes a estimar el valor integral de predios en propiedad horizontal, se han identificado dificultades relacionadas con la escasa o nula información en algunas locaciones de la ciudad. Para superar esta problemática, surge la propuesta de utilizar la información de ofertas de mercado para estimar las tasas de capitalización de renta (TCR).

Las TCR para inmuebles son índices (usualmente expresados en tasa) que relacionan el canon de arrendamiento y el valor de venta de activos inmobiliarios; son útiles en el campo de la valuación inmobiliaria porque permiten estimar el valor de mercado de un inmueble a partir del valor de arriendo o alquiler<sup>1</sup>.

Por lo anterior, este estudio tiene como objetivo principal proponer una metodología para la estimación de las TCR y hacer la estimación para inmuebles residenciales, según estratos y diferentes desagregaciones geográficas como lo son las localidades y sectores catastrales.

<sup>1</sup>En el sector inmobiliario, son un referente para medir el retorno potencial en aquellas inversiones enfocadas a la compra de un inmueble para el alquiler o cuando los inmuebles hacen parte de portafolios de inversión donde son los subyacentes. En una escala más amplia, la de ciudad, permiten hacer seguimiento a la dinámica urbana, dadas las diferencias que el índice puede presentar en términos de ubicación, usos, estratos y edades de los inmuebles que componen el inventario de la ciudad, análisis que se incorpora en un segundo documento con base en el resultado del presente documento.

En la información disponible para la realización de este documento, se encuentran las ofertas recolectadas<sup>2</sup> por la UAECD tanto de fuentes primarias como secundarias (Finca Raíz, Galería Inmobiliaria y Properati) desde el primero de enero de 2017 hasta el 31 de diciembre de 2020. Una limitante que se debe tener en cuenta es que los predios no tienen información de arriendo y venta de manera simultánea, lo que quiere decir que para cada oferta solamente se tiene una de las dos mediciones. En este documento únicamente se tendrán en cuenta las ofertas correspondientes a predios residenciales, es decir apartamentos y casas.

Dentro de la literatura se encuentran diversos ejemplos de datos donde se hacen mediciones o encuestas a grupos de individuos periódicamente, pero no necesariamente a las mismas personas, lo cual imposibilita hacer seguimientos sobre unidades en particular, de manera similar a la problemática presentada, donde no se cuenta con información de arriendo y venta en los mismos predios. [Deaton \[1985\]](#) propone un modelo para una muestra “*pseudo-panel*”, que se construye con lo que denominan cohortes, las cuales representan grupos de individuos con características similares, con la condición que cada individuo pertenezca únicamente a una cohorte durante todo el análisis y mediante estos grupos se ajustan efectos dentro de la regresión correspondientes a cada una de las diferentes cohortes. Muchas publicaciones que abordan diferentes contextos se han elaborado a partir del modelo propuesto por [Deaton \[1985\]](#) (Ver [Tsai et al. \[2014\]](#), [Tovar et al. \[2012\]](#), [Sprietsma \[2012\]](#), [Antman and McKenzie \[2007\]](#), entre otros). La propuesta utilizada en este documento se basa en el uso de estas cohortes, agregando la información a nivel de sector catastral, tomando el promedio o la mediana como medidas de resumen, tanto para valores de arriendo como para valores de venta.

Las metodologías de regresión ajustadas por mínimos cuadrados ordinarios tienen como supuestos la independencia entre mediciones, lo cual no se cumple en este caso, ya que un sector catastral puede tener información en las diferentes vigencias, por lo que se propone la inclusión de factores en la regresión que permitan captar el efecto de las medidas repetidas sobre la variable respuesta. De esta manera, la metodología a utilizar corresponde a los modelos de efectos mixtos, donde los efectos fijos van a ser el conjunto de factores asociados al estrato, localidad y el valor de arriendo, mientras que los efectos aleatorios corresponden a la relación entre el valor de arriendo y de venta, segmentada para cada uno de los sectores catastrales.

La programación realizada se realizó en el software R [[R Core Team, 2021](#)], en donde el procesamiento de información se llevó a cabo mediante los paquetes relacionados en el *tidyverse* [[Wickham et al., 2019](#)], mientras que los controles de calidad utilizaron funciones de la librería *qcc* y *RecordLinkage* (Ver [Scrucca \[2004\]](#) y [Sariyar and Borg \[2020\]](#)). La estimación de los parámetros del modelo se realizó mediante el paquete *nlme* [Pinheiro et al. \[2021\]](#).

La organización del documento es la siguiente: la primera sección corresponde a esta intro-

---

<sup>2</sup>Las ofertas contenidas en la base consolidada por la UAECD, son capturadas en campo u obtenidas de fuentes secundarias (como portales inmobiliarios), las cuales son dispuestas por particulares o agencias inmobiliarias que proporcionan un valor de arriendo o de alquiler e información de ciertos atributos del inmueble. Teniendo en cuenta esto, los precios objeto de análisis corresponden a precios de expectativa y no a precios de transacción efectiva. Por tal razón no son comparables con resultados de ejercicios que trabajan con información de precios de cierre (en el caso de las ventas) o de alquiler neto al cual se descuentan valores puntuales de administración, impuesto predial, entre otros.

ducción. La segunda sección presenta los objetivos del estudio. La tercera sección plantea el marco conceptual en el que se soporta el estudio. La cuarta sección describe de manera detallada la base de datos utilizada y los controles de calidad aplicados. En la quinta sección se presenta de manera general el modelo seleccionado y los resultados desagregados por tipo de inmueble (apartamentos o casas), con la correspondiente validación de supuestos. Y en la sexta sección, se incluyen las principales conclusiones. Finalmente, en el apéndice 1 contiene una descripción más detallada del modelo lineal mixto.

## **2. Objetivos**

Para el estudio, además del objetivo general, se definieron cinco objetivos específicos.

### **2.1. Objetivo general**

Proponer una metodología para la estimación de las TCR y hacer la estimación del índice para inmuebles residenciales de Bogotá, según estratos y diferentes desagregaciones geográficas como lo son las localidades y sectores catastrales.

### **2.2. Objetivos específicos**

1. Identificar metodologías que sirvan de referencia para el desarrollo del estudio.
2. Realizar un análisis descriptivo de la base de datos disponible para la realización del ejercicio.
3. Detectar aquellas ofertas atípicas que puedan sesgar los resultados del estudio.
4. Calcular las tasas de capitalización de rentas para inmuebles de uso residencial en diferentes desagregaciones: estrato, localidad y sector catastral
5. Presentar una documentación metodológica de los resultados

## **3. Marco conceptual**

### **3.1. Universo de estudio**

El Universo de estudio para este documento corresponde al conjunto de predios urbanos en la ciudad de Bogotá en los años 2017 al 2020.

### 3.2 Población objetivo

## 3.2. Población objetivo

La población objetivo corresponde a los predios residenciales, tanto de propiedad horizontal (PH) como no propiedad horizontal (NPH) en la ciudad de Bogotá comprendidos en los años 2017 al 2020.

## 3.3. Conceptos principales

De acuerdo con [Godinez \[2000\]](#), la tasa de capitalización se relaciona con el rendimiento de un capital invertido. Este autor indica que en “*el caso del negocio de arrendamiento de inmuebles, a la tasa de capitalización se le llama tasa de capitalización de rentas y se obtiene de la relación de la renta anual entre el valor comercial*”.

Con el fin de no confundir el concepto de TCR con la rentabilidad total de la inversión con destino al alquiler, es de precisar que el primero está incorporado en el segundo. En otras palabras, la rentabilidad total tiene dos componentes: i) los ingresos por alquiler (capturados a través de las TCR) y el incremento nominal derivado de la diferencia en precios del inmueble (denominado valorización<sup>3</sup>).

La expresión matemática de la que se deduce la TCR se muestra en la ecuación , donde  $P_k$  es el precio del  $k$ -ésimo inmueble,  $R_k$  es su tasa de renta o alquiler e  $i$  representa la respectiva TCR.

$$P_k = \frac{R_k}{i} \quad (1)$$

En cuanto al numerador, el alquiler, la tasa puede tomar en consideración la renta bruta del inmueble o la neta, esta última calculada a partir de los ingresos brutos menos los descuentos por concepto de vacíos y cobranzas, gastos de operación y cuotas de mantenimiento, y por descuentos del régimen tributario (Medrano, Perlaza y Fuentes 2011). Otros autores, como en [de Propiedad Raíz de Medellín \[2020\]](#), deducen el gasto de administración, predial, administración inmobiliaria, vacancia y mantenimiento.

Según [Borrero \[2008\]](#), el alquiler opera como “un factor de ajuste entre la valorización del inmueble y la tasa media de ganancia de la economía más un factor de riesgo y un factor de liquidez. En el caso de los inmuebles el factor de riesgo es casi nulo pero el factor de liquidez debe existir dado que no es fácil convertir el inmueble en dinero líquido”.

Respecto del denominador, el precio del inmueble es de precisar concepto de valor de mercado, el cual se define como el precio de venta más probable, o lo que un inversionista típico está dispuesto a pagar por una propiedad. Por lo cual, al momento de hacer la medición de las TCR, algunos autores como [Medrano J. Perlaza M. y Fuentes \[2011\]](#) descuentan del precio, un porcentaje de tasa de negociación.

Conceptualmente, de la misma manera como una tasa de interés, que es el precio que tiene el dinero a través del tiempo en el mercado financiero, las TCR tienen periodicidades. Cuando

---

<sup>3</sup>Y que si se ajusta de acuerdo a la inflación, se obtiene la valorización real.

### 3.3 Conceptos principales

el cálculo de la tasa de capitalización se estima a partir del flujo de interés de sólo un valor mensual de arrendamiento, se obtiene una tasa nominal mensual. Cuando se estima a partir de los flujos de renta en un año, corresponde a una tasa anual.

Jaramillo Montoya S. [2000] indica que las tasas de rentabilidad de la inversión en bienes inmuebles se relacionan con el precio del inmueble, su ubicación, edad y estado legal y con la demanda del activo y la oferta de inmuebles similares, y con los criterios liquidez, los flujos de caja, el riesgo y los impuestos. Alzate [2014] cita que, contrario a la rentabilidad de una acción, que es la misma para todo un país, la de los inmuebles depende del sector donde se ubica la vivienda, y puede mostrar grandes variaciones entre sectores de una misma ciudad.

En Colombia, la ley 820 de 2003 estipula los límites legales para la fijación de la renta de arrendamiento. Textualmente dice: “El precio mensual del arrendamiento será fijado por las partes en moneda legal pero no podrá exceder el uno por ciento (1 %) del valor comercial del inmueble o de la parte de él que se dé en arriendo”.

Al respecto Borrero [2008] afirma que: *“es razonable afirmar que la renta de los inmuebles tiende al 1 % mensual en condiciones de baja y media inflación. La tasa de renta sale por diferencia entre la rentabilidad media nacional y la valorización de los inmuebles más un premio por la iliquidez del mercado inmobiliario. Naturalmente que esta tasa oscilará de acuerdo con el momento del ciclo económico, el mercado inmobiliario, el estrato y el uso del inmueble en la zona”*

Por su parte, la Resolución 620 de 2008, “Por la cual se establecen los procedimientos para los avalúos ordenados dentro del marco de la Ley 388 de 1997”, establece como parte de las técnicas valuatorias el método de capitalización de rentas<sup>4</sup> los aspectos para tener en cuenta en la selección de la información y en la aplicación del método:

1. Que dichos contratos tengan menos de un (1) año de suscritos.
2. Que el canon de arrendamiento no sobrepase los topes legales.
3. Que los montos relacionados con el pago de servicios públicos y las cuotas de administración, no se incluyan en el cálculo correspondiente para la aplicación del método.
4. Los arrendamientos a comparar deben referirse a inmuebles que tengan rentas de acuerdo con la norma de uso del terreno o de las construcciones.
5. Las rentas a tener en cuenta para el cálculo del valor comercial de la propiedad deben estar asociadas exclusivamente al inmueble y no a la rentabilidad de la actividad económica que en él se realiza.
6. La tasa de capitalización (i) utilizada en este método debe proceder de la relación calculada entre el canon de renta y el valor comercial de las propiedades similares

---

<sup>4</sup>“ARTÍCULO 2o. MÉTODO DE CAPITALIZACIÓN DE RENTAS O INGRESOS. Es la técnica valuatoria que busca establecer el valor comercial de un bien, a partir de las rentas o ingresos que se puedan obtener del mismo bien, o inmuebles semejantes y comparables por sus características físicas, de uso y ubicación, trayendo a valor presente la suma de los probables ingresos o rentas generadas en la vida remanente del bien objeto de avalúo, con una tasa de capitalización o interés”.

### 3.4 Antecedentes

al inmueble objeto de avalúo, en función del uso o usos existentes en el predio y de localización comparable”.

En este documento se consideran principalmente las ofertas de tipo residencial. Éstas se dividen en ofertas residenciales PH, las cuales corresponden a los apartamentos y por otro lado las NPH, que representan las casas.

### 3.4. Antecedentes

Son pocas las metodologías documentadas de estimación de las TCR. El estudio de (Medrano, Perlaza, & Fuentes, 2011) propone como metodología para estimar la tasa de capitalización de la renta en vivienda en propiedad horizontal estrato 3 en PH en Bogotá un modelo de regresión lineal. Posteriormente, la publicación de S.A. [2017] que tiene como objetivo la medición de inmuebles residenciales urbanos estrato 2, aunque no presenta de manera explícita la metodología, expone que, como resultado del ejercicio, las variables área de terreno, estado de conservación resultaron significativas y con mayor incidencia en la tasa de capitalización frente a las demás variables utilizadas, a saber, tipo de inmueble, área privada, ascensor, equipamiento, estado de conservación y edad. De manera general se concluye que, en este estrato, se observa una tasa mayor para apartamentos que para casas.

Por otra parte, la UAECB y Camacol Bogotá Cundinamarca, en el año 2018, realizaron un primer ejercicio conjunto para el cálculo de la Tasa de Renta Inmobiliaria, calculadas para las desagregaciones geográficas con información de ofertas disponibles en el periodo 2016-2018. Como parte de las recomendaciones y resultados de dicho estudio, se encontró que era necesario mejorar la georreferenciación de las ofertas y aumentar la muestra, toda vez que en ciertas áreas no se logró obtener la correspondiente TCR.

### 3.5. Controles de calidad sobre la base

La base de datos consta originalmente con 1.390.326 registros, incluyendo todos los tipos de predios. Un primer filtro realizado corresponde a dejar aquellas ofertas de predios residenciales, es decir, cuyo tipo de inmueble es apartamento o casa. El total de registros residenciales es 1.072.252. Por otro lado, se tuvo en cuenta diferentes exclusiones para eliminar inconsistencias, dentro de las cuales se incluyen los casos que no tienen estrato, sin área construida, sin identificador del lote (barmanpre) y ofertas rurales. Es importante tener en cuenta que las exclusiones no son excluyentes, lo que quiere decir que un predio puede tener estar en dos o más exclusiones. Luego de estas exclusiones, el número de predios en la base para trabajar es 704.408. En el cuadro 1 se muestran el número de predios excluidos en cada caso.

### 3.5 Controles de calidad sobre la base

Cuadro 1: Número de predios en cada una de las exclusiones.

Exclusión	N
Rurales	97.321
Sin área	254.640
Sin Barmanpre	81.918
Sin estrato	19.043

Debido a que la base de datos tiene información de diferentes fuentes y estas contienen información reportada directamente por los usuarios de plataformas digitales, puede darse el caso en que se tengan valores mal digitados o ciertas inconsistencias. Para evitar estos casos se realizaron controles sobre esta información, con el objetivo de que la información utilizada por el modelo lineal sea lo más depurada posible.

La información utilizada dentro de los diferentes modelos consiste en las áreas de construcción, precios reportados y los valores integrales. Considerando que se cuenta con datos, por un lado de ventas y arriendo, mientras que por otra parte predios en propiedad horizontal (PH) y no propiedad horizontal (NPH), además de diferentes estratos, que conformaron diferentes agrupaciones de manera inicial. Las agrupaciones dependen del año de recolección del dato, clase de predio (PH y NPH), estrato y tipo de oferta (arriendo y venta). Es decir que, por ejemplo, las ofertas de arriendo en de apartamentos no se van a comparar con ofertas en arriendo en NPH, ni entre vigencias o entre estratos, ya que de antemano se espera que tengan un comportamiento diferente. Una vez se conforman estas agrupaciones, lo ideal es realizar comparaciones entre ofertas que sean próximas desde una perspectiva geográfica. Estas comparaciones se realizan con el objetivo de determinar si la oferta es similar en sus características a ofertas, debido a que en otro caso la información recolectada en ese registro presentaría anomalías y debería reportarse como atípica, al mismo tiempo que ser excluida de los análisis a desarrollar.

Una manera de determinar registros próximos geográficamente, es seleccionando todos los que se encuentren en el mismo sector catastral, pero un primer inconveniente surge cuando en un sector solamente se encuentre una oferta o un número insuficiente de ofertas para realizar la comparación. Por este motivo, se observa la necesidad de realizar aglomeraciones más grandes de tal forma que se tengan grupos con un número suficiente de información para realizar la validación. Es importante mencionar que si el número de registros supera el umbral determinado, no se requiere una agrupación más grande y se aplicó el protocolo de comparación únicamente con las ofertas que se encuentran en ese barrio.

Con el objetivo de tener agrupaciones de ofertas lo suficientemente grandes y realizar la respectiva comparación, pero de tal forma que no se agregue demasiado, considerando la singularidad de cada sector, se propone agruparon sectores de acuerdo a su codificación. Se esperaría que los sectores cuya codificación es igual a diferencia del último dígito son más cercanos que dos sectores que tienen diferencias en sus dos últimos dígitos. Por ejemplo, se esperaría que los sectores “001101” y “001102” sean próximos geográficamente hablando. De



### 3.5 Controles de calidad sobre la base

esta manera, el algoritmo propuesto busca estas coincidencias a partir de la distancia de Levenshtein [Sariyar and Borg, 2020], la cual es un valor entre 0 y 1. Esta distancia toma un valor de 0 cuando las dos codificaciones no tienen dígitos en común en las mismas posiciones y toma el valor de 1 cuando son exactamente iguales. El algoritmo realiza la búsqueda iterativa de sectores similares, iniciando por los últimos dígitos, de forma que si no encuentra sectores con el mismo código y último dígito diferente, procede a buscar coincidencias exceptuando los dos últimos dígitos y así sucesivamente hasta superar el umbral definido. Una vez el algoritmo encuentra agrupaciones tal que todas superan dicho valor predefinido se detiene y no continua con el proceso de agregación. El número de predios determinado, como umbral es 10, lo cual quiere decir que si un sector catastral cuenta con menos de 10 registros a nivel de tipo de oferta, estrato, clase de predio y año, se buscarán opciones de colapso, de tal forma que se encuentren ofertas próximas para poder realizar la comparación.

Una vez se conformaron las diferentes agrupaciones se procedió a utilizar una carta de control multivariada [Montgomery, 2020], lo que se debe a que se están controlando diferentes características, es decir un proceso multivariado, al involucrar el área construida, precio y el valor integral de las diferentes ofertas. Al considerar una carta de control multivariada se hace uso de la relación implícita dentro de las variables, lo cual no sucede con métodos univariados. Los métodos de control de calidad usualmente consideran dos etapas, lo que no se tiene en cuenta dentro de este ejercicio ya que no se tuvieron en cuenta ofertas adicionales. El control de calidad es solo un paso adicional para lograr el objetivo del estudio, el cual radica en la estimación de la tasa de capitalización. La figura 1 muestra un ejemplo de una carta de control multivariada, donde las ofertas se ilustran a partir de puntos, las bandas determinan límites de tal forma que una oferta que está fuera de ellos será marcada como atípica y se muestran en el gráfico en color rojo.

Luego de este control de calidad, la base de datos quedó conformada por un total de 580.980 registros, con lo cual se procedió a realizar los diferentes cálculos de promedios por sector, de tal manera que en cada uno de ellos se tendrá un promedio de venta y de arriendo, por estrato, vigencia y clase de predio (apartamento o casa). Al realizar este cálculo se tuvo un inconveniente, ya que un sector puede tener ofertas de arriendo, más no de ventas o viceversa. Para remediar esto, en el caso en que se tenga información de arriendo y no de ventas, se toma el promedio de valor integral de la base catastral, de acuerdo a la misma desagregación. Mientras que en el recíproco, cuando se tiene el valor de venta, pero no en el de arriendo se utiliza el promedio de la localidad respectivo para imputar este valor y proceder con la construcción del modelo estadístico.

Considerando que se elaboró una serie de procedimientos, controles de calidad, depuraciones e imputaciones se realizó un análisis para excluir registros que tuvieran tasas de renta atípicas, que puedan sesgar los resultados del estudio. Para esta situación, se excluyeron aquellos casos con una tasa de capitalización superior al 1 % o inferiores al 0.33 %. Luego de este proceso, la tabla 2 muestra el número de sectores resultantes, la cantidad de registros de ofertas y de ventas utilizadas en cada caso, respectivamente, donde se observa que la cantidad de ofertas de venta es mayor a la de arriendo y se encuentran cantidades mayores en estratos del 2 al 4.

### 3.5 Controles de calidad sobre la base

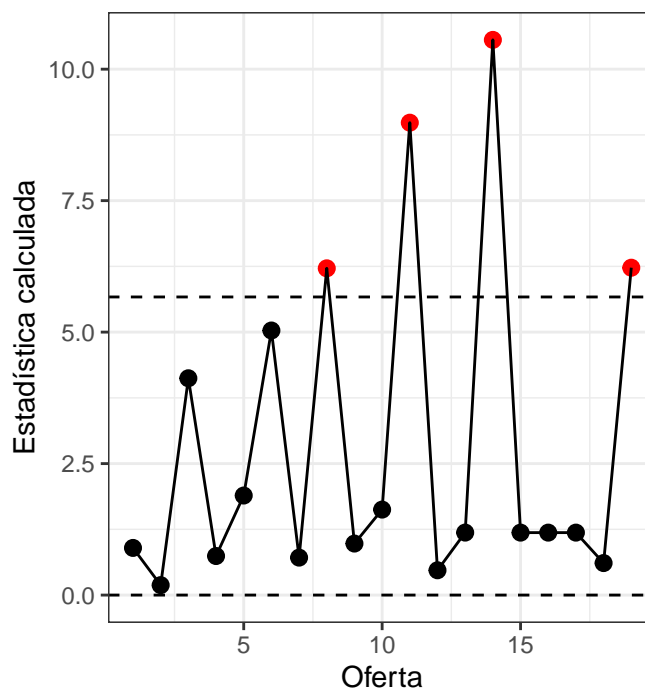


Figura 1: Ilustración de la carta de control multivariada.

Cuadro 2: Cantidad de predios y sectores según clase de predio, estrato y año, luego de la depuración de la base.

CLASE_PREDIO	AÑO	ESTRATO	N de sec- tores	N de ofertas de arriendo	N de ofertas de venta
N	2017	1	62	28	565
N	2017	2	188	227	3135
N	2017	3	224	748	5742
N	2017	4	70	580	1791
N	2017	5	30	227	737
N	2017	6	12	48	144
N	2018	1	36	29	310
N	2018	2	185	224	2962
N	2018	3	218	897	5526
N	2018	4	70	716	2334
N	2018	5	32	344	1097
N	2018	6	20	98	288
N	2019	1	51	29	540
N	2019	2	221	290	4453
N	2019	3	253	1235	8053

### 3.5 Controles de calidad sobre la base

Cuadro 2: Cantidad de predios y sectores según clase de predio, estrato y año, luego de la depuración de la base.  
*(continued)*

CLASE_PREDIO	AÑO	ESTRATO	N de sec- tores	N de ofertas de arriendo	N de ofertas de venta
N	2019	4	77	1249	3019
N	2019	5	31	505	1386
N	2019	6	19	234	500
N	2020	1	53	27	559
N	2020	2	213	354	4052
N	2020	3	214	1117	7252
N	2020	4	77	1284	3030
N	2020	5	29	565	1727
N	2020	6	17	261	635
P	2017	1	34	106	129
P	2017	2	258	3367	5761
P	2017	3	349	12915	17660
P	2017	4	134	12588	17850
P	2017	5	58	5681	9906
P	2017	6	36	7637	11267
P	2018	1	37	99	137
P	2018	2	283	3472	5877
P	2018	3	358	14677	18683
P	2018	4	140	14516	20623
P	2018	5	59	7225	11711
P	2018	6	38	9132	13263
P	2019	1	31	95	93
P	2019	2	272	3079	5398
P	2019	3	365	12872	15731
P	2019	4	153	14781	20792
P	2019	5	65	7462	12099
P	2019	6	39	9652	13653
P	2020	1	36	103	91
P	2020	2	281	3022	4625
P	2020	3	365	14732	15171
P	2020	4	153	19046	20629
P	2020	5	63	9054	13351
P	2020	6	39	12873	16992
Total	Total	Total	6048	209502	331329

## 4. Modelo aplicado

El objetivo de esta sección es ilustrar el modelo estadístico aplicado junto con una breve justificación de su uso y respectivos resultados. Para eliminar el efecto que pueda tener el área construida sobre los valores de venta y de arriendo totales se tomó el valor integral, que se calcula dividiendo los valores totales sobre el área construida, es decir, lo que se muestra en la ecuación (2). En este modelo la variable respuesta es el valor integral de venta, mientras que una de las variables independientes es el valor integral de arriendo.

$$Valor\_integral = \frac{Valor\_total}{Area\_construida} \quad (2)$$

Considerando que no se cuenta con información de ventas y arriendo de manera simultánea en un solo predio, se toma la propuesta dada en Deaton [1985], cuyas cohortes van a estar representadas por los diferentes sectores catastrales. De esta manera, el modelo estimado se presenta en la ecuación (3), donde  $P_{ij}$  y  $R_{ij}$  representan los valores integrales de venta y arriendo promedio, respectivamente, en el  $i$ -ésimo sector catastral en el  $j$ -ésimo tiempo,  $\alpha$  es el efecto global del valor integral de arriendo sobre el valor integral de venta. Por otro lado,  $\beta_k$  es el efecto del  $k$ -ésimo estrato sobre el efecto general de la relación entre valores integrales, denotado como  $\alpha$ ,  $\eta_l$  es el efecto de la  $l$ -ésima localidad sobre el efecto general de la relación entre integrales,  $\delta_m$  es el efecto del  $m$ -ésimo sector catastral sobre el efecto global y  $\kappa_p$  es el efecto del  $p$ -ésimo año, de manera análoga a los otros coeficientes. Por último  $\epsilon_{ij}$  representa a los errores de la regresión.

$$P_{ij} = \alpha R_{ij} + \beta_k R_{ij} + \eta_l R_{ij} + \kappa_p R_{ij} + \delta_m R_{ij} + \epsilon_{ij} \quad (3)$$

En el caso del modelo (3), los factores relacionados al estrato y localidad son efectos fijos, ya que en la base de ofertas utilizada se tiene información de todos los estratos y de todas las localidades, mientras que los efectos relacionados con los sectores catastrales son efectos aleatorios, debido a que no se tienen mediciones de todos los sectores en la ciudad de Bogotá, pero la inferencia a realizar se puede generalizar a la población de sectores. Como en el modelo (3) se incluyen efectos fijos y aleatorios, el modelo se denomina de efectos mixtos.

En este modelo se supone lo siguiente

$$\sum_k \beta_k = \sum_l \eta_l = 0,$$

lo cual se debe a temas relacionados con la estimabilidad de los parámetros. Por otro lado, en cuanto a los factores aleatorios, se supone que los coeficientes y errores son independientes e idénticamente distribuidos y a su vez  $\delta_m \sim N(0, \sigma_\delta^2)$  y  $\epsilon_m \sim N(0, \sigma_\epsilon^2)$ . Es necesario tener en cuenta que el modelo no incluye intercepto, considerando los objetivos del estudio. El valor integral de arriendo, denotado por  $R_{ij}$  se incluye en todas las covariables, ya que el objetivo es determinar la relación entre los valores totales y de arriendo. Para esclarecer el proceso de estimación de la tasa de rentabilidad, el lector puede notar que la ecuación (3) se puede reexpresar como sigue

## 4.1 Apartamentos

$$P_{ij} = [\alpha + \beta_k + \eta_l + \kappa_p + \delta_m] R_{ij} + \epsilon_{ij},$$

donde se observa que la tasa de renta, se va a calcular a partir de la suma de los respectivos coeficientes, según sea el caso.

Una vez se realizó la estimación de los coeficientes del modelo, se realizó un análisis para determinar puntos de apalancamiento o de influencia, con el fin de encontrar aquellos casos donde un solo registro pueda influir y cambiar los coeficientes del modelo de manera importante. Para ello se realizó el cálculo de la distancia de *Cook* y de los *DFBETAS*, que son medidas de diagnóstico del modelo. Detalles adicionales relacionados con modelos lineales de efectos mixtos o estas medidas de diagnóstico se pueden consultar en el anexo de este documento, en [Montgomery et al. \[2012\]](#), [Hinkelmann \[2011\]](#) o [Melo et al. \[2007\]](#).

A partir de este modelo se estima un coeficiente, mayor a 1, que nos permite conocer la razón entre el valor integral de venta y el valor integral o dicho de otra manera un valor que al multiplicarlo por el valor de arriendo, nos permita estimar el valor total de venta, pero para temas interpretativos con otras áreas de la Unidad e inclusive con la literatura relacionada, en los resultados presentados se va a mostrar el inverso de ese coeficiente, que va a ser un valor entre 0 y 1.

### 4.1. Apartamentos

A continuación se presentan los resultados de las diferentes tasas de rentabilidad para el caso de ofertas en propiedad horizontal. La figura 2 muestra diferentes diagramas de dispersión correspondientes a valores promedio de arriendo y de venta para los estratos en cada recuadro y diferentes años, representados por colores, para el caso de apartamentos. Allí se observa una relación lineal, lo cual justifica el uso del modelo (3) para este grupo de ofertas, al igual que una pendiente más empinada a medida que el estrato aumenta, lo que se refleja en una tasa de renta más baja para estratos altos y un comportamiento similar año a tras año.

#### 4.1.1. Resultados

La tabla 3 y la tabla 4 muestran los resultados correspondientes a las tasas de renta a nivel de estrato y año, y a nivel de localidad y año. En el primer cuadro se observa que a medida que el estrato aumenta, la tasa de rentabilidad es menor. Las tasas de renta de los estratos 1 y 2 oscilan al rededor de 0.55 %. En el estrato 3 se tienen valores cercanos al 0.51 %, mientras que en el 4 y 5 se tienen valores cercanos al 0.49 % y por último en el 6 se tiene una tasa de renta cercana al 0.46 % a través de los diferentes años considerados. En el cuadro 4 se observa que las localidades con mayores tasas de renta son Chapinero, La Candelaria y San Cristobal. Por otro lado, la localidad con menor tasa de rentabilidad es Puente Aranda. Las figuras 3 y 4 muestran los mismos resultados de una manera gráfica.

#### 4.1 Apartamentos

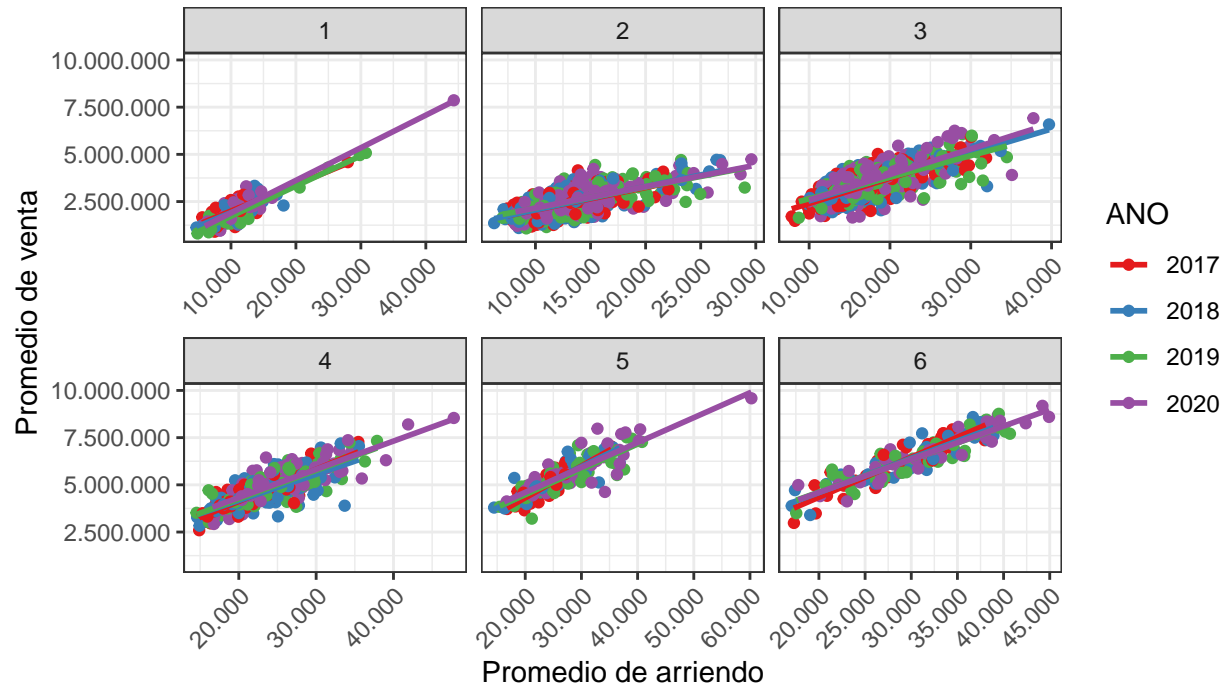


Figura 2: Diagrama de dispersión de los valores integrales de venta y de arriendo con sus respectivas rectas de regresión ajustadas por estrato y vigencia para apartamentos.

Cuadro 3: Tasas de capitalización para apartamentos según estratos y años.

AÑO	ESTRATO	TASA_RENTA
2017	1	0.00539
2018	1	0.00539
2019	1	0.00545
2020	1	0.00540
2017	2	0.00547
2018	2	0.00547
2019	2	0.00553
2020	2	0.00548
2017	3	0.00514
2018	3	0.00514
2019	3	0.00519
2020	3	0.00515
2017	4	0.00489
2018	4	0.00489
2019	4	0.00494

#### 4.1 Apartamentos

Cuadro 3: Tasas de capitalización para apartamentos según estratos y años. (*continued*)

AÑO	ESTRATO	TASA_RENTA
2020	4	0.00490
2017	5	0.00485
2018	5	0.00484
2019	5	0.00489
2020	5	0.00485
2017	6	0.00459
2018	6	0.00459
2019	6	0.00463
2020	6	0.00460

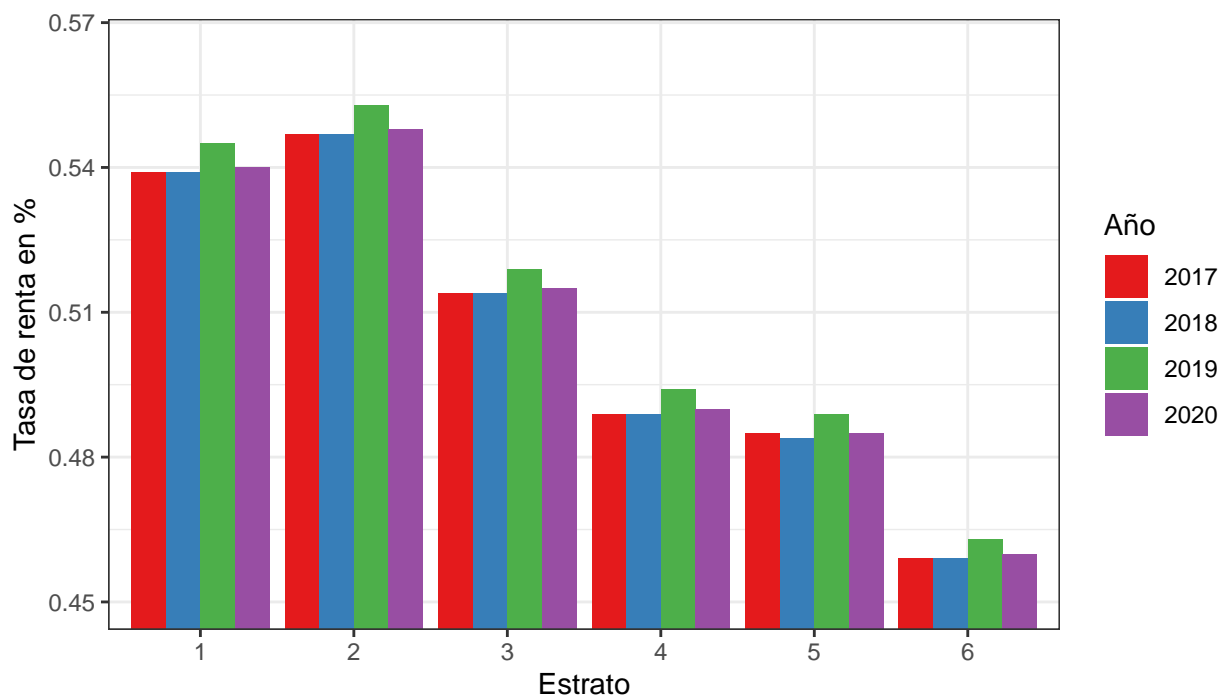


Figura 3: Tasas de capitalización para apartamentos según estratos y años.

Cuadro 4: Tasas de capitalización para apartamentos según localidades y años.

AÑO	CODIGO_LOCALIDAD	NOMBRE_LOCALIDAD	TASA_RENTA
2017	1	USAQUEN	0.00539

#### 4.1 Apartamentos

Cuadro 4: Tasas de capitalización para apartamentos  
según localidades y años. *(continued)*

AÑO	CODIGO_LOCALIDAD	NOMBRE_LOCALIDAD	TASA_RENTA
2018	1	USAQUEN	0.00539
2019	1	USAQUEN	0.00545
2020	1	USAQUEN	0.00540
2017	2	CHAPINERO	0.00574
2018	2	CHAPINERO	0.00574
2019	2	CHAPINERO	0.00581
2020	2	CHAPINERO	0.00575
2017	3	SANTA FE	0.00580
2018	3	SANTA FE	0.00580
2019	3	SANTA FE	0.00587
2020	3	SANTA FE	0.00581
2017	4	SAN CRISTOBAL	0.00552
2018	4	SAN CRISTOBAL	0.00552
2019	4	SAN CRISTOBAL	0.00558
2020	4	SAN CRISTOBAL	0.00553
2017	5	USME	0.00563
2018	5	USME	0.00562
2019	5	USME	0.00569
2020	5	USME	0.00563
2017	6	TUNJUELITO	0.00531
2018	6	TUNJUELITO	0.00531
2019	6	TUNJUELITO	0.00537
2020	6	TUNJUELITO	0.00532
2017	7	BOSA	0.00509
2018	7	BOSA	0.00509
2019	7	BOSA	0.00514
2020	7	BOSA	0.00510
2017	8	KENNEDY	0.00505
2018	8	KENNEDY	0.00505
2019	8	KENNEDY	0.00510
2020	8	KENNEDY	0.00506
2017	9	FONTIBON	0.00502
2018	9	FONTIBON	0.00502
2019	9	FONTIBON	0.00507
2020	9	FONTIBON	0.00503
2017	10	ENGATIVA	0.00507
2018	10	ENGATIVA	0.00506



#### 4.1 Apartamentos

Cuadro 4: Tasas de capitalización para apartamentos  
según localidades y años. *(continued)*

AÑO	CODIGO_LOCALIDAD	NOMBRE_LOCALIDAD	TASA_RENTA
2019	10	ENGATIVA	0.00512
2020	10	ENGATIVA	0.00507
2017	11	SUBA	0.00531
2018	11	SUBA	0.00530
2019	11	SUBA	0.00537
2020	11	SUBA	0.00532
2017	12	BARRIOS UNIDOS	0.00547
2018	12	BARRIOS UNIDOS	0.00546
2019	12	BARRIOS UNIDOS	0.00553
2020	12	BARRIOS UNIDOS	0.00547
2017	13	TEUSAQUILLO	0.00555
2018	13	TEUSAQUILLO	0.00554
2019	13	TEUSAQUILLO	0.00561
2020	13	TEUSAQUILLO	0.00556
2017	14	LOS MARTIRES	0.00545
2018	14	LOS MARTIRES	0.00544
2019	14	LOS MARTIRES	0.00551
2020	14	LOS MARTIRES	0.00545
2017	15	ANTONIO NARINO	0.00488
2018	15	ANTONIO NARINO	0.00487
2019	15	ANTONIO NARINO	0.00492
2020	15	ANTONIO NARINO	0.00488
2017	16	PUENTE ARANDA	0.00456
2018	16	PUENTE ARANDA	0.00456
2019	16	PUENTE ARANDA	0.00460
2020	16	PUENTE ARANDA	0.00456
2017	17	LA CANDELARIA	0.00592
2018	17	LA CANDELARIA	0.00591
2019	17	LA CANDELARIA	0.00599
2020	17	LA CANDELARIA	0.00592
2017	18	RAFAEL URIBE URIBE	0.00507
2018	18	RAFAEL URIBE URIBE	0.00506
2019	18	RAFAEL URIBE URIBE	0.00512
2020	18	RAFAEL URIBE URIBE	0.00507
2017	19	CIUDAD BOLIVAR	0.00544
2018	19	CIUDAD BOLIVAR	0.00543
2019	19	CIUDAD BOLIVAR	0.00549

## 4.2 Casas

Cuadro 4: Tasas de capitalización para apartamentos según localidades y años. (*continued*)

AÑO	CODIGO_LOCALIDAD	NOMBRE_LOCALIDAD	TASA_RENTA
2020	19	CIUDAD BOLIVAR	0.00544

### 4.1.2. Validación de supuestos

A manera de verificación de los supuestos del modelo para el grupo de ofertas de apartamentos, se realizó un análisis sobre los residuales del modelo. Las figuras 5 y 6 son gráficos de dispersión y de cajas y bigotes, que ilustran la relación entre los residuales con los valores ajustados, con los estratos y años, respectivamente. El objetivo de estas dos figuras es buscar patrones en los gráficos. Como se puede observar, no se encuentra una relación en los gráficos presentados.

Por otro lado, se validó el supuesto de normalidad incluido en los efectos aleatorios del modelo dado en la ecuación (3). Para ello, las figuras 7 y 8 muestran las densidades estimadas para los efectos aleatorios asociados a los sectores catastrales y a los residuales del modelo, respectivamente. En ambos casos se observa una forma acampanada, por lo cual se puede suponer que en ambos casos, los valores provienen de una distribución normal.

## 4.2. Casas

A continuación se presentan los resultados de las diferentes tasas de rentabilidad para el caso de ofertas en no propiedad horizontal. La figura 9 muestra diferentes diagramas de dispersión correspondientes a valores promedio de arriendo y de venta para los estratos en cada recuadro y diferentes años, representados por colores, para el caso de casas. Allí se observa una relación lineal, lo cual justifica el uso del modelo (3) para este grupo de ofertas, al igual que una pendiente más empinada a medida que el estrato aumenta y un comportamiento similar año a través de los años.

### 4.2.1. Resultados

El procedimiento en esta sección es análogo a lo presentado en el caso de propiedad horizontal. Las tablas 5 y 6 muestran los resultados para casas a nivel de estratos y localidades en las diferentes vigencias. Allí se observa, al igual que en el caso de la propiedad horizontal, que a medida que el estrato aumenta la tasa de rentabilidad es más baja. Las localidades con una mayor tasa de rentabilidad en casas son Chapinero, Los Mártires y Santa Fé, mientras que las de menor tasa son Bosa y Puente Aranda. Las figuras 10 y 11 muestran los mismos resultados de una manera gráfica. Para el caso de casas, a diferencia del de apartamentos, en el 2020 se observa una diferencia más notoria con respecto a los otros años, ya que en dicho

## 4.2 Casas

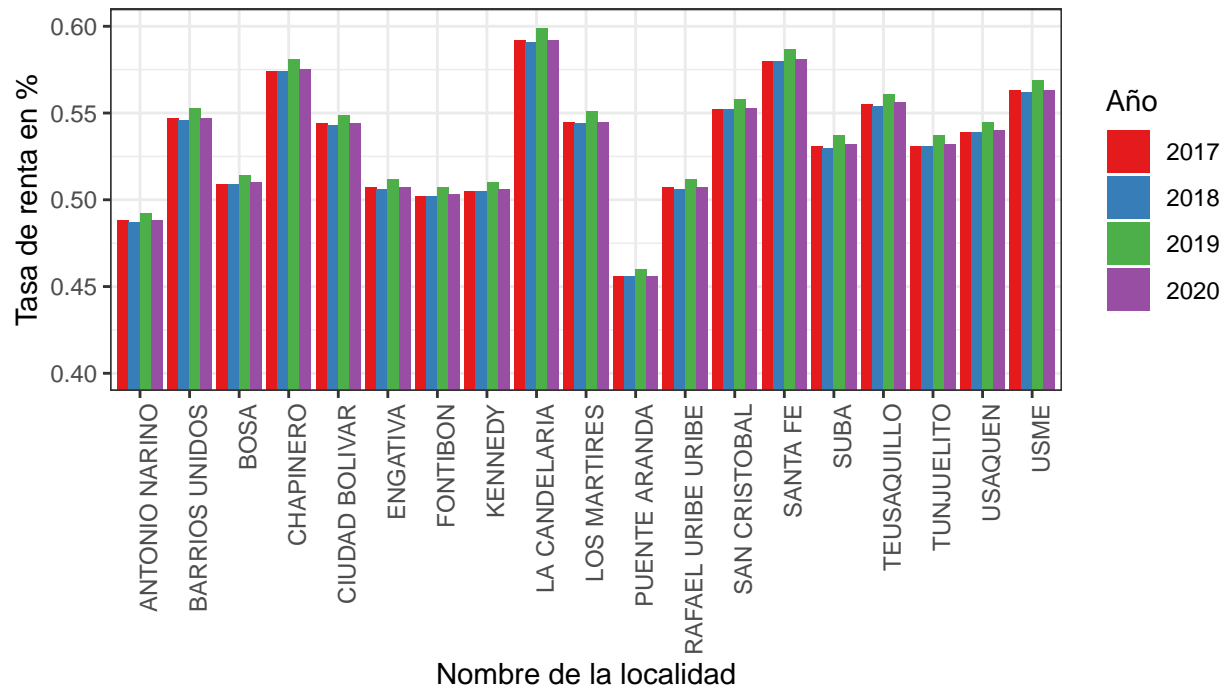


Figura 4: Tasas de capitalización para apartamentos según localidades y años.

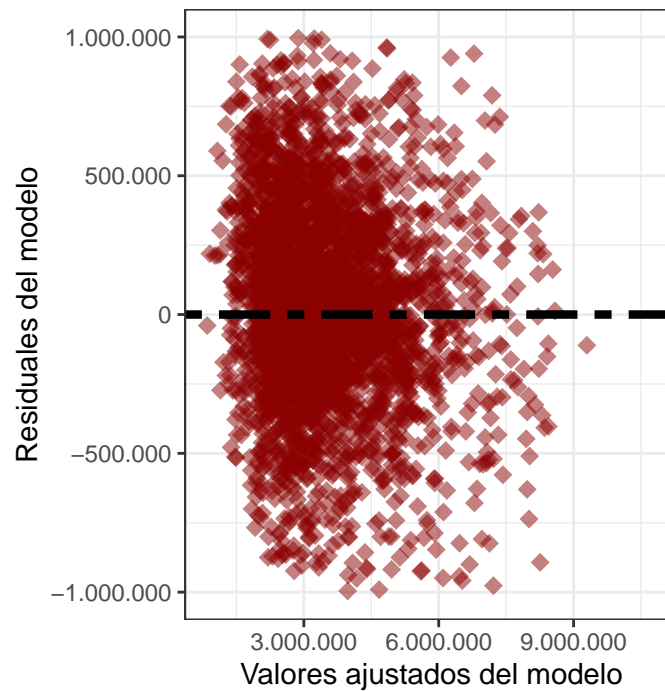


Figura 5: Diagrama de dispersión de los residuales de la regresión versus los valores ajustados en el modelo de apartamentos.

## 4.2 Casas

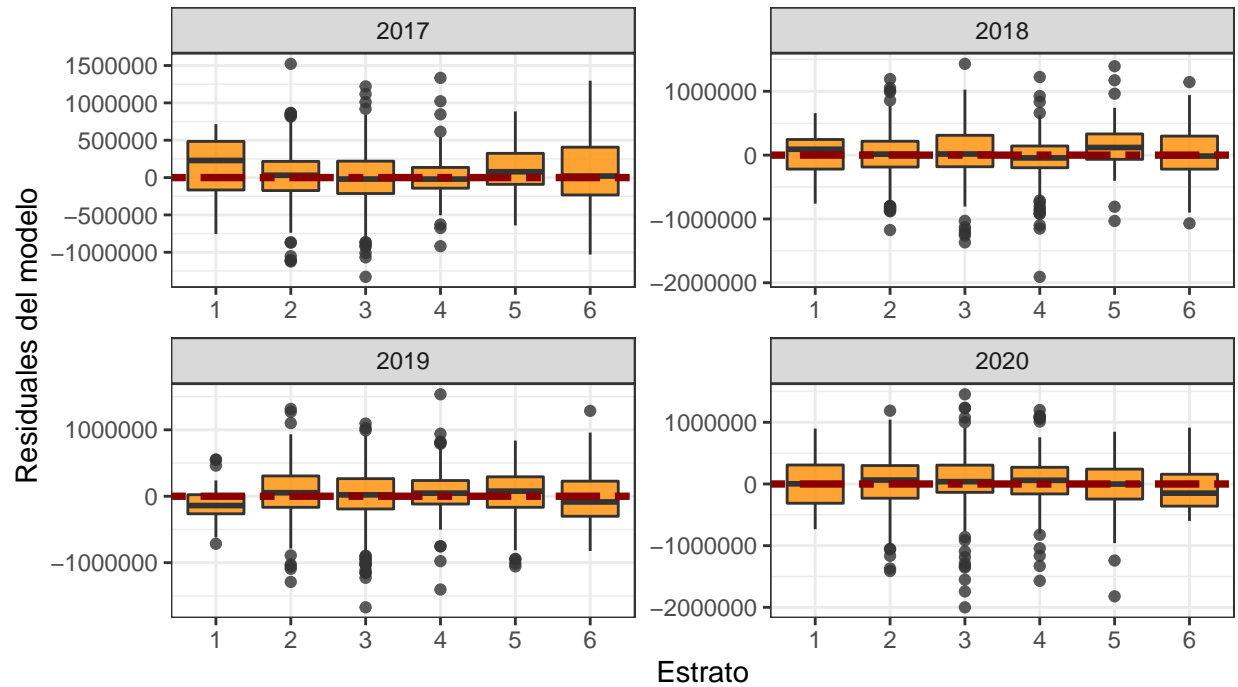


Figura 6: Diagramas de cajas y bigotes de los residuales de la regresión versus los estratos para diferentes vigencias en el modelo de apartamentos.

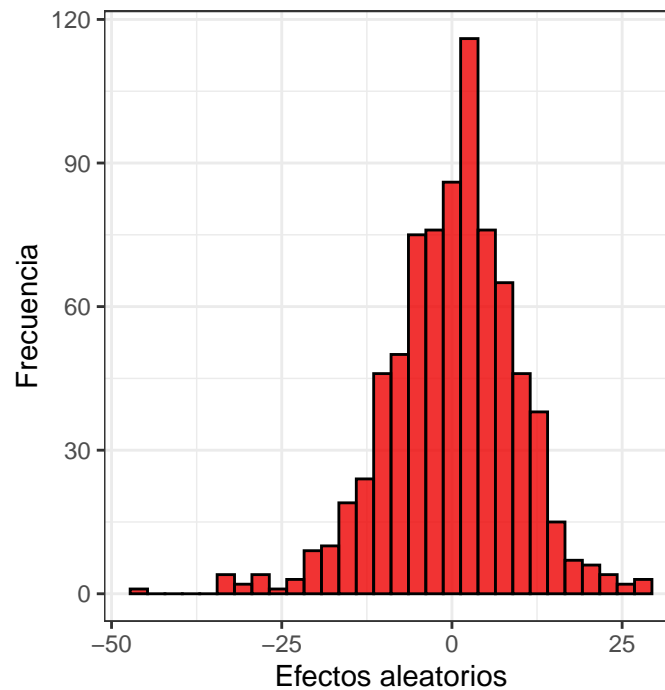


Figura 7: Densidad estimada de los efectos aleatorios asociados a los sectores en el modelo de apartamentos.

## 4.2 Casas

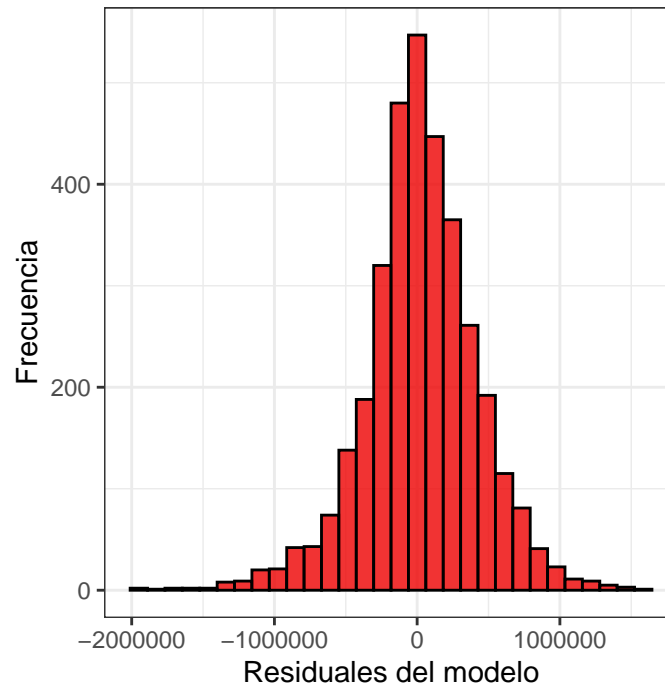


Figura 8: Densidad estimada de los residuales obtenidos en la regresión para el caso de apartamentos.

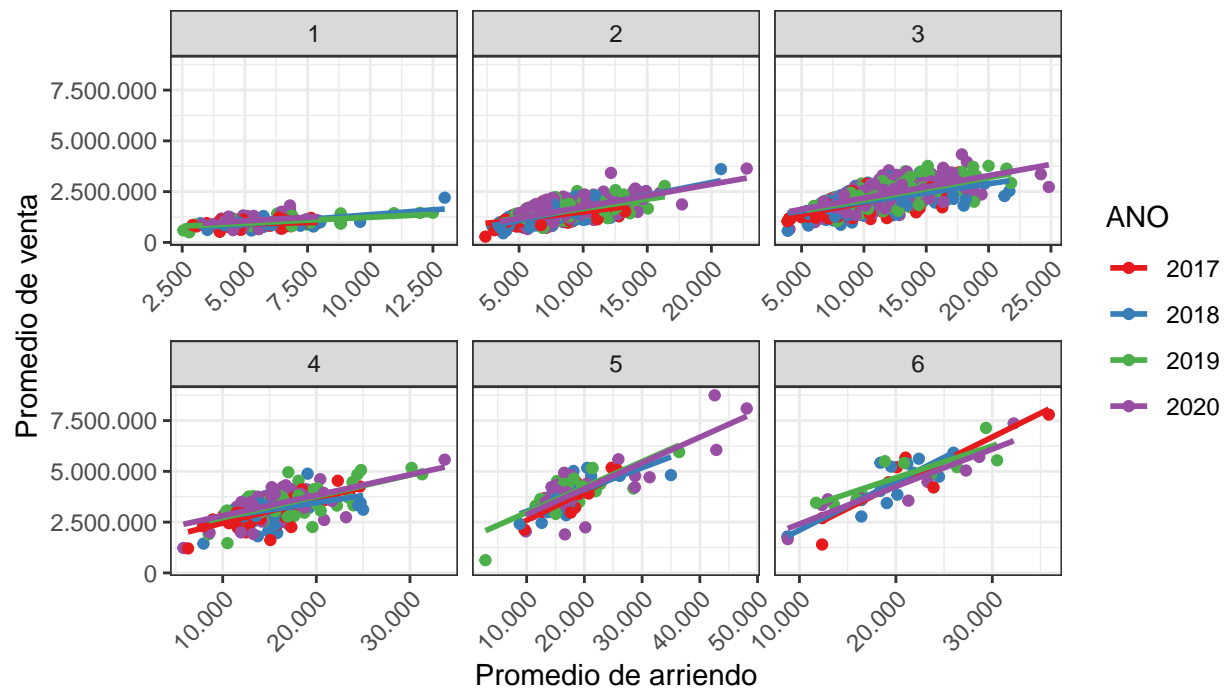


Figura 9: Diagrama de dispersión de los valores integrales de venta y de arriendo con sus respectivas rectas de regresión ajustadas por estrato y vigencia para apartamentos.

## 4.2 Casas

año se tiene una tasa de renta más baja en comparación con los otros, mientras que en 2019 se tuvo la tasa más alta, seguida de 2017 y de 2018.

Cuadro 5: Tasas de capitalización para casas según estrato y años.

AÑO	ESTRATO	TASA_RENTA
2017	1	0.00590
2018	1	0.00581
2019	1	0.00597
2020	1	0.00569
2017	2	0.00532
2018	2	0.00524
2019	2	0.00538
2020	2	0.00515
2017	3	0.00491
2018	3	0.00484
2019	3	0.00496
2020	3	0.00476
2017	4	0.00448
2018	4	0.00442
2019	4	0.00452
2020	4	0.00435
2017	5	0.00453
2018	5	0.00447
2019	5	0.00457
2020	5	0.00440
2017	6	0.00443
2018	6	0.00438
2019	6	0.00447
2020	6	0.00431

Cuadro 6: Tasas de capitalización para casas según localidades y años.

AÑO	CODIGO_LOCALIDAD	NOMBRE_LOCALIDAD	TASA_RENTA
2017	1	USAQUEN	0.00590
2018	1	USAQUEN	0.00581
2019	1	USAQUEN	0.00597
2020	1	USAQUEN	0.00569
2017	2	CHAPINERO	0.00822

## 4.2 Casas

Cuadro 6: Tasas de capitalización para casas según localidades y años. *(continued)*

AÑO	CODIGO_LOCALIDAD	NOMBRE_LOCALIDAD	TASA_RENTA
2018	2	CHAPINERO	0.00803
2019	2	CHAPINERO	0.00836
2020	2	CHAPINERO	0.00781
2017	3	SANTA FE	0.00843
2018	3	SANTA FE	0.00823
2019	3	SANTA FE	0.00857
2020	3	SANTA FE	0.00799
2017	4	SAN CRISTOBAL	0.00671
2018	4	SAN CRISTOBAL	0.00659
2019	4	SAN CRISTOBAL	0.00680
2020	4	SAN CRISTOBAL	0.00643
2017	5	USME	0.00589
2018	5	USME	0.00579
2019	5	USME	0.00596
2020	5	USME	0.00567
2017	6	TUNJUELITO	0.00622
2018	6	TUNJUELITO	0.00612
2019	6	TUNJUELITO	0.00630
2020	6	TUNJUELITO	0.00598
2017	7	BOSA	0.00554
2018	7	BOSA	0.00545
2019	7	BOSA	0.00560
2020	7	BOSA	0.00535
2017	8	KENNEDY	0.00599
2018	8	KENNEDY	0.00589
2019	8	KENNEDY	0.00606
2020	8	KENNEDY	0.00577
2017	9	FONTIBON	0.00626
2018	9	FONTIBON	0.00615
2019	9	FONTIBON	0.00634
2020	9	FONTIBON	0.00601
2017	10	ENGATIVA	0.00589
2018	10	ENGATIVA	0.00579
2019	10	ENGATIVA	0.00595
2020	10	ENGATIVA	0.00567
2017	11	SUBA	0.00633
2018	11	SUBA	0.00622

## 4.2 Casas

Cuadro 6: Tasas de capitalización para casas según localidades y años. *(continued)*

AÑO	CODIGO_LOCALIDAD	NOMBRE_LOCALIDAD	TASA_RENTA
2019	11	SUBA	0.00641
2020	11	SUBA	0.00608
2017	12	BARRIOS UNIDOS	0.00666
2018	12	BARRIOS UNIDOS	0.00653
2019	12	BARRIOS UNIDOS	0.00675
2020	12	BARRIOS UNIDOS	0.00638
2017	13	TEUSAQUILLO	0.00698
2018	13	TEUSAQUILLO	0.00685
2019	13	TEUSAQUILLO	0.00708
2020	13	TEUSAQUILLO	0.00668
2017	14	LOS MARTIRES	0.00933
2018	14	LOS MARTIRES	0.00908
2019	14	LOS MARTIRES	0.00950
2020	14	LOS MARTIRES	0.00880
2017	15	ANTONIO NARINO	0.00643
2018	15	ANTONIO NARINO	0.00632
2019	15	ANTONIO NARINO	0.00652
2020	15	ANTONIO NARINO	0.00618
2017	16	PUENTE ARANDA	0.00551
2018	16	PUENTE ARANDA	0.00542
2019	16	PUENTE ARANDA	0.00557
2020	16	PUENTE ARANDA	0.00532
2017	17	LA CANDELARIA	0.00777
2018	17	LA CANDELARIA	0.00760
2019	17	LA CANDELARIA	0.00789
2020	17	LA CANDELARIA	0.00740
2017	18	RAFAEL URIBE URIBE	0.00619
2018	18	RAFAEL URIBE URIBE	0.00608
2019	18	RAFAEL URIBE URIBE	0.00627
2020	18	RAFAEL URIBE URIBE	0.00595
2017	19	CIUDAD BOLIVAR	0.00608
2018	19	CIUDAD BOLIVAR	0.00598
2019	19	CIUDAD BOLIVAR	0.00615
2020	19	CIUDAD BOLIVAR	0.00585



## 4.2 Casas

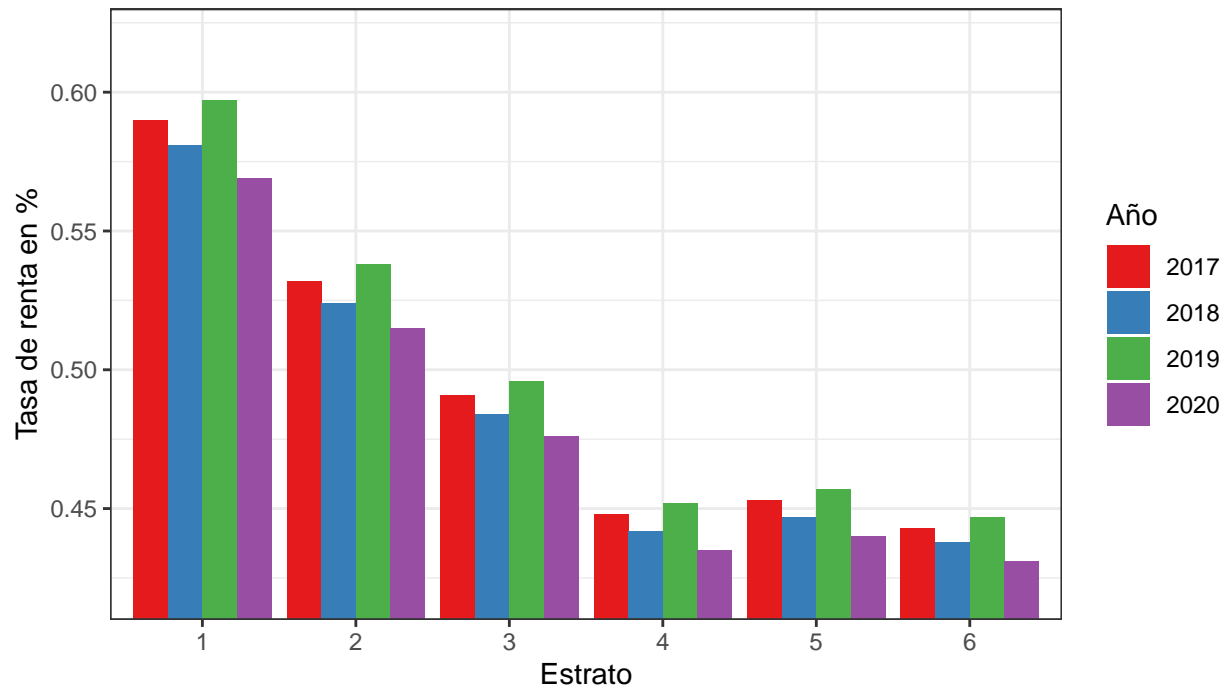


Figura 10: Tasas de capitalización para casas según estratos y años.

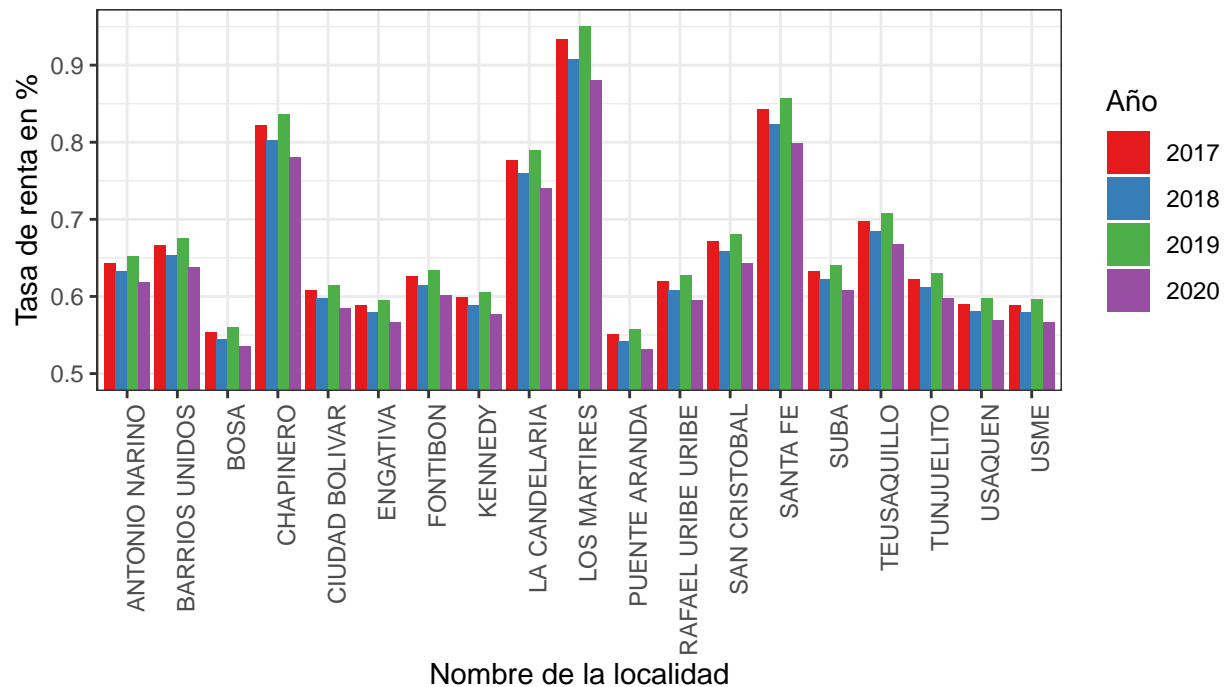


Figura 11: Tasas de capitalización para casas según localidades y años.

#### 4.2.2. Validación de supuestos

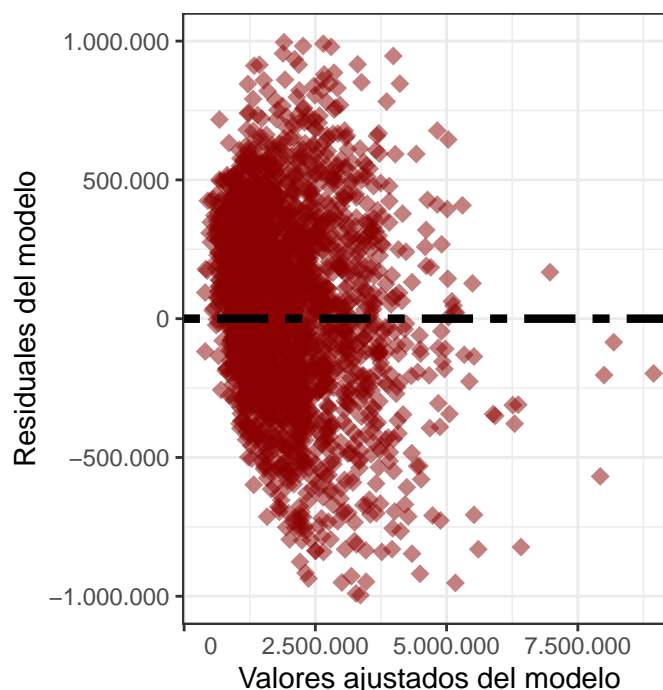


Figura 12: Diagrama de dispersión de los residuales de la regresión versus los valores ajustados en el modelo de casas.

A manera de verificación de los supuestos del modelo para el grupo de ofertas de apartamentos, se realizó un análisis sobre los residuales del modelo. Las figuras 12 y 13 son gráficos de dispersión y de cajas y bigotes, que ilustran la relación entre los residuales con los valores ajustados, con los estratos y años, respectivamente. Por otro lado, se validó el supuesto de normalidad incluido en los efectos aleatorios del modelo dado en la ecuación (3). Para ello, las figuras 14 y 15 muestran las densidades estimadas para los efectos aleatorios asociados a los sectores catastrales y a los residuales del modelo, respectivamente. Las conclusiones de estos gráficos son las mismas a las presentadas para la validación de supuestos en apartamentos.

Es importante mencionar que las tasas de rentabilidad obtenidas para las ofertas en propiedad horizontal son mayores, por lo general a las que corresponden a no propiedad horizontal. Por último, los resultados obtenidos a nivel como sector se calcularon y hacen parte de los resultados de este documento, pero teniendo en cuenta que son tablas con un número de filas considerable, se omiten en este escrito.

## 5. Conclusiones

En este trabajo se realizó la estimación de la tasa de rentabilidad para predios residenciales a diferentes niveles de desagregación, a partir de métodos de regresión tradicionales, utilizando la información de ofertas de mercado que han sido capturadas en los últimos años por la

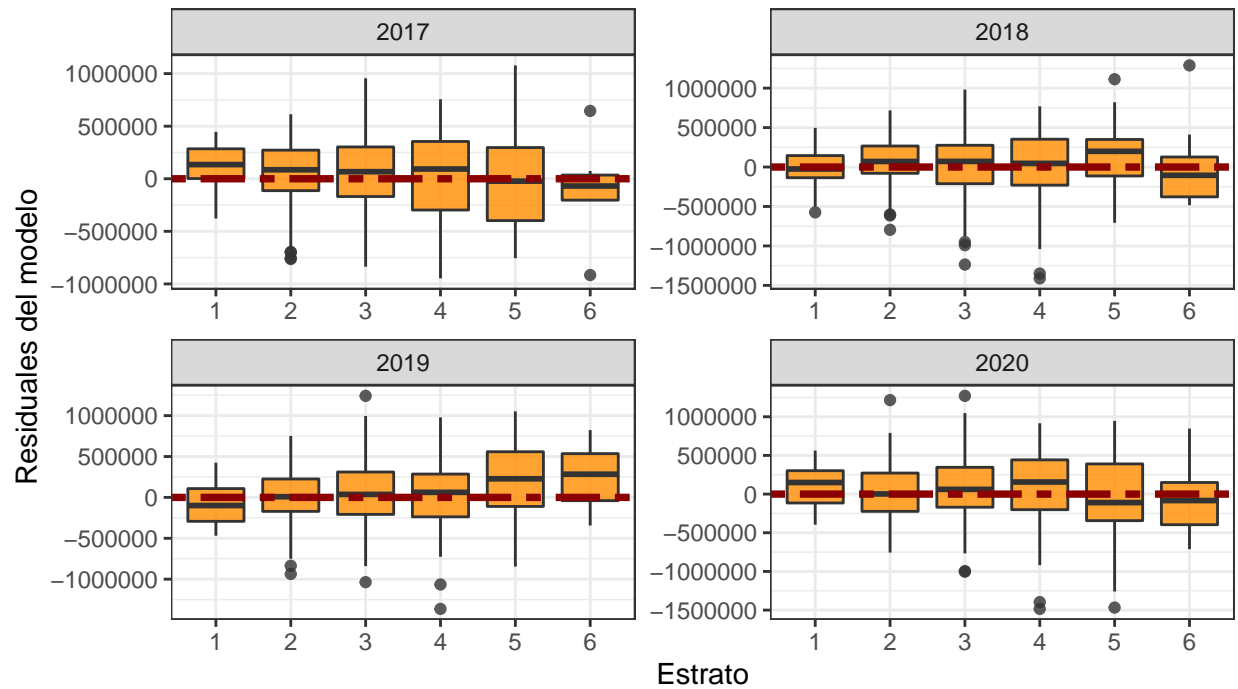


Figura 13: Diagramas de cajas y bigotes de los residuales de la regresión versus los estratos para diferentes vigencias en el modelo de casas.

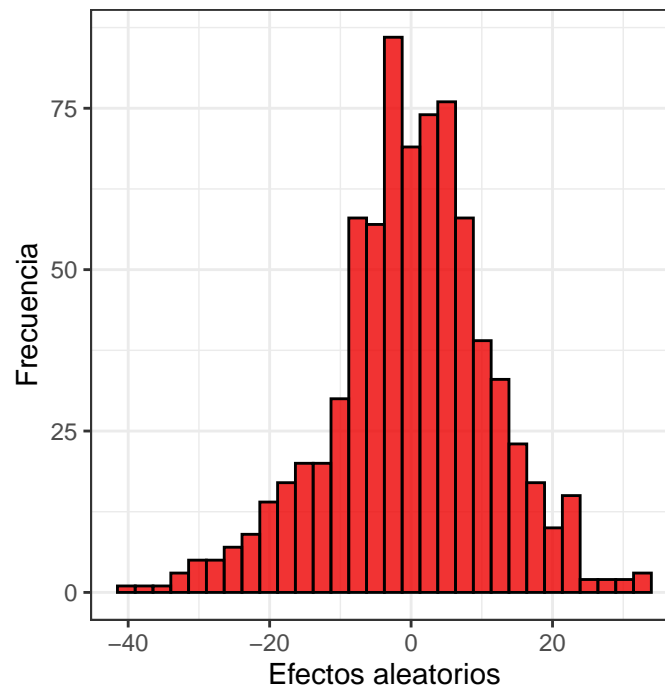


Figura 14: Densidad estimada de los efectos aleatorios asociados a los sectores en el modelo de casas.

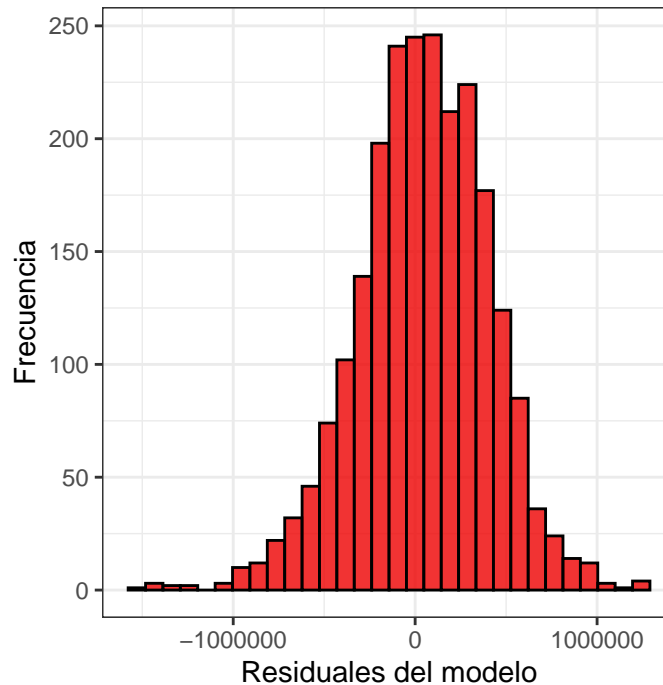


Figura 15: Densidad estimada de los residuales del modelo en el grupo de ofertas referente a casas.

UAECD, superando la limitación de trabajar con fuentes con precios de venta y arriendo para un mismo predio. En consecuencia, se constituye en una herramienta alternativa para obtener tasas de capitalización de rentas.

Los resultados obtenidos son a nivel de sector, año, clase de predio, localidad y estrato. En este documento se omiten los resultados a nivel de sector, debido a su tamaño. Las tasas de rentabilidad obtenidas para casas (NPH) son menores que las calculadas para apartamentos (PH). Se observa una relación inversamente proporcional entre el estrato y las tasas.

Para el caso de las casas, las localidades con una mayor tasa de rentabilidad en casas son Chapinero, Los Mártires y Santa Fé, mientras que las de menor tasa son Bosa y Puente Aranda, mientras que en el caso de apartamentos las localidades con mayores tasas de renta son Chapinero, La Candelaria y San Cristobal. Por otro lado, la localidad con menor tasa de rentabilidad es Puente Aranda.

Con la estimación realizada de las tasas de capitalización, aplicadas a las ofertas de arriendo, se obtiene una estimación de los valores de venta correspondientes, lo cual amplía la cantidad de ofertas disponibles para los procesos requeridos por la Unidad.

## 6. Trabajo futuro

- A partir de una depuración de la información, en futuros ejercicios se podría incluir la edad de los inmuebles como una variable independiente adicional en los modelos para

obtener su tasa de rentabilidad asociada.

- Realizar estimaciones para usos diferentes al residencial.

## 7. Anexo 1

En esta sección se presenta una descripción, de manera corta, del modelo lineal de efectos mixtos, el cual fue utilizado en el documento para realizar las estimaciones requeridas dentro de los objetivos. Adicional a esta descripción se adicionan algunos criterios de validación que también fueron utilizados.

### 7.1. Modelo lineal mixto

Los modelos lineales son una forma de explicar la dispersión de una o más respuestas aleatorias en términos de una serie de variables independientes (también conocidas como exógenas). En [West et al. \[2014, p. 5\]](#) se encuentra un recuento histórico referente a los modelos lineales desde sus inicios en 1861 hasta la actualidad, desde una perspectiva teórica, mencionando al igual avances importantes en cuanto a la implementación de las diferentes metodologías a nivel de paquetes estadísticos. Éstos tienen en cuenta una serie de supuestos lo que hace posible su planteamiento, estimación, interpretación y su respectiva evaluación. Las variables independientes pueden ser clasificadas como efectos fijos o efectos aleatorios. Acorde con [Me-lo et al. \[2007, p. 6\]](#), cuando al finalizar el experimento las conclusiones se formulan sobre un número preestablecido de tratamientos el modelo se denomina de efectos fijos y en este caso la inferencia estadística se hace sobre los efectos medios de los tratamientos, por lo cual aquellas situaciones en que se desean realizar comparaciones o contrastes entre niveles de un factor en búsqueda de diferencias, éste se considera como fijo. Si los niveles de un atributo son una muestra aleatoria de una población de posibles selecciones, es decir, las conclusiones se formulan sobre un número mayor de tratamientos a los usados en el experimento, el modelo se dice de efectos aleatorios, y en este caso, la inferencia estadística se hace sobre las varianzas de los mismos. Los modelos que incluyen ambos se denominan de efectos mixtos (MLM).

En términos matriciales, según [Pinheiro and Bates \[2006, p. 58\]](#), el modelo lineal general con efectos mixtos viene dado de la forma

$$\mathbf{y}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\epsilon}_i, \text{ con } i = 1, \dots, n \text{ y } n = r \times a \quad (4)$$

donde  $\mathbf{y}_i$  es el vector de respuestas,  $\mathbf{X}_i$  y  $\mathbf{Z}_i$  son las matrices de diseño de los efectos fijos y aleatorios asociadas al  $i$ -ésimo individuo, respectivamente. Los vectores  $\boldsymbol{\beta}$  y  $\mathbf{b}_i$  de dimensiones  $a$  y  $q$ , que contienen los coeficientes asociados a los efectos fijos y aleatorios considerados, respectivamente, mientras que  $\boldsymbol{\epsilon}_i$  es el vector de errores de mediciones. Como supuestos distribucionales sobre este modelo se asume que  $\mathbf{b}_i \sim \mathbf{N}_q(\mathbf{0}_q, \boldsymbol{\Psi}_{q \times q})$  y  $\boldsymbol{\epsilon}_i \sim \mathbf{N}_t(\mathbf{0}_t, \boldsymbol{\Sigma}_{t \times t})$ . A su vez los vectores  $\boldsymbol{\epsilon}_i$  y  $\mathbf{b}_i$  se asumen independientes, es decir que  $Cov(\boldsymbol{\epsilon}_i, \mathbf{b}_i) = 0$ . La matriz

## REFERENCIAS

$\Sigma$  asociada a los residuales dentro de las mediciones de la misma unidad puede considerar diferentes estructuras, es decir que se puede asociar a procesos autocorrelacionados, por ejemplo los que se encuentran en la literatura relacionada con series de tiempo, como los modelos *ARIMA* [Wei, 2006], o como los que se estudian en la estadística espacial, que se caracterizan mediante su función de variograma [Schabenberger and Gotway, 2017].

Las metodologías de estimación de los parámetros incluidos dentro del modelo son máxima verosimilitud o máxima verosimilitud restringida, a pesar de que en la literatura se encuentran diferentes propuestas. En este documento se utilizó la primera, ya que es la que se incluye por defecto dentro de la función *lme* del paquete *nlme* [Pinheiro et al., 2021] en R.

## Referencias

- D. Alzate. Tesis de maestría: Estimación del índice de rentabilidad inmobiliaria sectorial – iris para vivienda urbana en el municipio de medellín. url <https://repositorio.unal.edu.co/handle/unal/75163>, 2014.
- Francisca Antman and David J McKenzie. Earnings mobility and measurement error: A pseudo-panel approach. *Economic Development and Cultural Change*, 56(1):125–161, 2007.
- O. Borrero. Avalúo de inmuebles y garantías. bhandar ediciones ltda. url <https://docplayer.es/21369286-Capitulo-9-avaluo-de-arrendamientos-autor-oscar-borrero-ochoa-forma-parte-del-libro-avaluo-de-inmuebles-y-garantias.html>, 2008.
- Lonja de Propiedad Raíz de Medellín. Análisis inmobiliarios 2020 - no. 203 tasas de capitalización en vivienda. url <https://www.lonja.org.co/download/203-tasas-de-capitalizacion-en-vivienda>, 2020.
- Angus Deaton. Panel data from time series of cross-sections. *Journal of Econometrics*, 30(1-2):109–126, 1985.
- M. Godinez. Tesis de maestría: Tasas de capitalización de mercado de casas. url <http://bvirtual.ucol.mx/consultaxcategoria.php?categoria=3id=5358seccion=0>, 2000.
- Klaus Hinkelmann. *Design and analysis of experiments, special designs and applications*, volume 3. John Wiley and Sons, Blacksburg, 2011.
- y Agudelo Torres J. (s.f.) Jaramillo Montoya S. Tesis de maestría: Tasas de capitalización de mercado de casas. url <https://revistas.javeriana.edu.co/index.php/cvyu/issue/view/1291>, 2000.
- H. Medrano J. Perlaza M. y Fuentes. Estimación de tasas de capitalización de rentas de vivienda urbana estrato 3 sometidas al régimen de propiedad horizontal en bogotá mediante modelos econométricos. url <https://revistas.udistrital.edu.co/index.php/UDGeo/article/view/3651>, 2011.
- O Melo, L López, and S Melo. Diseño de experimentos: métodos y aplicaciones. *Editorial Universidad Nacional de Colombia. Bogotá*, 2007.

## REFERENCIAS

- Douglas C Montgomery. *Introduction to statistical quality control*. John Wiley & Sons, 2020.
- Douglas C Montgomery, Elizabeth A Peck, and G Geoffrey Vining. *Introduction to Linear Regression Analysis*, volume 821. John Wiley and Sons, New York, 2012.
- José Pinheiro and Douglas Bates. *Mixed-effects models in S and S-PLUS*. Springer Science & Business Media, New York, 2006.
- Jose Pinheiro, Douglas Bates, Saikat DebRoy, Deepayan Sarkar, and R Core Team. *nlme: Linear and Nonlinear Mixed Effects Models*, 2021. URL <https://CRAN.R-project.org/package=nlme>. R package version 3.1-152.
- R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2021. URL <https://www.R-project.org/>.
- Valorar S.A. Tasas de capitalización de renta en vivienda urbana estrato 2. url<http://valorar.com/2017/10/03/tasas-de-capitalizacion-de-renta>, 2017.
- Murat Sariyar and Andreas Borg. *RecordLinkage: Record Linkage Functions for Linking and Deduplicating Data Sets*, 2020. URL <https://CRAN.R-project.org/package=RecordLinkage>. R package version 0.4-12.1.
- Oliver Schabenberger and Carol A Gotway. *Statistical methods for spatial data analysis*. Chapman and Hall/CRC, Boca Raton, 2017.
- Luca Scrucca. qcc: an r package for quality control charting and statistical process control. *R News*, 4/1:11–17, 2004. URL <https://cran.r-project.org/doc/Rnews/>.
- Maresa Sprietsma. Computers as pedagogical tools in Brazil: a pseudo-panel analysis. *Education Economics*, 20(1):19–32, 2012.
- Ainhoa Oguiza Tovar, Inmaculada Gallastegui Zulaica, and Vicente Núñez-Antón. Analysis of pseudo-panel data with dependent samples. *Journal of Applied Statistics*, 39(9):1921–1937, 2012.
- Chi-Hong Tsai, Corinne Mulley, and Geoffrey Clifton. A review of pseudo panel data approach in estimating short-run and long-run public transport demand elasticities. *Transport Reviews*, 34(1):102–121, 2014.
- William WS Wei. Time series analysis. In *The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2*. 2006.
- Brady T West, Kathleen B Welch, and Andrzej T Galecki. *Linear mixed models: a practical guide using Statistical Software*. CRC Press, Boca Raton, 2014.
- Hadley Wickham, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D’Agostino McGowan, Romain François, Garrett Grolemond, Alex Hayes, Lionel Henry, Jim Hester, Max Kuhn, Thomas Lin Pedersen, Evan Miller, Stephan Milton Bache, Kirill Müller, Jeroen Ooms, David Robinson, Dana Paige Seidel, Vitalie Spinu, Kohske Takahashi, Davis Vaughan, Claus Wilke, Kara Woo, and Hiroaki Yutani. Welcome to the tidyverse. *Journal of Open Source Software*, 4(43):1686, 2019. doi: 10.21105/joss.01686.