Capítulo 4 - Indicadores socioeconómicos y su relación con la estratificación y la información catastral

Javier Acosta
Alexis Maluendas
Guillermo Rivas

SciELO Books / SciELO Livros / SciELO Libros

ACOSTA, J., MALUENDAS, A., and RIVAS, G. Indicadores socioeconómicos y su relación con la estratificación y la información catastral. In: SEPÚLVEDA RICO, C.E., LÓPEZ CAMACHO, D., and GALLEGO ACEVEDO, J.M., eds. *Los límites de la estratificación: en busca de alternativas* [online]. Bogotá: Editorial Universidad del Rosario: Alcaldía Mayor de Bogotá D.C., 2014, pp. 83-107. ISBN: 978-958-738-537-3. <https://doi.org/10.7476/9789587385373.0006>.

All the contents of this work, except where otherwise noted, is licensed under a [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Todo o conteúdo deste trabalho, exceto quando houver ressalva, é publicado sob a licença [Creative Commons Atribuição 4.0](https://creativecommons.org/licenses/by/4.0/).

Todo el contenido de esta obra, excepto donde se indique lo contrario, está bajo licencia de la licencia [Creative Commons Reconocimiento 4.0](https://creativecommons.org/licenses/by/4.0/).

Capítulo 4

Indicadores socioeconómicos y su relación con la estratificación y la información catastral

Javier Acosta, Alexis Maluendas y Guillermo Rivas*

Palabras clave: indicadores socioeconómicos, capacidad de pago, índices de bienestar, correlaciones grupales, errores de inclusión.

Introducción

La riqueza de la información catastral para detectar diferencias en los precios de la tierra y la vivienda ha sido un referente para buscar aproximarse a la capacidad de pago de los hogares. La visión impositiva del catastro se ha enriquecido con otras visiones que permiten tener más información del territorio y las relaciones de quienes lo ocupan. Desde una perspectiva analítica y con algunos resultados empíricos, en el capítulo anterior, Bernal y Tejedor sustentaron las propiedades y cualidades de la información catastral para captar las distintas familias de rentas de suelo y la dinámica de formación de los precios, reflejando la diferenciación socioespacial existente en la ciudad.

También, se estableció el estrecho vínculo entre los precios y las condiciones socioeconómicas de los hogares que habitan los inmuebles residenciales y se propuso una medida del valor unitario integral (VUI), que expresa el avalúo

* Consultores independientes.

por unidad de área y que tiene en cuenta los dos componentes del avalúo: el suelo y la construcción. Esta variable se construye eliminando el estrato en el cálculo para evitar sesgos vigentes en esta clasificación.

Con la georreferenciación del valor unitario integral de los predios clasificados en seis grupos, para comparar su distribución con los estratos socioeconómicos actuales, el capítulo 3 ilustró cómo el VUI logra revelar diferencias que no alcanza a mostrar la estratificación. Esto sugiere que la variable catastral contiene información mucho más rica que permite captar de mejor manera una mayor heterogeneidad en el interior de cada localidad.

Teniendo presente lo anterior, en este capítulo, se busca cumplir con dos objetivos: en primer lugar, profundizar el análisis empírico y evidenciar las fallencias que presenta la estratificación socioeconómica actual, en cuanto a su capacidad para diferenciar grupos con diferentes condiciones socioeconómicas en Bogotá, D.C. En segundo lugar, reforzar los hallazgos realizados desde el punto de vista analítico en el capítulo 3, por medio de la evidencia empírica sobre la estrecha relación entre la información catastral (particularmente el VUI) y las condiciones socioeconómicas de los hogares.

Para hacer esto, se plantean cinco indicadores que dan cuenta, desde diferentes perspectivas, de las características socioeconómicas de los hogares. Se hace primero un análisis inicial de la estratificación vigente y su comportamiento en relación con las características socioeconómicas, para luego recurrir a diferentes métodos estadísticos que evalúen los niveles y el sentido de las asociaciones entre un conjunto de variables catastrales y de acceso a equipamientos con los indicadores sugeridos.

La medición y caracterización de estas relaciones permitirá posteriormente proponer alternativas para la clasificación de los inmuebles residenciales que involucren explícitamente las relaciones halladas. Es decir que, una vez identificada la correspondencia entre las características de los inmuebles y las condiciones socioeconómicas de los hogares que los habitan, se podrá incluir como insumo para una eventual propuesta metodológica de estratificación (capítulo 5).

4.1. Indicadores de las condiciones socioeconómicas

4.1.1. ASPECTOS CONCEPTUALES

Medir las condiciones socioeconómicas de los hogares es un tema complejo. En la literatura disponible al respecto, se encuentra de manera recurrente dos enfoques: uno unidimensional, más comúnmente utilizado, y otro multidimensional, que busca ampliar el contexto de análisis del bienestar socioeconómico. Para evaluar el primero, se considera generalmente la medición del ingreso o el gasto de los hogares como insumo para medir su capacidad de pago. Al tomar el ingreso, se evalúa la utilidad de los individuos sobre el poder de compra de bienes y servicios, y la satisfacción potencial de las necesidades; mientras que, al considerar el gasto como indicador de capacidad de pago, se evalúa el poder de compra efectivo.

El enfoque multidimensional trabaja con los conceptos de funcionamientos y capacidades de Sen (2002), quien frente al ingreso plantea dos tipos de limitaciones: pertinencia y dificultad para medirlo. En el primer caso, lo considera un medio para convertir los recursos en funcionamientos, pero dicha conversión no es universal y excluye aquellos funcionamientos para los cuales el dinero no es el medio más adecuado¹.

4.1.2. INDICADORES SOCIOECONÓMICOS

Esta sección se ha subdividido en una primera parte que presenta los indicadores unidimensionales, y una segunda que expone los detalles y las principales diferencias de los indicadores multidimensionales.

4.1.2.1. *Indicadores unidimensionales*

Se utilizan dos medidas socioeconómicas unidimensionales para medir la capacidad de pago de los hogares: i) índice de capacidad de pago corriente (ICPc) y ii) índice de capacidad de pago priorizada (ICPp); ambos trabajan el concepto de capacidad de pago a partir de la variable gasto *per capita* de los hogares. Su construcción se relaciona con criterios objetivos, como la renta, incorporando los flujos de ingresos provenientes de la actividad económica

¹ Sen entiende el concepto de funcionamientos como las condiciones de vida que pueden ser o no alcanzadas por los individuos, mientras que las capacidades son entendidas por Sen como la habilidad para alcanzar dichas condiciones de vida.

corriente de los individuos. Adicionalmente, permiten introducir elementos progresivos, como la exclusión de las rentas de subsistencia o las que cubren las necesidades básicas (CID, 2012), útiles, entre otras cosas, en el diseño de políticas redistributivas.

Econometría (2006) plantea la construcción de un indicador de capacidad de pago, reconociendo que “... en términos de la teoría económica que se basa en funciones de utilidad y decisiones del consumidor, el indicador ideal de capacidad de pago es el ingreso permanente de los hogares una vez descontado los gastos de subsistencia”:

$$ICP_i = YP_i - GS_i$$

donde ICP_i : indicador de capacidad de pago para el hogar i
 YP_i : ingreso permanente del hogar i
 GS_i : gastos de subsistencia del hogar i

Retomando este planteamiento, se incorpora el gasto total (G) como *proxy* débil del ingreso permanente y el gasto de alimentos (GA) como *proxy*, también débil, de los gastos de subsistencia; ambos en términos *per capita* para eliminar el efecto del tamaño de los hogares. De esta forma, se construye el indicador de capacidad de pago corriente (ICPc), para un hogar i , como:

$$ICPc_i = G_i - GA_i$$

De otro lado, el ICPp es la aproximación que hace el Centro de Investigación y Desarrollo de la Facultad de Economía de la Universidad Nacional de Colombia (CID-UNAL) (CID, 2012), el cual se aleja del enfoque convencional del consumidor y la capacidad de pago, al introducir un indicador basado en ordenamientos jerárquicos o lexicográficos de la necesidades. A partir de este enfoque, se reconoce que no existe una completa sustituibilidad de las preferencias y que el gasto de los consumidores más bien sigue un orden jerárquico de necesidades, en el que existen unas necesidades más urgentes que otras (CID, 2012). El gasto mínimo en bienes básicos (GMBB) se construye de tal manera que los hogares que tengan un gasto igual o menor a dicho gasto mínimo no tienen capacidad de pago y, por lo tanto, el indicador es igual a cero, y, para los hogares que están por encima de tal umbral, la capacidad

de pago va aumentando conforme los hogares van incrementando el gasto, hasta tomar un valor máximo de 1. En términos generales, el procedimiento seguido por el CID cumplió las siguientes etapas: 1) Dividir el total de hogares de la muestra de la *EMB2011* en tres grupos de acuerdo con su nivel de ingreso; esto con el fin de tomar en consideración el hecho de que no todos los hogares alcanzan a consumir toda la oferta de bienes y servicios del mercado. 2) Definir tres grupos de bienes: básicos, de confort y bienes de lujo; esto, en conjunto con la definición de los hogares, permite identificar niveles de vida diferentes. 3) Estimación de modelos de demanda sobre los bienes jerarquizados; particularmente estiman el sistema lineal de gastos, toda vez que uno de sus parámetros de la forma estructural corresponde al gasto de subsistencia para cada grupo de bienes considerado en el sistema.

4.1.2.2. Indicadores multidimensionales

Como indicadores para estimar el nivel de bienestar de los hogares desde un enfoque multidimensional se proponen el índice de condiciones de vida (ICV), construido por el Departamento Nacional de Planeación en 2004; el índice de calidad de vida urbana (ICVU), desarrollado por el CID (2012) para el Distrito Capital; y el índice de bienestar socioeconómico (IBS), elaborado por el DANE (2011) como parte de su propuesta para definir una nueva metodología de estratificación socioeconómica.

El primero de ellos (ICV) es una medida tradicional de la calidad de vida, que ha sido calculada recurrentemente para determinar la evolución de las condiciones de vida de las personas en Colombia. El segundo (ICVU) es una propuesta que involucra aspectos urbanos como calidad ambiental, no discriminación, solidaridad y asociación. Y el tercero (IBS) tiene en cuenta aspectos como dotación de enseres y protección frente a fluctuaciones económicas. Los tres indicadores tienen en común que su base conceptual parte de un enfoque de funcionamientos y capacidades (Sen, 2002) y su metodología de construcción es similar: 1) elección de funcionamientos como estándar de vida, 2) definición de variables para medir cada funcionamiento, 3) elección de la herramienta estadística para definir las ponderaciones de cada variable y 4) estandarización del indicador resultante para que tome valores en el rango de 0 a 100, siendo 0 las peores condiciones y 100 el máximo nivel de bienestar.

Las diferencias propias de cada indicador en las distintas etapas van a estar condicionadas por las dimensiones incorporadas y el tipo de ponderación de variables. La tabla 4.1 sintetiza las principales diferencias:²

Tabla 4.1. Principales diferencias entre los indicadores multidimensionales utilizados

Aspecto	ICV	ICVU	IBS
Funcionamientos o dimensiones	1) Servicios básicos de la vivienda 2) Capital humano - educación 3) Sociodemográfico 4) Calidad de la vivienda	1) Salud y alimentación 2) Condiciones de la vivienda 3) Calidad ambiental 4) Equipamientos y dotaciones urbanas 5) Movilidad 6) Seguridad ciudadana 7) Trabajo 8) Educación 9) Ocio y recreación 10) Solidaridad y asociación 11) No discriminación 12) Capacidad de pago	1) Vivir en un hogar seguro y saludable 2) Habitar una vivienda digna 3) Acceder a instalaciones sanitarias 4) Acceder a la dotación de enseres 5) Sentirse a salvo ante fluctuaciones económicas
Metodología estadística	1) Cuantificación de variables a partir del ACP cualitativas 2) Construcción de un indicador por dimensión con ACP 3) Construcción del índice sintético final por ACP sobre los indicadores por dimensión 4) Estandarización del índice a un rango de variación de 0 a 100	1) Cuantificación de variables a partir del ACP cualitativas por dimensión y subdimensión 2) Construcción de un indicador por dimensión con ACP 3) Estandarización a un rango de variación de 0 a 100 4) Construcción de un indicador sintético global como el promedio simple de los indicadores por dimensión	1) Cuantificación de variables a partir del ACM y construcción de un indicador por dimensión 2) Estandarización a un rango de variación de 0 a 100 3) Construcción de un indicador sintético global como el promedio simple de los indicadores por dimensión

ACP: análisis de componentes principales

ACM: análisis de correspondencia múltiple

Fuente: elaboración de los autores a partir de CID (2012), DANE (2011) e información del DNP.

4.1.3. RESUMEN DE ÍNDICES PROPUESTOS

En la figura 4.1, se muestran las distribuciones de frecuencias de los índices de capacidad de pago para Bogotá (ICPc e ICPp) a partir de la *Encuesta multipropósito de Bogotá 2011 (EMB2011)*. Ambos tienen distribuciones con un

² En el anexo 4, se describe de forma resumida la construcción de cada uno de los indicadores multidimensionales.

patrón de comportamiento esperado al presentar asimetrías marcadas hacia la derecha, con lo cual se está indicando que la gran mayoría de hogares, 60% o más, tienen capacidad de pago baja.

Figura 4.1. Distribución índice de capacidad de pago en Bogotá 2011

Fuente: elaboración de los autores a partir de la EMB 2011.

En el caso del ICPp, alrededor del 20% de los hogares no tiene capacidad de pago, lo cual, en este caso, significa que presentan un gasto *per capita* por debajo del gasto mínimo en bienes básicos (GMBB).

Figura 4.2. Distribución de los índices en Bogotá a partir de la EMB2011

Fuente: elaboración de los autores a partir de la EMB 2011.

La figura 4.2 ilustra las distribuciones de los índices multidimensionales, todas con un comportamiento diferente entre sí y en comparación con las medidas de capacidad de pago. El ICVU tiene una distribución simétrica, indicando que la gran mayoría de los hogares se sitúa en el rango que va de 0,4 a 0,7, con un valor central alrededor de 0,55, y los restantes se distribuyen de manera similar en las colas; es decir, hay relativamente pocos hogares en Bogotá con condiciones de vida urbanas extremas (muy desfavorables o muy favorables) respecto a equipamientos y dotaciones urbanas, calidad ambiental, movilidad, ocio y recreación y demás dimensiones consideradas en el indicador. En el caso del ICV, la distribución presenta una asimetría hacia la izquierda, indicando que buena parte de los hogares, por lo menos el 70 %, goza de condiciones favorables en términos de servicios básicos y calidad de la vivienda y educación. Por su parte, el IBS es relativamente simétrico alrededor de 0,82, con una distribución alargada hacia la izquierda, mostrando que una baja proporción de los hogares, 8 % o menos, tiene condiciones desfavorables respecto a vivienda digna, condiciones sanitarias, acceso a dotación de enseres y protección a fluctuaciones económicas.

4.1.3.1. Análisis de los indicadores según localidades

Esta sección presenta una descripción general de las medidas de bienestar, por localidades, y corrobora su validez en reflejar, de manera complementaria, diferenciales socioeconómicos de los hogares. En este sentido se ilustran las similitudes en cuanto a ordenamiento por localidades, pero de igual manera, sus diferencias respecto a varianza entre ellas (tabla 4.2 y figura 4.1).

Tabla 4.2. Comportamiento del promedio de los indicadores socioeconómicos según localidades urbanas en Bogotá, D.C.

Localidad	IBS	ICV	ICVU (%)	ICPc (millones de pesos)	ICPp (% con CP)
Usme	0,739	0,852	0,460	0,284	0,631
Ciudad Bolívar	0,745	0,857	0,480	0,309	0,618
San Cristóbal	0,755	0,869	0,490	0,338	0,659
Bosa	0,756	0,870	0,500	0,346	0,716
Rafael Uribe Uribe	0,761	0,882	0,510	0,372	0,719
Tunjuelito	0,773	0,889	0,550	0,451	0,788

Continúa

Localidad	IBS	ICV	ICVU (%)	ICPc (millones de pesos)	ICPp (% con CP)
Kennedy	0,796	0,899	0,530	0,548	0,824
Antonio Nariño	0,793	0,911	0,560	0,594	0,887
Los Mártires	0,786	0,910	0,520	0,656	0,866
Puente Aranda	0,817	0,926	0,570	0,662	0,876
Engativá	0,821	0,921	0,560	0,692	0,888
Santa Fe	0,780	0,899	0,520	0,868	0,798
Suba	0,831	0,923	0,560	0,881	0,880
La Candelaria	0,774	0,913	0,540	0,944	0,832
Barrios Unidos	0,834	0,933	0,580	0,955	0,904
Fontibón	0,845	0,928	0,590	1,007	0,886
Usaquén	0,877	0,947	0,620	1,626	0,924
Teusaquillo	0,887	0,961	0,630	1,630	0,940
Chapinero	0,891	0,967	0,620	2,576	0,924

Fuente: elaboración de los autores a partir de la EMB 2011.

La tabla 4.2 está ordenada de manera ascendente según el ICPc. En primer lugar, se observa una correlación positiva entre todos los indicadores, de manera que las localidades que tienen valores bajos en las medidas de la capacidad de pago también tienen valores bajos en los indicadores de calidad de vida. En segundo lugar, la varianza de cada indicador muestra comportamientos diferentes a lo largo de las localidades. Respecto a los multidimensionales, el ICVU es el que presenta mayor variabilidad, patrón explicado probablemente al definirlo en sus inicios particularmente para Bogotá, con el fin de captar precisamente la gran diversidad de condiciones socioeconómicas que se presentan en la ciudad.

La figura 4.3 compara los promedios de cada indicador por localidad. Evidencia mayores distancias para el ICPc, destacando los promedios superiores de las localidades de Teusaquillo, Usaquén y Chapinero. Se observa también que en las localidades donde dominan ampliamente los estratos 1 y 2, como Ciudad Bolívar y Usme, se dan los valores más bajos en todos los indicadores, mientras que en las localidades donde se ubican la mayoría de los estratos 4, 5 y 6, como Chapinero y Usaquén, éstos reportan valores altos.

Figura 4.3. Comparación de los promedios de los indicadores por localidad

Fuente: elaboración de los autores a partir de la EMB 2011.

Las medidas seleccionadas para medir las condiciones de vida y la capacidad de pago de los hogares reflejan entonces aspectos diferenciales de estándar de vida, lo cual permite tener mayor robustez al estudiar la estratificación actual en términos de posibles errores de inclusión (hogares sin condiciones para ser subsidiados son clasificados en estratos bajos) y exclusión (hogares con las condiciones para ser subsidiados son clasificados en estratos altos); así como al estudiar la relación entre las condiciones socioeconómicas de los hogares con información complementaria, (variables catastrales por ejemplo) que lleve a nuevas metodologías de estratificación de inmuebles y definición de subsidios en SPD.

4.2. Problemas de la estratificación actual

La estratificación socioeconómica que opera actualmente en Bogotá para el cobro de los servicios públicos domiciliarios, si bien es una clasificación de inmuebles implementada a partir de características arquitectónicas y de entorno de estos, tiene como propósito implícito hacer una aproximación a la clasificación de los hogares que los habitan, de tal manera que permita la determinación de subsidios y contribuciones de los usuarios. En otras palabras, se establece un sistema para asignar subsidios a aquellos domicilios de ‘menores condiciones socioeconómicas’ y recaudar una contribución a aquellos de ‘mayores condiciones socioeconómicas’.

Con el fin de analizar si la estratificación vigente de inmuebles residenciales tiene la capacidad adecuada de clasificar y discriminar entre los diferentes *niveles socioeconómicos* expresados a través de los deciles de los indicadores propuestos, se evalúa en esta sección la probabilidad de que un hogar pertenezca a un estrato socioeconómico dado el *nivel socioeconómico*. Este ejercicio se realiza para cada uno de los cinco indicadores presentados en la sección anterior, de acuerdo con el siguiente procedimiento:

- 1) Se clasifican los hogares por deciles de cada indicador.
- 2) Dentro de cada decil, se calcula el número de hogares que pertenecen a cada uno de los estratos socioeconómicos y el número total de hogares en el decil.
- 3) Se define la probabilidad de que el hogar pertenezca a un estrato, como el cociente entre el número de hogares del estrato i en el decil j y el

número total de hogares del decil j , como se puede ver en la siguiente expresión:

$$P(\text{Hogar } c \text{ Estrato } i \mid \text{Decil } j) = \frac{\text{Número de hogares del estrato } i \text{ en el decil } j}{\text{Número de hogares en el decil } j}$$

$$i = 1, 2, \dots, 6; j = 1, 2, \dots, 10.$$

- 4) Si la ‘estratificación socioeconómica’ actual, basada en la clasificación de inmuebles con la metodología definida a mediados de los noventa, cataloga adecuadamente los hogares, lo que se espera ver en los gráficos de probabilidad son seis líneas donde las áreas debajo de estas se superponen poco. En un escenario ideal, la probabilidad de pertenecer a estratos bajos debería disminuir a medida que aumenta el decil de capacidad de pago o bienestar. De igual manera, a medida que los hogares tienen una mayor capacidad de pago, debe aumentar la probabilidad de habitar en un inmueble de estrato alto. Esta relación ideal entre probabilidad de asignación de estrato y capacidad de pago (o bienestar) se puede ver de dos formas: por un lado, el primer decil, por ejemplo, debe tener una mayor probabilidad de habitar en un inmueble de estrato 1, seguido por la probabilidad de que un inmueble esté en estrato 2 y una probabilidad muy cercana a cero de estar en estrato 6. Por otro lado, la probabilidad de estar clasificado en estrato 1 debe ser mayor para el decil 1 y descender gradualmente a medida que los deciles aumentan.

En la figura 4.4, se presentan las probabilidades o frecuencias relativas de hogares pertenecientes a los diferentes estratos, eje vertical, según los deciles de los cinco indicadores, ejes horizontales.

El resultado es contundente y muy similar para los cinco indicadores propuestos. Si bien es cierto que la probabilidad de estar en estrato 1 (línea roja) presenta un patrón de comportamiento esperado, en el sentido de tener la mayor probabilidad de pertenecer a este estrato en el primer decil y de disminuir a medida que se incrementan los deciles, y, de manera inversa, para los estratos 4, 5 y 6 (a medida que aumenta el decil, se incrementa la probabilidad de pertenecer a ellos), la distribución de probabilidad de pertenecer a estratos 2 o 3 domina a las demás distribuciones a lo largo de todos los deciles.

Figura 4.4. Probabilidad de pertenecer al estrato socioeconómico i , dados los deciles de cada uno de los indicadores socioeconómicos

Fuente: elaboración de los autores a partir de la EMB 2011.

Es decir, para hogares ubicados en deciles 1 al 6, la probabilidad de estar asignados a un estrato 2 es significativamente mayor comparada con los de-

más estratos, mientras que en los deciles 7 al 9 existe una probabilidad muy superior de estar en estrato 3.

Esta relación de predominancia de asignación de estrato 2 o 3 para cualquier decil, tanto de capacidad de pago como de calidad de vida o de bienestar, evidencia el hecho de que la ‘estratificación socioeconómica’ actual no logra discriminar de manera clara los hogares a partir de sus condiciones de vida a pesar de existir cierto grado de correlación entre asignación de estrato y decil de condición de vida.

4.3. Relación entre variables catastrales e indicadores socioeconómicos

Las deficiencias de la estratificación socioeconómica para expresar la dinámica socioeconómica dan pie para evaluar la pertinencia de incorporar otros insumos en la metodología de clasificación de inmuebles residenciales en servicios públicos domiciliarios, como la información catastral y de acceso a equipamientos. Para indagar sobre esta posibilidad, se presentan dos tipos de ejercicios; el primero es un análisis exploratorio de datos con medidas clásicas de correlación para seleccionar un conjunto más reducido de variables. Con el segundo análisis, que utiliza herramientas estadísticas multivariadas, se busca estudiar el nivel de asociación existente entre las variables seleccionadas en el primer análisis y destacar las variables catastrales y de acceso que mejor discriminan el componente socioeconómico.

La información catastral hace referencia al conjunto de variables que son recolectadas y administradas por el catastro distrital, las cuales describen física y económicamente los predios y su ubicación en el espacio urbano de la ciudad (capítulo 3). Dentro de este grupo, el valor unitario integral (VUI), obedece a la construcción descrita en la sección 3.2.1 del capítulo 3, y representa el valor del avalúo sobre el área construida, para inmuebles en propiedad horizontal (PH), y la sumatoria de los valores de metro cuadrado del terreno y de la construcción para no PH.

Las mediciones de acceso a equipamientos y dotación de bienes públicos y privados corresponden a otro grupo de variables calculadas a partir de la información de la *EMB2011* y de la geodatabase de la SDP. Mayor detalle respecto a la construcción de este conjunto de variables se detalla en las secciones 1 y 3 del anexo 3.

4.3.1. ANÁLISIS EXPLORATORIO INICIAL

Con el fin de depurar el número de variables a ser utilizadas en los análisis posteriores, se estiman las correlaciones lineales entre los indicadores socioeconómicos y las variables catastrales y de acceso a equipamientos. En la tabla 4.3, se muestran los coeficientes de correlación y el nivel de significancia entre los dos grupos.

Tabla 4.3. Coeficiente de correlación lineal entre indicadores socioeconómicos y variables catastrales y de acceso

	Variables catastrales y de acceso	ICPp		ICPc		ICVU		ICV		IBS	
Variables catastrales	Valor unitario integral sin efecto de estrato	0,29	***	0,67	***	0,48	***	0,46	***	0,57	***
	Puntaje	0,27	***	0,57	***	0,50	***	0,50	***	0,60	***
	Acabados principales	0,27	***	0,54	***	0,49	***	0,49	***	0,59	***
	Valor m ² construcción	0,27	***	0,59	***	0,43	***	0,42	***	0,53	***
	Valor m ² terreno	0,28	***	0,56	***	0,43	***	0,42	***	0,48	***
	Estructura	0,25	***	0,54	***	0,42	***	0,42	***	0,55	***
	Baño	0,25	***	0,51	***	0,45	***	0,44	***	0,52	***
	Cocina	0,24	***	0,48	***	0,45	***	0,45	***	0,53	***
	Avalúo	0,16	***	0,28	***	0,24	***	0,21	***	0,24	***
	Área de construcción	-0,05	***	-0,14	***	-0,07	***	-0,12	***	-0,18	***
	Área de terreno	-0,04	***	-0,12	***	-0,09	***	-0,10	***	-0,15	***
	Vetustez	-0,05	***	-0,14	***	-0,05	***	-0,06	***	-0,20	***
Variables de acceso a equipamientos	Distancia colegio	0,18	***	0,33	***	0,25	***	0,23	***	0,31	***
	Índice acceso b. públicos	0,06	***	0,13	***	0,41	***	0,19	***	0,16	***
	Distancia hospital nivel III	-0,12	***	-0,18	***	-0,14	***	-0,16	***	-0,14	***
	Índice acceso b. privados	0,05	***	0,11	***	0,34	***	0,15	***	0,13	***
	Distancia UBA	0,08	***	0,12	***	0,13	***	0,13	***	0,15	***
	Distancia jardín	0,07	***	0,17	***	0,10	***	0,08	***	0,16	***

	Variables catastrales y de acceso	ICPp		ICPc		ICVU		ICV		IBS	
Variables de acceso a equipamientos	Índice problemas sector	0,05	***	0,09	***	0,32	***	0,05	***	0,10	***
	Ditancia al trabajo - máximo	-0,07	***	-0,15	***	-0,13	***	-0,10	***	-0,01	
	Distancia a relleno	0,03	***	0,04	***	0,05	***	0,05	***	0,04	***
	Índice global de parques	-0,01		0,02	**	0,08	***	0,06	***	0,11	***
	Distancia a ríos	0,05	***	0,11	***	0,01		0,01		0,03	***
	Distancia a CAI	-0,04	***	-0,06	***	-0,01	*	-0,04	***	-0,00	
	Distancia a parques	0,00		-0,01		-0,04	***	-0,03	***	-0,08	***
	Distancia CADE	0,02	***	0,01	*	0,00		-0,01		0,03	***
	Distancia a humedales	-0,00		0,03	***	-0,00		-0,02	**	-0,04	***

*** Estadísticamente significativas al 99 %.

** Estadísticamente significativas al 95 %.

* Estadísticamente significativas al 90 %.

Correlaciones en color rojo son estadísticamente no significativas.

Fuente: elaboración de los autores a partir de la EMB 2011.

Del conjunto de variables catastrales, se destacan las correlaciones de los valores por unidad de área (valor m² terreno, valor m² construcción y valor unitario integral) y el puntaje catastral, los cuales presentan los niveles de asociación lineal más altos con los indicadores socioeconómicos. Los componentes del puntaje, por su parte, muestran correlaciones positivas con los indicadores socioeconómicos, aunque ninguno evidencia correlaciones superiores a las observadas con el puntaje total, por lo que no es necesario incluir estos desagregados. Del conjunto de variables de acceso, como las más relevantes están las distancias al colegio y a un hospital nivel III, y los indicadores de acceso a bienes públicos y privados.

Las variables asociadas a las áreas de los predios y la vetustez (edad de la construcción), a pesar de tener correlaciones estadísticamente significativas, muestran niveles bajos de asociación con las socioeconómicas³. Lo mismo

³ Se infiere de estos resultados que el tamaño del predio no se asocia con los niveles socioeconómicos, fenómeno que puede ser explicado por la gran proporción de predios en PH en la ciudad.

ocurre con las distancias a CAI, parques, CADE y humedales que presentan correlaciones más bajas y, en algunos casos, no significativas estadísticamente.

Las correlaciones lineales del valor unitario integral con los indicadores socioeconómicos resultan significativas y con niveles más altos (hasta del 67%) a los observados con el avalúo catastral directamente. Por otro lado, los componentes del puntaje muestran correlaciones importantes con los indicadores socioeconómicos, aunque ninguno de los cuatro componentes evidencia correlaciones superiores a las observadas con el puntaje total, por lo que no es necesario incluir estos desagregados.

A pesar de que una gran parte de las variables de acceso resulta con correlaciones significativas estadísticamente, la mayoría de las que hacen referencia a distancias a equipamientos presentan signos contraintuitivos. Este es el caso de las distancias al colegio, a UBA, a jardín, a parques y a CADE. Este fenómeno se presenta porque el sentido de la relación se comporta de diferente forma dependiendo del nivel socioeconómico. Por ejemplo, en los niveles socioeconómicos altos, los establecimientos educativos a los que asisten los estudiantes se encuentran a gran distancia del hogar, incluso, en muchos casos, fuera de la ciudad, lo que distorsiona la medida de correlación.

Así las cosas, el conjunto de variables catastrales se focaliza en el siguiente subconjunto para efectos de los análisis multivariados, al mostrar las correlaciones más altas con las medidas socioeconómicas: valor m^2 terreno, valor m^2 de construcción, valor unitario integral y puntaje. Adicionalmente, se incluyen las variables de acceso: distancia a hospital nivel III, índice de acceso a bienes públicos, índice de acceso a bienes privados y distancia al trabajo.

4.3.2. CORRELACIONES O ASOCIACIONES ENTRE LOS DIFERENTES CONSTRUCTOS

En esta sección, se desarrollan cuatro análisis multivariados para observar la estructura de las correlaciones entre el conjunto de los indicadores socioeconómicos y el conjunto de las variables catastrales y de acceso:

- 1) *Análisis de componentes principales (ACP)*, con el fin de determinar si existen asociaciones entre todo el conjunto de variables evaluadas, y qué variables se asocian más con otras.
- 2) *Análisis de correlación canónica (ACC)*. El ACC permite establecer el nivel de correlación entre dos bloques de variables. En este caso calcula

el grado en que las variables catastrales pueden explicar, en conjunto, los fenómenos socioeconómicos expresados en los cinco indicadores juntos.

- 3) *Análisis discriminante lineal (ADL)*. Permite identificar, de un grupo de covariables, cuáles discriminan mejor en una clasificación de una variable de interés. En este caso, determina las variables catastrales y de acceso que discriminan de mejor forma cada uno de los indicadores socioeconómicos por aparte.
- 4) *Análisis de regresión logística (ARL)*. Similar al ADL, aunque, además de identificar las variables catastrales que discriminan mejor, presenta la medida en que lo hacen.

4.3.2.1. Resultados del ACP

Se procede a realizar el ACP con las variables estandarizadas, para evitar influencia de las unidades de medida. Los resultados se presentan en la figura 4.5. En el gráfico de la izquierda, se muestra el porcentaje de la inercia (varianza), que logra explicar cada uno de los componentes generados, y, en el de la derecha, la proyección de cada una de las variables de análisis sobre los dos primeros componentes. En el círculo de correlaciones, el largo de las líneas asociadas a cada variable establece el grado de relación con cada componente y la dirección en que se proyectan expresa el sentido de la relación (positiva o negativa).

Del gráfico de la izquierda, se deduce que, con los dos primeros componentes (porcentajes de las dos primeras barras), se retiene aproximadamente un 56 % de la varianza total de las variables de análisis. Se construye entonces el círculo de correlaciones a partir de los dos primeros componentes (gráfico de la derecha), donde se observa cómo los cinco indicadores socioeconómicos se proyectan en el mismo sentido de las variables valor m² terreno (VM2T), valor m² construcción (VM2C), valor unitario integral (VUI) y puntaje (*Punt*), indicando correlaciones positivas. Los indicadores de acceso a bienes públicos (Ind.Acc.Pub) y privados (Ind.Acc.Pri) se proyectan en un sentido más próximo al ICVU, como era de esperarse, lo que indica mayor asociación a este indicador. Finalmente, las variables distancia al trabajo (Dist.Trab) y distancia a hospital nivel III (Dis.Hos.NIII) se proyectan en el sentido contrario de los indicadores y las variables catastrales, indicando una correlación inversa (altos valores de los indicadores se asocian a bajos valores de estas variables),

Figura 4.5. Inercia de los ejes factoriales generados por el ACP y círculo de correlaciones

Fuente: elaboración de los autores a partir de la EMB 2011.

sin embargo, las líneas proyectadas son cortas, manifestando, a su vez, una baja correlación.

Del ACP anterior, se concluye que ICPc, ICPp, IBS e ICV están altamente correlacionados con las variables catastrales asociadas a valores de construcción, suelo y valor unitario integral, así como al puntaje catastral.

4.3.2.2. Resultados del ACC

Se realiza un ACC tomando dos conjuntos de variables: el primero conformado por los cinco indicadores socioeconómicos revisados en la primera sección de este capítulo; el segundo, por el conjunto de variables catastrales y de acceso definidas en el primer análisis de esta sección. A través del ACC, se construyen una serie de variables canónicas para cada conjunto, que son el resultado de combinaciones lineales sobre las variables originales. La primera variable canónica presenta la más alta correlación que puede existir entre ambos conjuntos.

La figura 4.6 (izquierda) muestra, en porcentaje, la correlación alcanzada por las variables canónicas generadas. Se observa en la primera barra que la más alta correlación encontrada entre los dos conjuntos es del 76,7%, siendo estadísticamente significativa.

Figura 4.6. Correlaciones canónicas y cargas sobre la variable canónica socioeconómica

Fuente: elaboración de los autores a partir de la EMB 2011.

En la figura 4.6 (derecha), se registran las cargas de las variables catastrales sobre la primera variable canónica del componente socioeconómico, ordenadas de mayor a menor en valores absolutos. El valor unitario integral y el puntaje de la construcción presentan las más altas correlaciones con la variable canónica socioeconómica, indicando que estas dos variables logran explicar las características socioeconómicas en alrededor de un 74 %.

4.3.2.3. Resultados del ADL

A través de este procedimiento, se identifican las variables catastrales y de acceso que discriminan mejor entre unas clases previamente definidas de los indicadores socioeconómicos. Para realizar el ejercicio, se clasifican inicialmente los cinco indicadores socioeconómicos en deciles. Cada una de estas clasificaciones se utiliza como respuesta en un modelo discriminante lineal, teniendo como explicativas el conjunto de variables catastrales y de acceso seleccionadas en la sección 4.3.1.

En la tabla 4.3, se encuentra el resultado de realizar el proceso *stepwise* sobre los deciles de cada uno de los cinco indicadores, mostrando los pasos de la selección de variables de acuerdo con su importancia, el coeficiente de

ajuste (R^2) y el estadístico Wilks' lambda, que mide la proporción de la varianza de los indicadores socioeconómicos que no es explicada por cada una de las variables catastrales y de acceso⁴.

Se puede entonces inferir, primero, que las variables catastrales permiten discriminar mejor en los deciles de ICPc y con menor calidad el ICPp. Adicionalmente, el valor unitario integral es la variable que mejor discrimina los dos indicadores de capacidad de pago (ICPp e ICPc), mientras que el puntaje discrimina mejor en los indicadores de calidad de vida (ICVU, ICV e IBS).

En relación con las variables de acceso, el índice de acceso a bienes privados es el que menos discrimina, contrastado con el índice de acceso a bienes públicos que discrimina mejor, especialmente en el ICVU.

Tabla 4.4. *Stepwise* para los deciles del ICPp

Indicador socioeconómico	Paso	Variable	R-cuadrado parcial	Valor F	Pr > F	Wilks' lambda	Pr < lambda
ICPp	1	Valor unitario integral	0,108	159,35	<,0001	0,892	<,0001
	2	Puntaje	0,009	11,46	<,0001	0,884	<,0001
	3	Valor m ² terreno	0,009	11,29	<,0001	0,876	<,0001
	4	Valor m ² construcción	0,008	10,57	<,0001	0,869	<,0001
	5	Distancia hospital nivel III	0,007	8,53	<,0001	0,864	<,0001
ICPp	6	Distancia al trabajo	0,003	3,4	0,000	0,862	<,0001
	7	Índice de acceso a b. públicos	0,002	2,57	0,006	0,860	<,0001
ICPc	1	Valor unitario integral	0,507	1358,17	<,0001	0,493	<,0001
	2	Puntaje	0,077	109,17	<,0001	0,455	<,0001
	3	Valor m ² construcción	0,027	36,8	<,0001	0,443	<,0001
	4	Valor m ² terreno	0,029	38,67	<,0001	0,430	<,0001
	5	Índice de acceso a b. públicos	0,018	23,43	<,0001	0,422	<,0001
	6	Distancia hospital nivel III	0,013	17,06	<,0001	0,417	<,0001
	7	Distancia al trabajo	0,003	3,88	<,0001	0,416	<,0001
	8	Índice de acceso a b. privados	0,002	2,05	0,030	0,415	<,0001

⁴ El estadístico Wilks' lambda indica que cada variable catastral o de acceso explica más las diferencias socioeconómicas (según cada indicador) en la medida que se acerca a 0 y explica menos esas diferencias en la medida que se acerca a 1.

Indicador socioeconómico	Paso	Variable	R-cuadrado parcial	Valor F	Pr > F	Wilks' lambda	Pr < lambda
ICVU	1	Puntaje	0,249	435,04	<,0001	0,751	<,0001
	2	Índice de acceso a b. públicos	0,139	211,07	<,0001	0,647	<,0001
	3	Valor unitario integral	0,019	25,8	<,0001	0,634	<,0001
	4	Valor m ² construcción	0,008	10,29	<,0001	0,629	<,0001
	5	Distancia hospital nivel III	0,002	2,17	0,021	0,628	<,0001
	6	Distancia al trabajo	0,001	1,6	0,109	0,628	<,0001
ICV	1	Puntaje	0,302	566,72	<,0001	0,698	<,0001
	2	Valor m ² construcción	0,026	35,59	<,0001	0,680	<,0001
	3	Valor m ² terreno	0,016	21,72	<,0001	0,669	<,0001
	4	Índice de acceso a b. públicos	0,010	12,98	<,0001	0,662	<,0001
	5	Distancia hospital nivel III	0,008	10,74	<,0001	0,657	<,0001
	6	Valor unitario integral	0,005	6,35	<,0001	0,654	<,0001
	7	Índice de acceso a b. privados	0,001	1,57	0,118	0,653	<,0001
IBS	1	Puntaje	0,426	975,11	<,0001	0,574	<,0001
	2	Valor unitario integral	0,047	64,66	<,0001	0,547	<,0001
	3	Distancia al trabajo	0,014	18,96	<,0001	0,540	<,0001
	4	Valor m ² construcción	0,009	11,61	<,0001	0,535	<,0001
	5	Índice de acceso a b. públicos	0,007	8,93	<,0001	0,531	<,0001
	6	Valor m ² terreno	0,003	3,97	<,0001	0,530	<,0001
	7	Distancia hospital nivel III	0,002	2,2	0,019	0,529	<,0001

Fuente: elaboración de los autores a partir de la EMB 2011.

4.3.2.4. Resultados del ARL

Para reforzar los resultados del poder discriminante de las variables catastrales y de acceso, se realiza la estimación de un modelo lineal generalizado (MLG), que permite establecer la ponderación de cada covariable en la diferenciación de las clases generadas de la variable dependiente (indicadores socioeconómicos).

Para llevar a cabo este análisis, se parte nuevamente de la clasificación en deciles de cada indicador y, posteriormente, se estandarizan, por media y varianza, las variables catastrales y de acceso. Adicionalmente, se agrega la

clase de predio (PH y NPH) y se estima un MLG en la forma de una regresión logística ordinal. El modelo estructural tiene la siguiente forma general:

$$P [hogar_i \in decil_j] = f (Z_{puntaje}, Z_{valor\ unitario\ integral}, clase\ predio_{PH-NPH})$$

A partir de los parámetros estimados, es posible establecer las ponderaciones de cada variable que expresan la magnitud de su asociación con la clasificación en deciles de los indicadores. Estas ponderaciones pueden verse en la tabla 4.5.

Tabla 4.5. Peso de las variables catastrales y de acceso en la discriminación de las variables socioeconómicas

Variable	ICPp	ICPe	ICVU	ICV	IBS
Puntaje	0,3	0,87	0,62	0,77	0,82
Valor unitario integral	0,23	0,51	0,51	0,09	0,9
Valor metro cuadrado de terreno	0,28	0,66	0,07	0,39	0,08
Valor metro cuadrado de construcción	-0,17	0,08	-0,2	-0,39	-0,24
Distancia a hospital de nivel III	-0,06	-0,11	-0,04	-0,09	-0,04
Índice de acceso a bienes públicos	-0,009	0,12	0,72	0,09	0,1
Índice de acceso a bienes privados	0,004	0,03	-0,04	0,05	-0,005
Máxima distancia al trabajo por hogar	-0,05	0,00	0,005	0,001	0,21
Predio en propiedad horizontal	0,19	0,24	0,36	0,93	0,59

* Los valores en color rojo son estadísticamente no significativos (5%).

Fuente: elaboración de los autores a partir de la EMB 2011.

Para los cinco modelos estimados, uno por indicador, las variables catastrales son las que mayor peso tienen en la discriminación por deciles de los indicadores socioeconómicos. Nuevamente sobresalen el valor unitario integral y el puntaje de la construcción, sumándose ahora la condición de la propiedad (PH o NPH) con una ponderación importante en el poder discriminante de los indicadores socioeconómicos. Por su parte, las variables de acceso han mostrado ser útiles, sin embargo, su ponderación es bastante baja en relación con las catastrales.

Conclusiones y recomendaciones

Se ha explorado a lo largo de este capítulo el comportamiento de una batería de indicadores en la ciudad de Bogotá, para posteriormente determinar el nivel de asociación de estos con las variables catastrales. El principal hallazgo de este análisis ha sido poner en evidencia que existe una estrecha relación entre las variables físicas y de entorno, expresadas en la información catastral y de acceso, con un conjunto particular de indicadores socioeconómicos que buscan aproximarse a las condiciones socioeconómicas de los hogares.

Especialmente, se ha destacado, a través de diferentes análisis, el alto grado de asociación del *valor unitario integral* y del *puntaje de la construcción* con las condiciones socioeconómicas de los hogares. El último de los análisis realizados, adicionalmente, ha resaltado el poder discriminante del régimen de propiedad horizontal.

La importancia del valor unitario integral, expresada en las altas correlaciones con los indicadores socioeconómicos, se robustece al explorar sus propiedades intrínsecas, pues esta variable se ha construido buscando condensar los aspectos más relevantes en la determinación de los precios de los inmuebles: el entorno y la construcción. En el puntaje de la construcción, si bien sobresale también por sus altas correlaciones con los indicadores socioeconómicos, su naturaleza es expresar solamente el componente de la construcción y no del entorno. Estos elementos, sumados al hecho de que en la construcción del valor unitario integral se involucra el puntaje de la construcción como un atributo dentro de los modelos econométricos, hacen que se destaque como la variable más completa y con más altas correlaciones con las características socioeconómicas en la ciudad de Bogotá.

Otro importante resultado ha sido la baja capacidad de la estratificación vigente para expresar las fluctuaciones de los indicadores socioeconómicos a lo largo de la escala de sus respectivos deciles; las probabilidades de pertenecer a los estratos 2 y 3 son las más altas en toda la escala de capacidad de pago y de calidad de vida. Esto es una señal de que la estratificación vigente no está diferenciando adecuadamente la población en términos de sus características socioeconómicas. Este aspecto será evaluado con mayor detalle más adelante, al comparar la estratificación vigente con otras alternativas de clasificación.