



**Universidad Autónoma de Baja California  
Facultad de Economía y Relaciones Internacionales**

**Doctorado en Ciencias Económicas**

**"Mercados de vivienda transfronterizos entre México, Estados Unidos y  
Canadá: submercados e información asimétrica"**

**Atenea De La Cruz Brito**

**Director de Tesis: Dr. Emilio Hernández Gómez  
Co-Director: Dr. Paavo Monkkonen**

Tijuana, B.C. 2022

## **DEDICATORIA**

*A mi esposo Alberto Díaz Ramírez,  
por tu inmenso amor.*

## **AGRADECIMIENTOS**

La presente investigación fue realizada con el apoyo recibido por parte del Consejo Nacional de Ciencia y Tecnología, CONACyT, así como con el soporte académico de la Facultad de Economía y Relaciones Internacionales FEYRI, de la Universidad Autónoma de Baja California, UABC.

Agradezco a mi director de tesis, el Dr. Emilio Hernández Gómez, por su apoyo y paciencia, aprecio mucho su tiempo, dedicación y amabilidad, tanto durante la etapa lectiva del doctorado, como en el proceso de planteamiento y realización de la investigación. Gracias por siempre haber tenido la puerta de su subículo abierta, aprecio enormemente sus comentarios, reflexiones y sugerencias tanto para mi tesis como para mi proyecto de carrera. En este mismo tenor, agradezco a mi co-director el Dr. Paavo Monkkonen por su generoso apoyo y confianza, gracias por su guía en la elaboración del diseño de la investigación, el cual fue central en el planteamiento general de este estudio. Aprecio mucho su tiempo y atención durante todo el proceso de mi doctorado así como durante mi tesis de maestría, su motivación y aliento ha sido de gran ayuda en el proceso de la investigación. Quiero agradecer a la Dra. Erika Chávez Nungaray por su apoyo desde el primer día y hasta el día de hoy tanto en todo lo relacionado con el doctorado y la tesis así como en el ámbito académico y profesional. Agradezco de igual forma al Dr. Noé Arón Fuentes Flores, integrante fundamental de mi comité tutorial quien me brindó sus comentarios y observaciones, los cuales fueron fundamentales para la conclusión de esta investigación. Aprecio y reconozco de manera especial su tiempo y apoyo, tanto en esta tesis como en mi tesis de maestría. Agradezco al Dr. Julio Ernesto Torres Coto Maizer por su interés en este proyecto de investigación y por haber aceptado formar parte del comité tutoría.

Igualmente, reconozco a la Facultad de Economía y Relaciones Internacionales, tanto al cuerpo académico del Doctorado en Ciencias Económicas como al personal administrativo por el apoyo recibido. Agradezco a la Dra. Ana Isabel Acosta Martínez, Coordinadora del Programa de Doctorado en Ciencias Económicas, por sus atenciones durante el periodo de mis estudios. Asimismo, agradezco la orientación, amabilidad, atenciones y apoyo de Raquel Tirado Hernández, asistente de Posgrado de la FEYRI.

Mi agradecimiento más especial es para mi esposo Alberto Díaz Ramírez, quien siempre me ha impulsado a lograr mis metas y me ha alentado a dar lo mejor de mi, brindadome todo su apoyo y atención, que ha sido fundamental durante toda mi formación de posgrado así como para cumplir con esta tesis de investigación. Gracias mi vida por tu paciencia y compromiso, así como por tu ejemplo de tenacidad y excelencia que te caracterizan. Agradezco a mis padres Jesús De La Cruz Patiño y Audrey Brito Ramos, por su motivación y seguimiento atento al avance de esta tesis.

Por último, agradezco a mis compañeros de generación del Doctorado en Ciencias Económicas, así como a los compañeros de otras generaciones de los programas de maestría y doctorado con quienes tuve la oportunidad de coincidir, agradezco sus pláticas, comentarios y cortesías. A todos ellos, muchas gracias.

# ÍNDICE

<b>INTRODUCCIÓN</b> .....	1
<b>I. MARCO CONTEXTUAL DE LA REGIÓN FRONTERIZA</b> .....	10
1.1. El espacio geográfico y su ámbito .....	10
1.2. Indicadores de los mercados de vivienda .....	20
<b>II. MARCO TEÓRICO</b> .....	29
2.1. Estado de la cuestión .....	30
2.1.1. <i>Economía urbana</i> .....	30
2.1.2. <i>Precios hedónicos</i> .....	32
2.1.3. <i>Mercados multijurisdiccionales</i> .....	34
2.2. Revisión crítica de la bibliografía.....	36
2.2.1. <i>Submercados</i> .....	36
2.2.2. <i>Información asimétrica</i> .....	41
<b>III. METODOLOGÍA</b> .....	46
3.1. Diseño de investigación general .....	46
3.1.1. <i>Fuentes de datos</i> .....	46
3.1.2. <i>Modelo base</i> .....	48
3.2. Metodología y datos por eje de investigación .....	50
3.2.1. <i>Submercados</i> .....	50
3.2.2. <i>Información asimétrica</i> .....	56
<b>IV. MODELO DE PRECIOS HEDÓNICOS</b> .....	58
4.1. Preprocesamiento y análisis exploratorio .....	58
4.1.1. <i>Preparación de datos: recopilación, limpieza y transformación</i> .....	58
4.1.2. <i>Análisis exploratorio de datos (EDA)</i> .....	60
4.1.3. <i>Análisis exploratorio de datos espaciales (ESDA)</i> .....	67
4.4. Modelo de precios hedónicos .....	68

4.4.1. Modelo base .....	68
<b>V. IDENTIFICACIÓN DE SUBMERCADOS</b> .....	75
5.1. Submercados por características del vecindario .....	76
5.1.1. Análisis de Componentes Principales (PCA).....	76
5.1.2 Descripción de los indicadores del entorno urbano .....	77
5.1.3 Resultados del análisis de componentes principales .....	80
5.2. Submercados por características de la vivienda .....	92
5.2.1. Análisis de Conglomerados (K-Means).....	92
5.2.2. Análisis de Conglomerados Jerárquico (HC).....	95
<b>VI. DISCUSIÓN Y CONCLUSIONES</b> .....	98
<b>VI. BIBLIOGRAFÍA</b> .....	103
<b>VII. ANEXOS</b> .....	111

## ÍNDICE TABLAS, GRÁFICAS Y MAPAS

Tabla 1. Claves de aglomeraciones del área de estudio .....	13
Tabla 2. Distribución de la tenencia de la vivienda, 2019 (%).....	20
Tabla 3. Condiciones de la vivienda, 2019 (%) .....	20
Tabla 4. Ingreso mediano disponible (precios corrientes, USD).....	22
Tabla 5. Índice de Gini.....	22
Tabla 6 Definición operacional de las variables e hipótesis de submercados .....	41
Tabla 7 Definición operacional de las variables e hipótesis de información asimétrica .....	45
Tabla 8 Descripción de variables .....	59
Tabla 9 Estadística descriptiva, ciudades fronterizas México-Estados Unidos .....	61
Tabla 10 Estadística descriptiva, ciudades fronterizas Estados Unidos-Canadá.....	63
Tabla 11 Resumen de resultados modelos OLS, semilog y semilog con ruido blanco HC para las ciudades fronterizas México-Estados Unidos .....	69
Tabla 12 Resumen de resultados modelos OLS, semilog y semilog con ruido blanco HC para las ciudades fronterizas Estados Unidos-Canadá.....	72
Tabla 13 Clasificación vecindarios, combinación de los dos primeros factores principales resultantes del PCA.....	77
Tabla 14 Variables del conjunto de datos IDU2021 .....	78
Tabla 15 Variables de los conjuntos de datos LAI17, NWI21, PK22 y PLS20.....	79
Tabla 15 Variables del conjunto de datos PMD2020.....	80
Tabla 17 Resumen FAMD: número de factores y varianza acumulada.....	81
Tabla 18 Contribución de las variables originales en los primeros dos componentes principales por ciudad, FAMD.....	81
Tabla 19 Resumen PCA: número de factores y varianza acumulada.....	83
Tabla 20 Contribución de las variables originales en los primeros dos componentes principales por ciudad, PCA .....	84
Tabla 21 Resumen PCA: número de factores y varianza acumulada.....	85
Tabla 22 Contribución de las variables originales en los primeros dos componentes principales por ciudad, PCA .....	85
Tabla 23 Resumen PCA: número de factores y varianza acumulada.....	89
Tabla 24 Contribución de las variables originales en los primeros dos componentes principales por ciudad, PCA .....	90

Gráfica 1. Cruce anual de personas a Estados Unidos provenientes de México y Canadá por puertos de entrada seleccionados, 1996-2021.....	14
Gráfica 2. Índice de Precios de Vivienda SHF, trimestral 2005-2020 (Variación %) .....	23
Gráfica 3. Índice de Precios de Vivienda FRED, trimestral 2000-2020 (Variación %).....	24
Gráfica 4. Índice de Precios de Vivienda FRED, anual 2000-2020 (Variación %) .....	25
Gráfica 5. Índice de Precios de Vivienda Teranet-National Bank House, anual 2000-2020 (Variación %).....	25
Gráfica 6. Cruce mensual de personas a Estados Unidos provenientes de México y Canadá por puertos de entrada seleccionados, 2018-2021. ....	27
Gráfica 7. Dispersión de precios en dólares y metros cuadrados de las viviendas de la frontera México-Estados Unidos, ciudades seleccionadas .....	65
Gráfica 8. Dispersión de precios en dólares y metros cuadrados de las viviendas de la frontera Estados Unidos-Canadá, ciudades seleccionadas.....	66
Gráfica 8. Índice de Moran para las ciudades fronterizas entre México, Estados Unidos y Canadá, ciudades seleccionadas.....	67
Gráfica 10. Observaciones agrupadas por variables cualitativas del entorno urbano, FAMD México .....	82
Gráfica 11. Círculo de correlación, contribución de las variables en las dimensiones 1 y 2 del entorno urbano, PCA Estados Unidos.....	87
Gráfica 8. Cuadrado del eigenvalor de cada observación del entorno urbano ( $\cos^2$ ), PCA Canadá .	91
Gráfica 13. Clústeres resultantes, clasificación por K-Medias, frontera México-Estados Unidos....	93
Gráfica 14. Clústeres resultantes, clasificación por HC, Estados Unidos .....	96
Gráfica 15. Clasificación de ciudades fronterizas entre México, Estados Unidos y Canadá.....	101
Gráfica 16. Población de las conurbaciones fronterizas por país (%).....	102
Mapa 1.Regiones ecológicas entre México, Estados Unidos y Canadá. ....	11
Mapa 2. Conurbaciones y cruces fronterizos seleccionados entre México, Estados Unidos y Canadá. ....	12
Mapa 3. Detalle de ciudades entre México y Estados Unidos, noroeste. ....	15
Mapa 4. Detalle de ciudades entre México y Estados Unidos, noreste. ....	16
Mapa 5. Detalle de ciudades entre Estados Unidos y Canadá, noroeste. ....	18
Mapa 6. Detalle de ciudades entre Estados Unidos y Canadá, noreste. ....	19

## INTRODUCCIÓN

---

El espacio fronterizo entre México, Estados Unidos y Canadá, se caracteriza por una contigüidad geográfica e interacción social que históricamente ha dado pie al establecimiento de comunidades en sus límites territoriales. Este contexto de cercanía e intercambio permite a los residentes de la región binacional acceder a bienes y servicios en el país vecino, entre ellos la vivienda. De este modo, existen en el área mercados inmobiliarios residenciales que se extienden a través de las fronteras, por tanto tienen la cualidad de ser transfronterizos.

En la presente investigación se estudian los mercados de bienes raíces residenciales en las fronteras de Norteamérica a partir de dos ejes temáticos: los submercados de vivienda y la información asimétrica. En primer lugar, la vivienda es un bien inmueble, durable y heterogéneo, esta última dimensión se refiere al hecho de que las propiedades presentan características intrínsecas y extrínsecas distintas, a partir de las cuales se segmenta la oferta en submercados. En la práctica, los submercados de vivienda son producto de la coexistencia de conjuntos diferenciados de residencias con características similares, así como zonas o vecindarios con atributos y amenidades particulares que son demandadas por diversos posibles compradores.

La identificación de submercados de vivienda y su caracterización permite dar cuenta de los segmentos en los cuales se divide el mercado así como de su composición y nivel de precios con el fin de capturar de forma más específica el comportamiento de cada mercado. Para ello, en este estudio se delimitan los submercados a partir de tres ejercicios criterios: por condiciones extrínsecas de los vecindarios, por atributos intrínsecos de las casas y por regímenes espaciales.

En segundo lugar, el mercado de vivienda se caracteriza por exhibir problemas de información o fallos de mercado, en particular asimetrías de información entre los agentes, sean estos vendedores, compradores o profesionales inmobiliarios. Esto ocurre cuando dichos participantes poseen diferentes niveles de conocimiento sobre el mercado. Los mercados de vivienda con información asimétrica tienden a presentar una mayor dispersión de precios en bienes similares, así como costos más altos en el proceso de búsqueda. La

irrupción de sitios web especializados y aplicaciones digitales para la oferta de bienes raíces, ha permitido que los agentes puedan obtener información a menor costo.

El grado de adopción y desarrollo de dichas plataformas digitales en cada país y ciudad del área de estudio varía. Con el fin de evaluar las diferencias entre los niveles de información de cada mercado de vivienda local y compararlo con su contraparte al otro lado de la frontera, se elabora un índice de información a partir de la evaluación de la calidad y cantidad de datos disponibles en portales de internet. Asimismo, se estudia la relación entre este índice y la dispersión de precios.

La motivación social de esta investigación radica en el hecho de que la vivienda representa uno de los mayores activos de los hogares y es el componente más importante en la riqueza de las familias, por lo cual tiene un papel relevante en la reducción de la desigualdad (Soaita, Maclaren, & Gibb, 2020). Al mismo tiempo, la vivienda constituye una necesidad básica reconocida por leyes nacionales e internacionales como un derecho<sup>1</sup>, ya que brinda cobijo, protección, seguridad, descanso, sanidad, privacidad y comodidad a las personas. Además, proporciona un espacio para el libre desarrollo de la personalidad en sus dimensiones biológica, psicológica y sociocultural, así como para el establecimiento de las relaciones sociales.

En la práctica, la satisfacción de esta necesidad básica dependerá del acceso y las condiciones de habitación de las personas. En este tenor, es relevante el concepto de asequibilidad a la vivienda, la cual indica la capacidad de los hogares para costear la adquisición de un bien inmueble a precio de mercado. En las comunidades fronterizas de Norteamérica, la demanda residencial de algunos hogares se satisface en el otro lado de la frontera de esta región limítrofe. De esta manera, la elección de vivienda está sujeta a las preferencias de los consumidores, su restricción presupuestaria, la oferta conjunta de inmuebles en las ciudades colindantes y las barreras o costos de transacción asociados al acceso al mercado de vivienda contiguo.

---

<sup>1</sup> El derecho a la vivienda que asegure un nivel de vida adecuado se encuentra plasmado en el artículo 25.1 de la Declaración Universal de los Derechos Humanos (ONU, 1948), así como por el Pacto Internacional de Derechos Económicos, Sociales y Culturales de 1966, firmado por los tres países y ratificado por México y Canadá. En cuanto a las legislaciones de cada país, el artículo 4° de la Constitución Política de los Estados Unidos Mexicanos (1917) establece el derecho a una vivienda digna y decorosa como garantía individual. Recientemente Canadá aprobó el reconocimiento a este derecho a través del *National Housing Strategy Act* (2019), mientras que en los Estados Unidos está ausente en la legislación federal.

Es así como la demanda de vivienda en un lado de la frontera está en función de la demanda local aunada a la demanda de su contraparte fronteriza o foránea. Por consiguiente, la valoración de precios de un mercado no está determinada únicamente por el nivel de ingreso y demanda local, sino por la influencia de la demanda foránea, en función de su magnitud, dirección y peso relativo.

De este modo, el impacto de la demanda foránea se traslada a los precios de la ciudad contigua y con ello al nivel de asequibilidad, que a su vez afecta en el bienestar de las personas. El efecto de este fenómeno puede ocasionar que la vivienda sea aún más inasequible en mercados que ya sufrían un aumento significativo de precios. Por lo anterior, las condiciones del mercado de vivienda y el efecto de la demanda proveniente del otro lado de la frontera requieren de especial atención.

Sin embargo, el estudio de la demanda foránea y sus efectos en la oferta, demanda y precios de los mercados de vivienda en las fronteras es exiguo. Tradicionalmente, el estudio del consumo a través de las fronteras se ha enfocado en dos tipos de productos: los bienes muebles, es decir, aquellos que pueden trasladarse; y los servicios, aquellos que se reciben en el punto donde se producen (Bhagwati, 1984; Kravis & Lipsey, 1983). Ejemplo de ellos son los productos exentos de arancel o *duty free* y los servicios profesionales, respectivamente. Por tanto, existe una brecha en la bibliografía sobre los mercados de vivienda en las fronteras que se solventa con esta investigación.

El área de estudio entre México, Estados Unidos y Canadá, constituye un caso particular de mercados de vivienda bajo diferentes jurisdicciones político-administrativas, niveles de desarrollo económico e ingreso. Por un lado, el marco legal de operación de los mercados de vivienda es diferente en cada país, además de los planes y ordenamientos locales de cada ciudad o zona metropolitana. Es por ello que cada uno de los pares de ciudades fronterizas constituye un caso de mercado de vivienda multijurisdiccional. Por otro lado, la disparidad en niveles de desarrollo e ingreso de una ciudad frente a su contraparte del otro país, representan un abanico de posibilidades para la comparación de mercados de vivienda bajo diferentes condiciones de desarrollo económico.

De este modo, en virtud de la similitud de los pares de ciudades fronterizas y las diferencias nacionales entre los tres países, se responde a preguntas generalizadas sobre la estructura de los mercados de vivienda, sus características intrínsecas y extrínsecas así

como la valoración de su precio de no mercado desde un planteamiento comparado, ya que las ciudades fronterizas ofrecen una ventana a los procesos que se experimentan en otras ciudades.

Igualmente, existe un espacio en la línea de investigación de los submercados de vivienda para la aplicación y desarrollo de su identificación, a partir de un diseño metodológico que conjuga los tres tipos de submercados: geográficos o con base topográfica, por características cualitativas y submercados anidados. En primer lugar, los submercados geográficos son aquellos que diferencian los vecindarios en función de sus condiciones topográficas o propias del terreno, en conjunto con el espacio construido a través de amenidades presentes en el entorno urbano circundante.

En segundo lugar, los submercados por características intrínsecas de las viviendas agrupan a las residencias en función de la cantidad y calidad de sus atributos, entre los cuales destacan los metros cuadrados de terreno y construcción, número de habitaciones, baños, espacios de estacionamiento, calidad de acabados, espacios comunes, entre otros. En tercer lugar, los submercados anidados son aquellos resultantes de la incorporación de los dos submercados anteriores. Por tanto, a partir del estudio de los submercados de vivienda transfronterizos, se contribuye en la evaluación de esta metodología mixta mediante resultados robustos que brinden evidencia empírica y abonen en el debate académico.

La información es un recurso de vital importancia para oferentes y demandantes, sin embargo, en el mercado de vivienda este activo suele ser escaso y costoso, motivo por el cual la aplicación de este marco teórico es relevante en la práctica. Por tanto, con la incorporación del análisis de las asimetrías de información en el área de estudio, esta investigación aporta en evidencias a partir de un examen comparativo entre mercados bajo diferentes jurisdicciones.

En razón de lo anteriormente expuesto, el objetivo de esta investigación es analizar la segmentación de submercados de vivienda en un contexto fronterizo entre México, Estados Unidos y Canadá. Asimismo, medir los niveles de información de los mercados residenciales a partir del nivel de precios y su dispersión en cada ciudad. Con ello, se establece el estudio de los mercados transfronterizos como un caso útil para el análisis de los mercados de vivienda sujetos a múltiples jurisdicciones y niveles de desarrollo

económico, de forma que la metodología empleada sea replicable y contribuya a mejorar la comprensión de este tipo de mercados de bienes residenciales.

Para ello, se caracterizan los mercados de las conurbaciones limítrofes a partir de un estudio de sección cruzada con la metodología de precios hedónicos, enriquecido con la identificación de submercados para cada ciudad mediante el análisis de componentes principales, análisis de conglomerados, y regímenes espaciales, como la expresión de la segmentación por características del vecindario, de la vivienda, y conjunta, respectivamente. El estudio de los submercados de vivienda constituye el eje central de esta investigación.

Adicionalmente, se examina el efecto de la disponibilidad de información sobre el mercado de vivienda a través de páginas de internet especializadas, en la asimetría de la información de los agentes en mercados con diferentes niveles de información. A partir de la evaluación de la calidad y cantidad de información con un índice, y su relación con la dispersión de los precios en bienes inmobiliarios residenciales para cada mercado local. Esto a su vez permite estudiar el efecto de la irrupción de las plataformas digitales en el mercado de vivienda convencional.

Las fuentes de información utilizadas se catalogan en dos tipos, aquellos conjuntos de datos generados por las agencias estadísticas de cada país y los que se obtienen de la oferta publicada en páginas web y aplicaciones digitales. Los primeros incluyen para México a los *Censos de Población y Vivienda* 2000, 2010 y 2020, la *Encuesta Nacional de Ingreso-Gasto de los Hogares* (ENIGH) de 2000 a 2018 y las Características del Entorno Urbano del *Inventario Nacional de Vivienda* (INV) 2016, todos ellos elaborados por el Instituto Nacional de Estadística y Geografía (INEGI). De igual modo, datos producidos por agencias internacionales o centros de investigación como el *World Resources Institute* (WRI) y su Índice de Desigualdad Urbana 2021.

Para Estados Unidos los *US Census* de 2000, 2010 y 2020, y la *American Community Survey* (ACS) de 2000 a 2019, elaborados por el organismo estadístico federal *United States Census Bureau* (US Census). Así como fuentes de agencias gubernamentales estadounidenses como los registros de cruces fronterizos por puertos de entrada del *Bureau of Transportation Statistics* (BTS), el *Location Affordability Index* 2017 del *Department of Housing and Urban Development* (HUD), el *Public Libraries Survey* 2020 del *Institute of*

*Museum and Library Services (IMLS)*, el *National Walkability Index 2019* del *U.S. Environmental Protection Agency (EPA)*, y el conjunto de datos *U.S. ParkServe 2022* de la organización no gubernamental *Trust for Public Land (TPL)*.

En el caso de Canadá, se utilizaron los *Census of Population* de los años 2001, 2006, 2011, 2016 y 2021, y las bases del entorno de datos abiertos *Open Database Cultural and Art Facilities (ODCAF)* del año 2020, *Open Database Educational Facilities (ODEF)* de 2021, *Open Database Healthcare Facilities (ODHF)* de 2020, y el *Open Database Recreational and Sport Facilities (ODRSF)* de 2021. Todos ellos elaborados por la agencia estadística nacional *Statistics Canada (StatCan)*.

El segundo tipo de información comprende fuentes creadas expresamente para esta investigación. En primer lugar, la oferta de vivienda para cada mercado local presente en páginas web especializadas de franquicias de agentes inmobiliarios y plataformas comerciales virtuales de intermediarios de bienes raíces *–tech real-estate marketplace–*, tanto de compañías con presencia internacional como agencias nacionales o locales. Dichos datos fueron obtenidos por medio de técnicas de extracción de datos web y comprenden a las viviendas ofertadas, sus características intrínsecas y valoración del precio o precios de lista. En segundo lugar, la cantidad y calidad de la información disponible en páginas de internet especializadas en el mercado inmobiliario, así como su posición en índices de motores de búsqueda, alcance y flujo de visitas, recopilados mediante análisis de contenido y herramientas de optimización de posicionamiento en navegadores web *–SEO–*.

Como se expuesto anteriormente, los mercados de vivienda de las regiones limítrofes de Norteamérica operan en un contexto transfronterizo, ante el cual la presente investigación se plantea responder a la interrogante ¿cómo es la estructura de los mercados de vivienda de las ciudades fronterizas entre México, Estados Unidos y Canadá? A su vez, considerando como marco referencial la dimensión heterogénea de los bienes raíces residenciales, aunado a los problemas de información que subsisten en esta industria, surgen preguntas específicas, las cuales se articulan alrededor de cada uno de los dos ejes temáticos de este estudio: los submercados de vivienda y la información asimétrica.

La hipótesis central esgrime el hecho de que los mercados de vivienda de las ciudades en los límites territoriales de América del Norte, están inmersas en un contexto en el cual los agentes económicos pueden optar por consumir vivienda de uno u otro lado de la

frontera. Dicha contigüidad e interacción espacial tiene efectos bidireccionales de diferente magnitud en la segmentación de los submercados y las asimetrías de información, las cuales impactan la oferta, demanda y la valoración de precios de los bienes residenciales de la región. Por tanto, el objetivo general del presente estudio es explicar las características y segmentación de la oferta y demanda así como en el nivel y dispersión de los precios de los mercados de vivienda bajo un contexto de interacción espacial de las conurbaciones fronterizas entre México, Estados Unidos y Canadá.

A continuación se presentan las interrogantes y posibles respuestas para cada uno de los dos ejes temáticos que componen esta investigación. Dentro de la primera línea de estudio sobre los submercados, se indagan ¿cuáles son y qué características presentan los submercados de vivienda en la región fronteriza entre México, Estados Unidos y Canadá?, ¿cómo se relacionan dichos submercados con su contexto multijurisdiccional y de interacción espacial? Para responder a ello, se identifican los submercados de vivienda en las conurbaciones fronterizas del área de estudio, así como sus características comunes y disímiles. Se estima el efecto en magnitud y dirección de las características extrínsecas e intrínsecas de la vivienda en la segmentación de los submercados de vivienda fronterizos. Por último, se contrasta la capacidad explicativa del modelo de precios de vivienda con submercados frente a un modelo sin dicha identificación.

La respuesta hipotética a esta pregunta, es que el contexto fronterizo de la zona colindante entre México, Estados Unidos y Canadá, tiene efectos en la composición de los submercados de vivienda. De este modo, la identificación de los submercados por el entorno del vecindario, por características de la vivienda y la conjugación de los dos anteriores en submercado anidados, permite caracterizar de manera más nítida a los mercados de vivienda. Por consiguiente, la incorporación de los submercados en la especificación del modelo de precios, presenta un mayor poder explicativo de su comportamiento. Al mismo tiempo, esta segmentación permite señalar y comparar mercados de bienes inmobiliarios residenciales contiguos pero pertenecientes a jurisdicciones distintas.

A partir del segundo eje temático surge la interrogante sobre ¿cuál es el efecto de la disponibilidad de información en internet sobre los mercados de vivienda de la zona fronteriza entre México, Estados Unidos y Canadá en la asimetría de la información de los

agentes? La estrategia para inquirir su respuesta se basa en la evaluación del nivel de acceso, la cantidad y calidad de la información web sobre los submercados de vivienda en las conurbaciones fronterizas de Norteamérica. De igual modo, en la estimación del efecto de la disponibilidad de fuentes digitales en la asimetría de la información de los mercados de vivienda del área de estudio. Así como la comprobación del efecto de los recursos informativos en línea especializados en el mercado residencial, en la dispersión de los precios de vivienda de la zona colindante entre los tres países.

Como posible respuesta se plantea que la publicación de nuevas fuentes especializadas en el mercado de vivienda, disponibles para consultar en internet, producen una reducción en la dispersión de los precios para propiedades con características similares en las áreas geográficas con mayor y mejor cobertura en las páginas web de bienes raíces. Es decir, que las ciudades con plataformas de mercados virtuales más desarrollados cuentan con una menor dispersión de precios, y con ello, menores costos de información e ignorancia en el mercado. La irrupción de las empresas de tecnologías de la información en el sector inmobiliario representa un factor a considerar en el estudio de los mercados de vivienda, particularmente, en espacios sometidos a jurisdicciones múltiples con diferentes niveles de información.

En suma, la vivienda es fundamental para el bienestar de las personas, por tanto el estudio de las condiciones del mercado residencial es de gran relevancia no sólo científica o económica, sino social. De este modo, el estudio de los mercados de vivienda en el espacio fronterizo entre México, Estados Unidos y Canadá, permite identificar el efecto de la contigüidad e interacción espacial en la región limítrofe entre estas naciones, sujeta a un entorno multijurisdiccional. Este análisis aporta elementos para discernir el papel los sumercados en el nivel de precios de vivienda y con ello avanzar en la comprensión de las condiciones imperantes en los mercados residenciales a través de fronteras o límites jurisdiccionales. En un marco más amplio, el estudio del consumo de vivienda en espacios fronterizos abona a la comprensión de la dinámica económica en la región.

Las partes subsecuentes de esta tesis comprenden seis capítulos de la investigación, incluyendo los submercados de vivienda y la información asimétrica en el mercado residencial. El primer capítulo esboza una introducción al área de estudio entre las fronteras de México, Estados Unidos y Canadá, en el cual se explican los criterios de selección de las

conurbaciones fronterizas. Además, se describen brevemente las características físicas y sociales de las ciudades de estudio, y se brinda un panorama general de los principales indicadores económicos relacionados con los bienes inmobiliarios residenciales.

El capítulo siguiente muestra el andamiaje teórico en el cual se basa esta investigación, dividido en dos apartados. El primero muestra el estado de la cuestión del campo de la economía urbana, la metodología de precios hedónicos y el estudio de las regiones multijurisdiccionales. El segundo presenta la revisión crítica de la bibliografía para cada uno de los dos ejes temáticos, en donde se identifican los aspectos más relevantes del estudio de los submercados y la información asimétrica en los mercados de vivienda, que no han sido suficientemente abordados por otros estudios.

El tercer capítulo describe el diseño de la investigación a partir de la metodología y datos utilizados. En esta sección se detallan las técnicas estadísticas, econométricas y de econometría espacial que se aplican en el análisis. Asimismo, se enuncian las diferentes fuentes de información consultadas y se describe brevemente el proceso a partir del cual se crearon bases de datos de la oferta de bienes residenciales en cada mercado.

Posteriormente, se describe el preprocesamiento, análisis exploratorio y estimación del modelo base. El quinto capítulo presenta los resultados del estudio sobre submercados de vivienda, por un lado la clasificación por características del vecindario seguido de la clasificación por características intrínsecas de las viviendas. Finalmente, se expone el capítulo de discusión de los resultados a partir del estado actual del conocimiento en la materia y sus implicaciones prácticas, así como las conclusiones generales de esta investigación y recomendaciones. En este mismo capítulo se incluye una clasificación de ciudades fronterizas, como un ejercicio de caracterización de las aglomeraciones del área de estudio en cuatro tipos a partir de su nivel de interacción y tamaño relativo.

## **I. MARCO CONTEXTUAL DE LA REGIÓN FRONTERIZA**

---

El espacio entre las fronteras políticas de México, Estados Unidos y Canadá, alberga comunidades que presentan diferentes niveles de conexión y complementariedad entre sí que van desde los pares de ciudades con poca relación, pasando por las ciudades hermanas hasta la integración funcional en aglomeraciones urbanas. Algunas de estas ciudades han experimentado un rápido crecimiento económico y demográfico en las últimas décadas. Este proceso se ha visto impulsado, entre otros factores, por el fortalecimiento de las relaciones comerciales de integración regional para América del Norte bajo el contexto de la globalización, ejemplo de ello es el TLCAN ahora T-MEC<sup>2</sup>. Consecuentemente, los nexos e intercambios no se restringen sólo a los bienes y servicios, ya que han sido acompañados de un flujo continuo de personas por vía terrestre en la región.

A continuación se brinda una descripción general del área de estudio entre México, Estados Unidos y Canadá, con énfasis en los principales rasgos del territorio que se comparten a través de las fronteras, los cuales han propiciado la conformación de subregiones y permitido el establecimiento de comunidades fronterizas. En primer lugar, se presentan las características del territorio y en segundo lugar, los aspectos socioeconómicos relacionados con el mercado de vivienda.

### **1.1. El espacio geográfico y su ámbito**

Norteamérica comprende a tres bastas naciones independientes que comprenden en conjunto alrededor de 24,323,000 kilómetros cuadrados, con dos límites internacionales entre los diez más largos del mundo<sup>3</sup>. La mayor parte de las fronteras entre estos países discurren por tierra, con algunas secciones delimitadas por ríos: entre México y Estados Unidos por el Río Colorado y el Río Bravo; entre Estados Unidos y Canadá por el Río

---

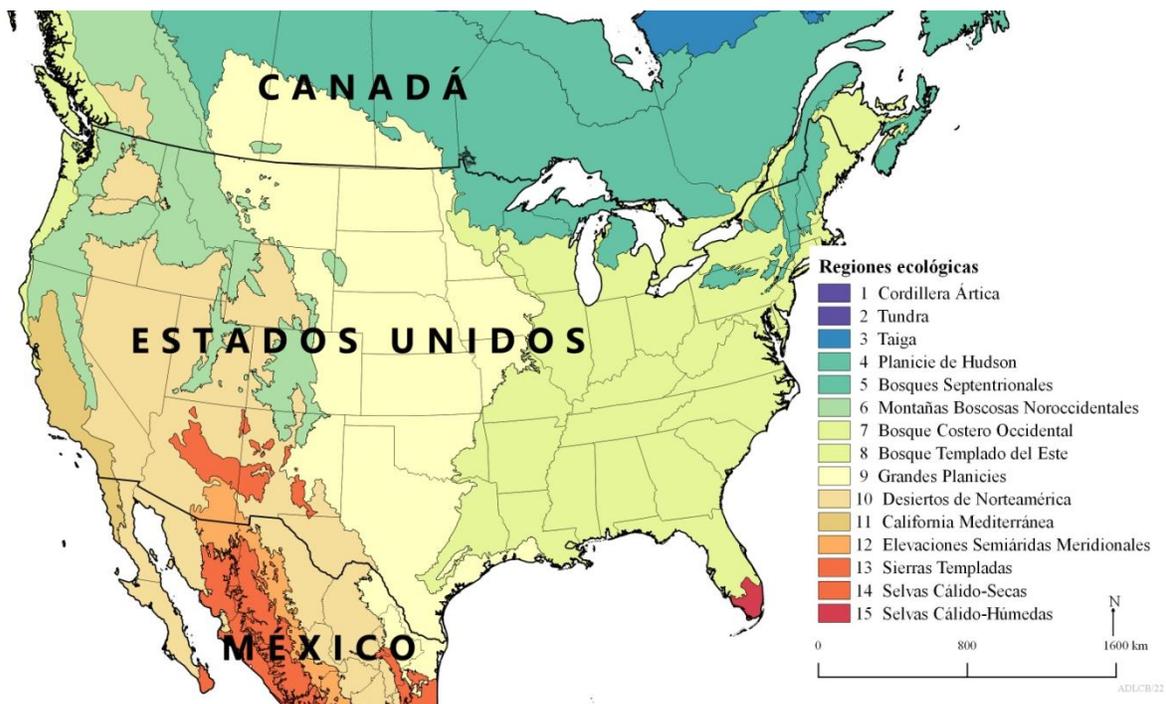
<sup>2</sup> El Tratado de Libre Comercio de América del Norte entre México, Estados Unidos y Canadá conocido como TLCAN o NAFTA por sus siglas en inglés, entró en vigor el 1° de enero de 1994 y fue actualizado y sustituido por el Tratado entre México, Estados Unidos y Canadá, nombrado T-MEC, USMCA o CUSMA en cada país respectivamente, el cual entró en vigor el 1° de julio de 2020.

<sup>3</sup> La frontera entre Estados Unidos y Canadá es la más larga del mundo con 8,891 km de extensión, mientras que la división entre México y Estados Unidos es de 3,111 km de largo, ocupando la décima posición (U.S. Census Bureau, 2011).

Columbia, el Lago de los Bosques y Río Lluvioso, los Grandes Lagos y Río San Lorenzo, el Río St. Croix, entre otros.

Este subcontinente septentrional está en su mayoría sobre la placa tectónica Norteamericana. En cuanto a sus regiones ecológicas, puede dividirse en 15 grandes áreas por sus ecosistemas, tipo, calidad y cantidad de recursos ambientales (ver mapa 1). A lo largo de la frontera Estados Unidos-Canadá se encuentran de oeste a este el Bosque Costero Occidental, las Montañas Boscosas Noroccidentales, los Desiertos de Norteamérica, las Grandes Planicies, los Bosques Septentrionales y los Bosques Templados del Este (Comisión para la Cooperación Ambiental, 1997). El frío extremo en el norte de Canadá de alrededor de 2°C a -20°C en promedio anual, en conjunto con otras condiciones medioambientales y sociales, contribuye a la concentración de dos tercios de la población canadiense en el sur del país, a menos de 100 kilómetros de la frontera estadounidense.

**Mapa 1. Regiones ecológicas entre México, Estados Unidos y Canadá.**

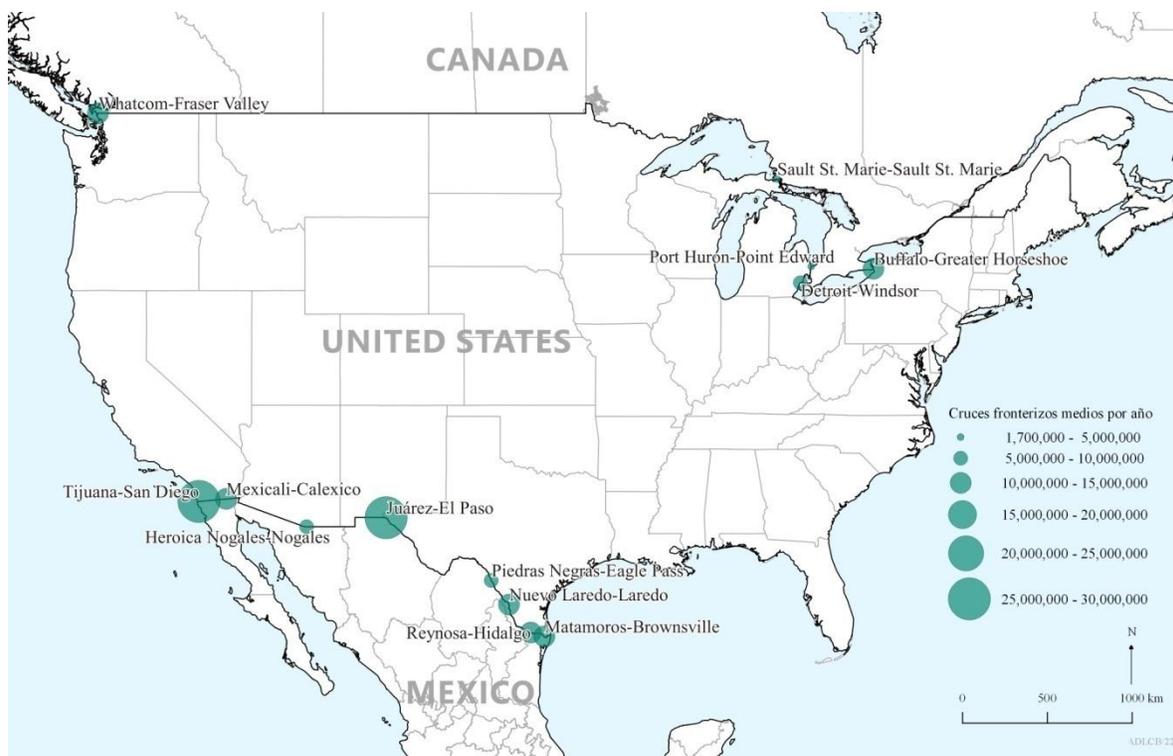


Fuente: Elaborado con datos de la *Comisión para la Cooperación Ambiental (CCA)*, 2006.

Entre México-Estados Unidos se distingue las regiones ecológicas de la California Mediterránea, los Desiertos de Norteamérica, las Elevaciones Semiáridas Meridionales y

las Grandes Planicies (Comisión para la Cooperación Ambiental, 1997). El clima en la frontera entre estos dos países es de templado a cálido y las condiciones más extremas se observan en el área donde colindan el sur de California con Arizona y el norte de Baja California con Sonora, con temperaturas que oscilan de los 10°C en invierno a los 45°C en verano. A pesar de las temperaturas cálidas extremas en ciudades como Mexicali, Baja California y Calexico, California, dichas comunidades prosperaron gracias a la disponibilidad de agua, tras la implementación de numerosas obras de infraestructura. Sin embargo, en la actualidad éstas y otras ciudades presentan estrés hídrico y vulnerabilidad ante los efectos del cambio climático, al igual que gran parte de la región ecológica del Desierto de Sonora y la costa del Golfo entre México y Estados Unidos (Aqueduct, 2021).

**Mapa 2. Conurbaciones y cruces fronterizos seleccionados entre México, Estados Unidos y Canadá.**



Notas: Media de cruces fronterizos de personas por año de 2015 a 2020. Cruces terrestres hacia Estados Unidos vía peatonal, en vehículo personal, autobús o tren. San Diego incluye los puertos de entrada de San Ysidro, Cross Border Xpress, Otay Mesa y Tecate. Calexico suma a los puertos de Calexico y East Calexico.

Fuente: Elaboración y cálculos propios con datos del *Bureau of Transportation Statistics (BTS)*, 2015-2020.

**Tabla 1. Claves de aglomeraciones del área de estudio**

México					Estados Unidos				
N°	Nombre	Abrev.	Municipio	Tipo	Edo.	Abrev.	Condado	Tipo	Edo.
01	Tijuana-San Diego	<b>TIJ</b>	Tijuana*, Tecate, Playas de Rosarito	ZM	02 BC	<b>SD</b>	San Diego	[MSA] San Diego*-Chula Vista-Carlsbad	06 CA
02	Mexicali-Calexico	<b>MXL</b>	Mexicali*	ZM	02 BC	<b>IMP</b>	Imperial	[MSA] El Centro	06 CA
03	Nogales-Nogales	<b>NOG</b>	Nogales	ZM	26 SON	<b>NOG</b>	Santa Cruz	[MicroSA] Tucson*-Nogales	04 AZ
04	Ciudad Juárez-El Paso	<b>CJZ</b>	Juárez	ZM	08 CHI	<b>ELP</b>	El Paso	[MSA] El Paso*-Las Cruces	48 TX
05	Piedras Negras-Eagle Pass	<b>PDN</b>	Nava, Piedras Negras*	ZM	05 COA	<b>EGP</b>	Maverick	[MicroSA] Eagle Pass	48 TX
06	Nuevo Laredo-Laredo	<b>NLD</b>	Nuevo Laredo	ZM	28 TAM	<b>LRD</b>	Webb	[MSA] Laredo	48 TX
07	Reynosa-McAllen	<b>REY</b>	Reynosa*, Río Bravo	ZM	28 TAM	<b>MCA</b>	Hidalgo	[MSA] McAllen*-Edinburg-Mission	48 TX
08	Matamoros-Brownsville	<b>MTM</b>	Matamoros	ZM	28 TAM	<b>BRW</b>	Cameron	[MSA] Brownsville*-Harlingen-Raymondville	48 TX
Canadá					Estados Unidos				
N°	Nombre	Abrev.	CMA/ CA Municipality	Tipo	Edo.	Abrev.	Condado	Tipo	Edo.
09	Metro Vancouver-Blaine	<b>MVR</b>	Vancouver	CMA	59 BC	<b>WHA</b>	Whatcom	[MSA] Bellingham	53 WA
10	Sault Sainte Marie-Sault Ste. Marie	<b>SSM</b>	Sault Sainte Marie	CA	35 ON	<b>SSM</b>	Chippewa	[MicroSA] Sault Ste. Marie	26 MI
11	Windsor-Detroit	<b>WIN</b>	Windsor	CMA	35 ON	<b>DET</b>	Wayne	[MSA] Detroit*-Dearborn-Livonia	26 MI
12	Point Edward-Port Huron	<b>PTE</b>	Sarnia	CA	35 ON	<b>PTH</b>	St. Clair	[MSA] Warren-Troy*-Farmington Hills	26 MI
13	Great Golden Horseshoe-Buffalo Niagara Falls	<b>GGH</b>	St. Catharines-Niagara Hamilton Toronto*	CMA	35 ON	<b>NIA</b>	Niagara, Erie	[MSA] Buffalo*-Cheektowaga	36 NY

Notas: \* Ciudad principal, cuando se trata de una aglomeración o MSA con dos o más ciudades enunciadas.

México: ZM- Zona Metropolitana.

Canadá: CA-Census Agglomeration.

CMA- Census Metropolitan Area.

Estados Unidos:MSA-Metropolitan Statistical Area

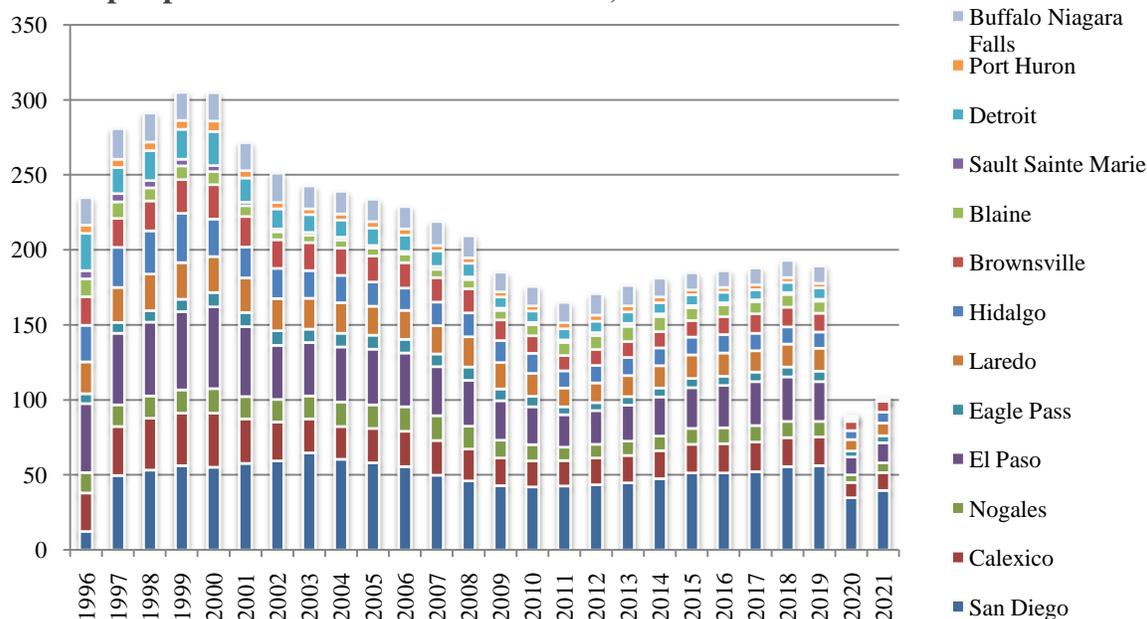
MicroSA-Micropolitan Statistical Area

Fuente: Elaboración propia con información de INEGI, US Census y StatCan.

En concreto, el área de estudio de esta investigación comprende ocho zonas metropolitanas en el norte de México, trece áreas estadísticas micro y metropolitanas en ambas fronteras de Estados Unidos y cinco áreas censales en el sur de Canadá (ver mapa 2 y tabla 1). Estas conurbaciones presentan en primer lugar, similitudes asociadas a su ubicación, como las condiciones geográficas, medioambientales, entre otras. En segundo lugar, muestran diferencias regionales a lo largo de una misma frontera, es el caso de las

diferencias entre la zona del Pacífico frente a la del Atlántico. En tercer lugar, exhiben contrastes al interior de los pares de ciudades, por ejemplo, los niveles de ingreso, las instituciones, entre otros. Esta heterogeneidad enriquece el estudio al permitir un análisis con variaciones regionales, nacionales y locales.

**Gráfica 1. Cruce anual de personas a Estados Unidos provenientes de México y Canadá por puertos de entrada seleccionados, 1996-2021.**



Notas: San Diego incluye los puertos de entrada de San Ysidro, Cross Border Xpress, Otay Mesa y Tecate. Calexico suma a los puertos de Calexico y East Calexico.

Fuente: Elaboración propia con datos del *Bureau of Transportation Statistics* (BTS), 2021.

Las trece conurbaciones del estudio fueron seleccionadas con base en los principios de contigüidad, interacción y peso. Primeramente, las aglomeraciones se encuentran ubicadas sobre la línea internacional o a una proximidad menor a 2.5 kilómetros de la misma. En segundo término, cuentan con el mayor flujo binacional de personas a través de sus puertos fronterizos<sup>4</sup>, concentrando alrededor del 70 por ciento de todos los cruces terrestres por las fronteras norteamericanas (ver gráfica 1). En tercer término, estas ciudades concentran una mayor población frente al resto de comunidades fronterizas de México, Estados Unidos y Canadá. Respecto a la delimitación temporal, el periodo de estudio incluye diferentes cortes de tiempo que van del año 2000 al 2021.

<sup>4</sup> Selección a partir del número de cruces de personas vía peatonal, en vehículo personal, autobús o tren, reportados por el *Bureau of Transportation Statistics* (BTS, 2029).

En la región limítrofe entre México y Estados Unidos, se incluyen ocho municipios y condados, a continuación se presentan en un orden de oeste a este. La zona metropolitana de Tijuana, Baja California y San Diego, California en la costa del Pacífico conforman la aglomeración binacional más grande de América del Norte, donde el condado estadounidense tiene una población mayor que su contraparte mexicana en razón de tres a dos, al contrario de lo que ocurre en el resto de esta frontera binacional. No obstante, la densidad de población en Tijuana es mayor con aproximadamente 900 habitantes por km<sup>2</sup>, dos tercios más que el resto de ciudades de la frontera México-Estados Unidos. Este territorio pertenece a la región ecológica de Salvia Costera Chaparral y Robledal de California.

**Mapa 3. Detalle de ciudades entre México y Estados Unidos, noroeste.**



Fuente: Elaboración propia.

El segundo par de ciudades en la frontera compartida entre Baja California y California es Mexicali-Calexico en el Condado de Imperial. Su región ecológica es el Desierto de Sonora que se extiende a través de California y Arizona en Estados Unidos y Baja California, Sonora y Sinaloa en México. Las principales industrias de estas ciudades son las relacionadas con el comercio internacional, además de la manufactura en Tijuana, la tecnología y los servicios financieros en San Diego, y la agroindustria y los servicios gubernamentales en Mexicali y Calexico. El desempeño económico del Condado de San Diego es de los más altos del país y el mejor en toda la frontera sur de Estados Unidos.

Las ciudades gemelas de Heroica Nogales, en Sonora y Nogales en Arizona, se ubican en el oeste del Archipiélago Madreño, región ecológica dentro del sistema del Piedemonte de la Sierra Madre Occidental, enclave montañoso conocido como las Islas del Cielo. Allí, las actividades económicas están relacionadas con su condición fronteriza e incluyen la logística y transporte, la venta al por menor, la agroindustria y los servicios gubernamentales de migración y control fronterizo.

Por su parte, Ciudad Juárez, Chihuahua y El Paso, Texas, conforman la segunda comunidad más poblada en los límites entre México-Estados Unidos. Esta aglomeración también tiene una vocación relacionada con el comercio internacional, pero más diversificada. En el caso de la ciudad mexicana, con plantas maquiladoras de diferentes tipos y en la ciudad estadounidense, con la sede de grandes compañías. Ambas comunidades se encuentran en la región ecológica del Desierto de Chihuahua, perteneciente al sistema de Desiertos Cálidos.

**Mapa 4. Detalle de ciudades entre México y Estados Unidos, noreste.**



Fuente: Elaboración propia.

Las ciudades de Piedras Negras, Coahuila y Eagle Pass, Texas, tienen una economía dedicada mayoritariamente al sector secundario, con plantas de manufactura relacionadas con metales y autopartes. En cuanto a las ciudades hermanas de Nuevo Laredo, Tamaulipas y Laredo, Texas, representan el principal corredor comercial entre México y Estados Unidos al concentrar más de un tercio de las importaciones y exportaciones binacionales,

las cuales atraviesan de sur a norte la carretera federal 85 y la interestatal 35. Tanto Piedras Negras-Eagle Pass como Nuevo Laredo-Laredo, están divididas por el Río Bravo y forman parte de la región ecológica de Llanuras del sur de Texas, Planicies y Colinas Interiores con Arbustos Xerofitos y Bosques de Roble, dentro del sistema de Planicie Semiárida de Tamaulipas y Texas.

Reynosa, Tamaulipas y McAllen en el Condado de Hidalgo, Texas, se extienden entre las regiones ecológicas al oeste por las Llanuras del sur de Texas, Planicies y Colinas Interiores con Arbustos Xerofitos y Bosques de Roble y por el este la Planicie Costera del Golfo Occidental. La principal actividad económica en Reynosa es la industria manufacturera y en McAllen el comercio internacional y la agricultura, producida en Valle de Río Grande. Finalmente, Matamoros, Tamaulipas y Brownsville en el Condado de Cameron, Texas se encuentran dentro de la región ecológica Planicie Costera del Golfo Occidental, perteneciente a la Planicie Costera de Texas y Louisiana. La economía de estas ciudades se basa en el comercio, la logística, el transporte portuario y terrestre, la manufactura y los servicios.

El área fronteriza entre Estados Unidos y Canadá contiene cinco aglomeraciones, las cuales se describen a continuación de poniente a oriente. En primer lugar, Blaine en el Condado de Whatcom, en el estado de Washington y la Región Metropolitana de Vancouver, también conocida como Metro Vancouver en el estado de British Columbia, pertenecen a la región de Lower Mainland, y se conectan a través de la Carretera del Pacífico. La economía de la ciudad estadounidense se basa en las actividades asociadas al contexto fronterizo y la agroindustria. Por su parte, la metrópolis canadiense es un polo económico de gran relevancia al concentrar diversas industrias que van desde el comercio internacional, el transporte marítimo y terrestre así como los servicios. Además, Metro Vancouver es la tercer aglomeración más importante de Canadá, después de la capital Toronto y Montréal. Ambas comunidades pertenecen a la región ecológica del estrecho de Georgia y Tierras Bajas de Puget, dentro del sistema del Bosque Costero Occidental.

**Mapa 5. Detalle de ciudades entre Estados Unidos y Canadá, noroeste.**



Fuente: Elaboración propia.

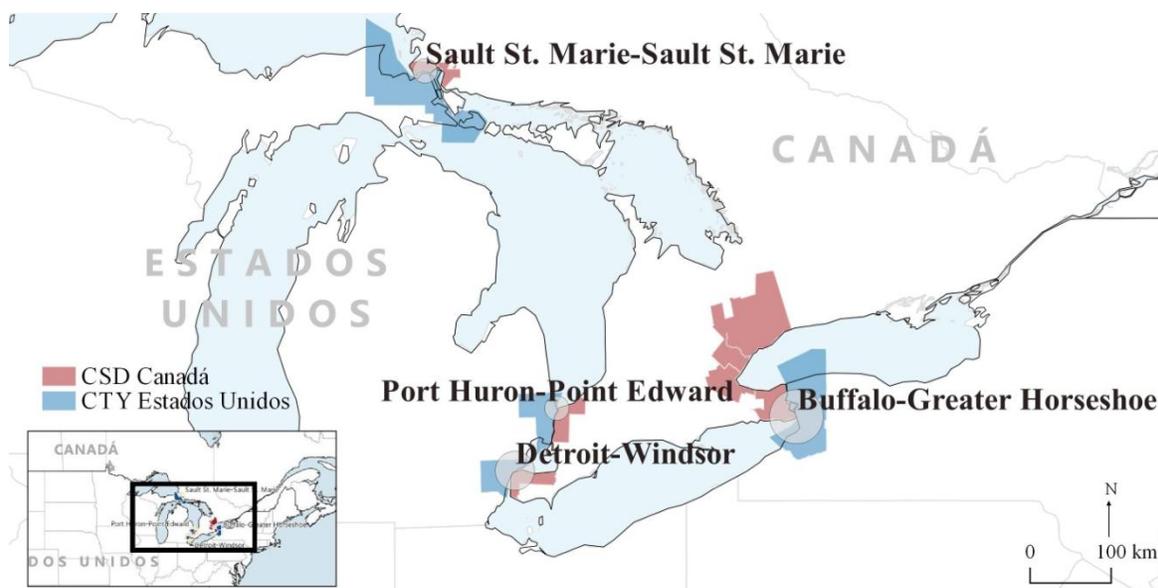
A más de dos mil ochocientos kilómetros al este se encuentran las ciudades gemelas de Sault Ste. Marie en el Condado de Chippewa, Michigan y Sault Sainte Marie, Ontario, divididas por el Río St. Mary. Ambas corresponden al sistema de Bosque Mixto en Escudo, la parte estadounidense está en la región ecológica de los Lagos y Bosques del Norte, en la península superior de Michigan y la parte canadiense en el Algonquin o Laurentides del Sur. La ciudad del lado estadounidense tiene una economía ligada al turismo y los servicios gubernamentales, mientras que su homónima canadiense mantiene las industrias metalúrgica, maderera y de comunicaciones.

En tercer lugar, la ciudad de Detroit en el Condado de Wayne, Michigan y Winsor, Ontario, se ubican al borde del Río Detroit. Estas ciudades han estado ligadas a la industria automotriz, el turismo y los servicios médicos. En Detroit, el cierre de varias fábricas de automóviles tuvo importantes efectos en la ciudad, entre ellos, la pérdida de población, el abandono de viviendas y el deterioro urbano. Por su parte, Port Huron, en el Condado de St. Clair Michigan y Point Edward, en Ontario, son un par de ciudades divididas por el Río St. Clair y el Lago Huron. La ciudad estadounidense cuenta con industrias acerera, papelera y de servicios de salud, mientras que su vecina canadiense brinda servicios turísticos y de entretenimiento.

La ciudad de Detroit se encuentra sobre la región ecológica de las Planicies del Lago Huron-Erie correspondiente a las Planicies Centrales de Estados Unidos. A su vez,

está rodeada por la región de las Llanuras de Deriva del Sur de Michigan-Norte de Indiana, zona en donde se encuentra Port Huron. Mientras que sus contrapartes canadienses, Winsor y Point Edward, se ubican en las Tierras Bajas del Lago Erie correspondientes al sistema de Plainicies con Bosque Mixto.

**Mapa 6. Detalle de ciudades entre Estados Unidos y Canadá, noreste.**



Fuente: Elaboración propia.

Finalmente, el área metropolitana de Buffalo-Niagara Falls, Nueva York y la aglomeración conocida como *Greater Golden Horseshoe* que integra a St. Catharines-Niagara, Ontario, representan el área metropolitana más grande de la frontera Estados Unidos-Canadá, y una de las regiones más pobladas en cada país. La ciudad de Buffalo se encuentra en la región ecológica de las Tierras Bajas del Este de los Grandes Lagos y Hudson, mientras que el área de la Gran Herradura Dorada se ubica en las Tierras Bajas del Lago Erie, ambos dentro del sistema de Planicies con Bosque Mixto. La economía de la región está diversificada, en la ciudad estadounidense se cuenta con industria metalúrgica, manufactura y servicios, a la vez que su contraparte canadiense engloba servicios financieros, industria acerera, automotriz y de servicios de transporte marítimo.

## 1.2. Indicadores de los mercados de vivienda

Respecto a las condiciones del mercado de vivienda en cada país, bajo estas líneas se presentan algunos indicadores económicos a escala nacional así como índices de precios para las ciudades de estudio, con el fin de brindar un panorama general de los mercados de vivienda. Dicha descripción se articula a partir del principio de satisfacción de la necesidad de vivienda, la cual depende de dos aspectos: por un lado, de las condiciones materiales del bien inmueble, y por el otro, del acceso efectivo a una residencia. Bajo la primera dimensión, se considera el estado general de las viviendas en relación a sus ocupantes, y en la segunda el nivel de asequibilidad de la vivienda.

**Tabla 2. Distribución de la tenencia de la vivienda, 2019 (%)**

	Propiedad absoluta	Propiedad con hipoteca	Alquiler a precio de mercado	Otros
OECD	43.1	25	23.1	5.8
Canadá	29.7	39	30.4	0.9
Estados Unidos	24.5	40	34.1	1.7
México	59.4	9	15.0	16.7

Fuente: *Indicators on the Housing Market*, OECD 2019.

**Tabla 3. Condiciones de la vivienda, 2019 (%)**

	Hacinamiento en viviendas de personas con bajos ingresos	Sobrecarga del coste del alquiler para inquilinos con bajos ingresos	Sobrecarga del coste de hipoteca para propietarios con bajos ingresos
OECD	16.0	33.4	23.9
Canadá	1.1	34.3	41.4
Estados Unidos	7.5	45.6	38.8
México	45.4	16.5	24.8

Fuente: *Indicators of Housing Conditions and Affordability*, OECD 2019.

En Norteamérica, la mayoría de la población es dueña de su vivienda (ver tabla 2), ya sea que se encuentre en proceso de pago o esté totalmente liquidada (OECD, 2019). En este indicador de posesión, destaca México, país donde casi dos tercios de los hogares son propietarios. Sin embargo, las condiciones de habitación de las personas no son las óptimas, puesto que casi la mitad de la población con bajos ingresos en México padece hacinamiento (ver tabla 3), además de tener uno de los promedios más bajos de habitaciones por persona con sólo 1, mientras en Estados Unidos el promedio es de 1.9 y

Canadá tiene el promedio más alto dentro de la OECD con 2.6 habitaciones por persona (OECD, 2017).

Respecto a la segunda dimensión relacionada con la asequibilidad, tanto Estados Unidos como Canadá presentan retos importantes, ya que al menos un tercio de su población con bajos ingresos se ve abrumada por los costos de alquiler y las hipotecas (OECD, 2019). A nivel metropolitano, las ciudades de Vancouver, Toronto y San Diego ocupan la segunda, quinta y doceava posición respectivamente, entre los mercados de vivienda más inaccesibles del mundo en el informe anual de Demographia (2021). En dicho documento se calcula la mediana múltiple como el cociente de los ingresos medios con relación al costo mediano del pago de una hipoteca. Es una medida del tipo tasa precio-ingreso y tiene cuatro categorías, en los extremos están los umbrales 3.0 y por debajo como mercado accesible y 5.1 y más para los mercados severamente inasequibles. Las ciudades arriba mencionadas están por arriba de los 8.0 puntos.

En específico, el distrito Metropolitano de Vancouver ha tenido un aumento sostenido en los precios reales del sector inmobiliario desde mediados de la década del 2000, experimentando los puntos más altos entre 2015 y 2016, año en el cual se establecieron controles en el mercado como el impuesto a los compradores extranjeros y a la vivienda deshabitada. Después de la aplicación de estas medidas, los precios se estabilizaron por un breve periodo, pero su tendencia sigue con un fuerte impulso a la alza (UBS, 2019).

Toronto, como parte del *Greater Golden Horseshoe*, también instauró medidas impositivas como las de Vancouver, además de controles de los alquileres y de las normas de las hipotecas, con el fin de apalea la falta de inventario de vivienda suficiente. A pesar de ello, avanzó una posición en el ranking de las ciudades más inasequibles y paso de una mediana múltiple de 8.6 en 2020 a 9.9 en 2021 (Demographia, 2020).

Por su parte, el sur de California ha experimentado aumentos en los precios de la vivienda de más del 50% en las últimas dos décadas, superando el aumento de los ingresos reales. A pesar de que se han otorgado permisos para la construcción, actualmente existe una falta de inventario que se ha visto agudizada por la zonificación estricta aunado a las bajas tasas de interés y las facilidades de crédito (UBS, 2019).

Pero no todos los mercados tienen una tendencia alcista, el área de estudio también comprende a ciudades evaluadas como accesibles o moderadas, es el caso de Buffalo y Detroit, la tercera y doceava más asequibles (Demographia, 2021). No resulta extraño, ya que estas ciudades han experimentado un profundo proceso de pérdida de empleos, situación que entre sus consecuencias provoca el éxodo de un número considerable de sus habitantes y el declive de estas urbes.

En suma, la inasequibilidad de la vivienda puede ser indicador de fenómenos exógenos como la demanda extranjera con fines de inversión, y endógenos como la zonificación estricta y la regulación del mercado de alquiler, entre otros. Por tanto, los mercados de vivienda presentarán diferentes niveles de riesgo ante las perturbaciones producto del aumento o caída de la inversión financiera en bienes inmuebles (UBS, 2019).

A pesar de que las ciudades mexicanas no forman parte de la clasificación internacional de mercados menos asequibles realizada por Demographia o UBS, existen elementos para señalar la importancia creciente de este tema, especialmente si se considera la diferencia de ingresos frente a sus vecinos estadounidenses (ver tabla 4) y los niveles de desigualdad que imperan en el país (ver tabla 5).

**Tabla 4. Ingreso mediano disponible (precios corrientes, USD)**

	2016	2018
Canadá	31171.14	35673.86
Estados Unidos	42635.00	35600.00*
México	2962.23	3464.92

Notas: Ingreso disponible, después de impuestos y transferencias.

\*Valor correspondiente a 2017.

Fuente: *Income Distribution Database*, OECD 2020.

**Tabla 5. Índice de Gini**

	2016	2018
Canadá	0.307	0.303
Estados Unidos	0.391	0.39*
México	0.458	0.418

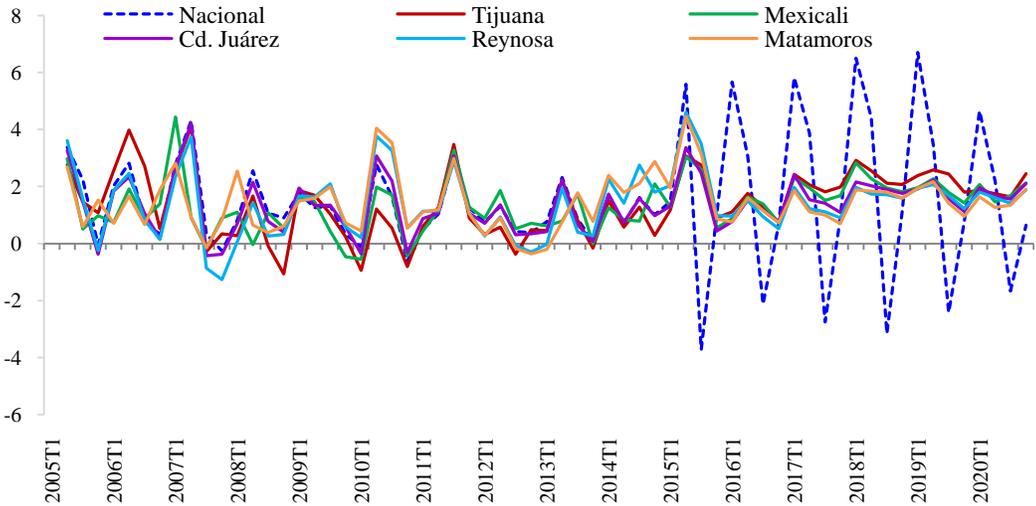
Notas: Ingreso disponible, después de impuestos y transferencias.

\*Cálculo correspondiente a 2017.

Fuente: *Income Distribution Database*, OECD 2020.

En cuanto al comportamiento de los precios a través del tiempo, los índices de precios son útiles para tener una perspectiva general de los cambios más significativos en los mercados y con ello identificar tendencias y movimientos anormales. Este precepto se hizo aún más relevante tras la crisis hipotecaria *subprime* que inició en Estados Unidos en 2007 y se extendió por el sistema financiero mundial hasta finales de 2008 aproximadamente. A continuación se muestran las variaciones de los índices de precios de vivienda para las ciudades fronterizas entre México, Estados Unidos y Canadá.

**Gráfica 2. Índice de Precios de Vivienda SHF, trimestral 2005-2020 (Variación %)**



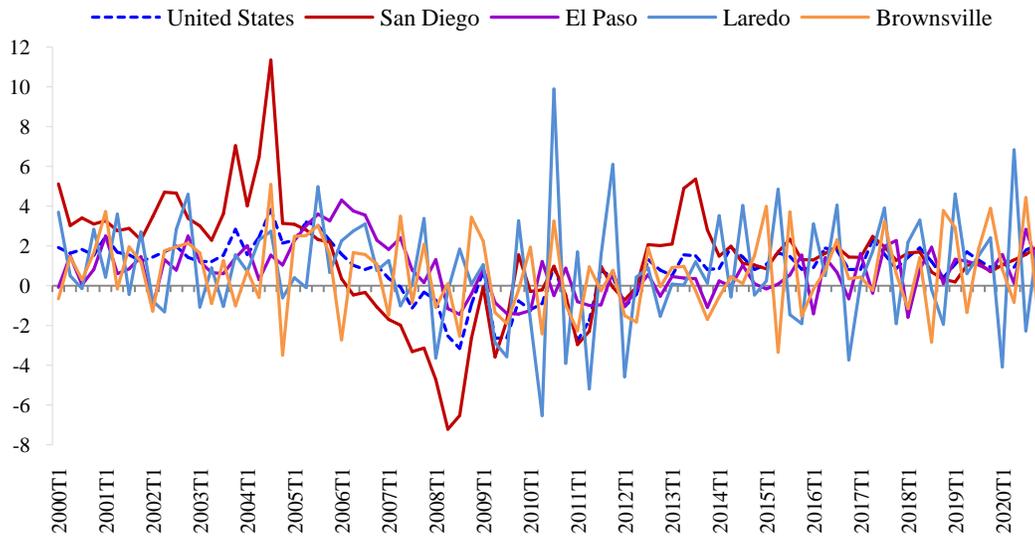
Fuente: Elaboración propia con base en Índice de Precios de Vivienda de la SHF, 2005-2020.

En México la elaboración de estas mediciones es reciente. El índice de precios de vivienda es elaborado por la Sociedad Hipotecaria Federal –SHF–, una dependencia gubernamental que construye el índice con base en los precios de viviendas con crédito hipotecario garantizado, previo avalúo. Su cobertura es nacional, por estados así como por municipios y zonas metropolitanas seleccionadas. Al estar calculado con base en hipotecas, no capta las transacciones realizadas fuera de este sistema, en particular las compras al contado y por medio de otros tipos de crédito.

En Estados Unidos, existen diferentes índices de precios, como referencia está el elaborado por la Agencia Federal Financiera de la Vivienda a través de la Reserva Federal, con series que se remontan a 1975. Incluye a todas las transacciones para la estimación a partir de los precios de venta y datos de avalúos de casas unifamiliares con una

metodología de ventas repetidas. Comprende estados, áreas estadísticas metropolitanas (MSA) y condados.

**Gráfica 3. Índice de Precios de Vivienda FRED, trimestral 2000-2020 (Variación %)**

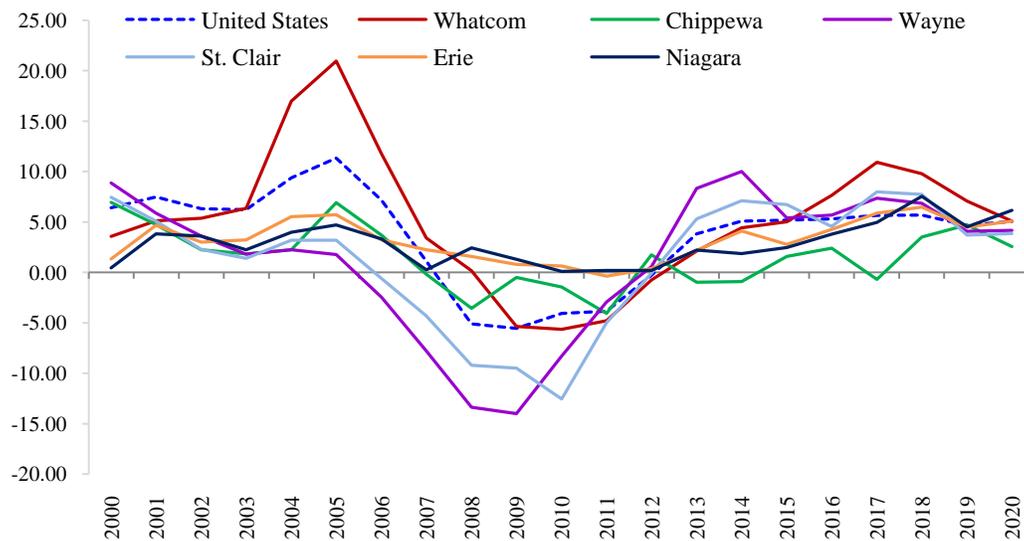


Fuente: Elaboración propia con base en Índice de Precios de Vivienda de la FRED, 2000-2020.

Para las ciudades en la frontera sur de los Estados Unidos, el índice de El Paso se comporta conforme al promedio nacional, mientras que Laredo experimenta la mayor variabilidad y San Diego los mayores aumentos y caídas (ver gráfica 3). Cabe resaltar que El Paso tuvo menores caídas que el resto de MSA durante el periodo posterior a la crisis inmobiliaria *subprime*.

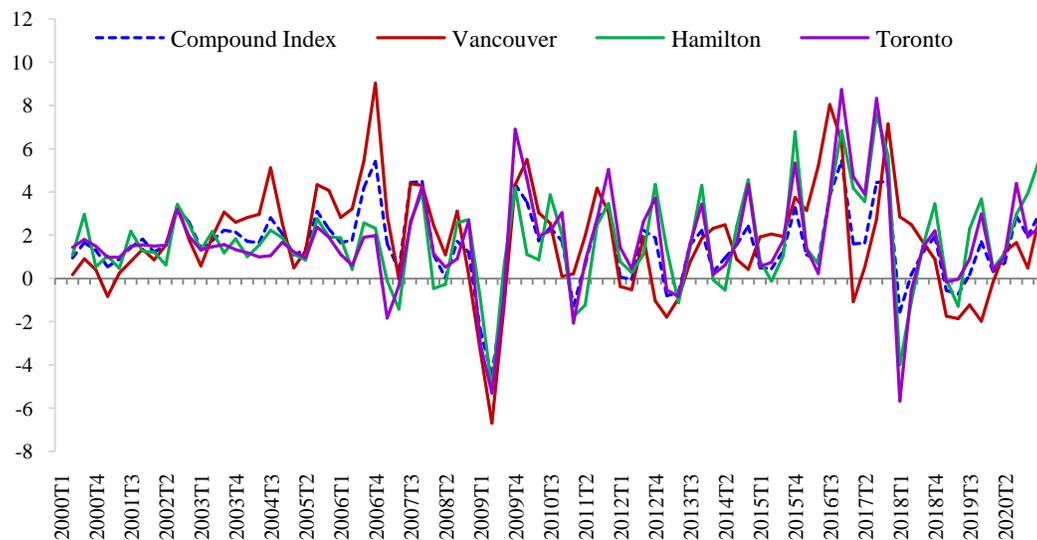
Para los condados de la frontera norte de los Estados Unidos, el índice anual muestra valores más extremos (ver gráfica 4). El mayor aumento anual lo experimentó el condado de Whatcom, frontera con la región metropolitana de Vancouver, con una ganancia de 20.97 puntos en 2005 y 10.91 puntos en 2017. De igual modo, se observan caídas más profundas donde Detroit y el resto del condado de Wayne cayeron 13.40 y 14.02 puntos en 2008 y 2009, respectivamente. Destaca que los condados de Erie y Niagara, se mantuvieron con variaciones positivas en la mayor parte del periodo después de la crisis de hipotecaria, en contraste con el resto de las ciudades estudiadas.

**Gráfica 4. Índice de Precios de Vivienda FRED, anual 2000-2020 (Variación %)**



Fuente: Elaboración propia con base en Índice de Precios de Vivienda de la FRED, 2000-2020.

**Gráfica 5. Índice de Precios de Vivienda Teranet-National Bank House, anual 2000-2020 (Variación %)**



Fuente: Elaboración propia con base en *Teranet-National Bank House Price Index*, 2000-2020.

En Canadá existen índices de precios de vivienda nueva elaborados por la autoridad estadística nacional, sin embargo el índice de *Teranet-National Bank* elaborado desde 1990, representa una medida más completa que integra todos los precios de las casas

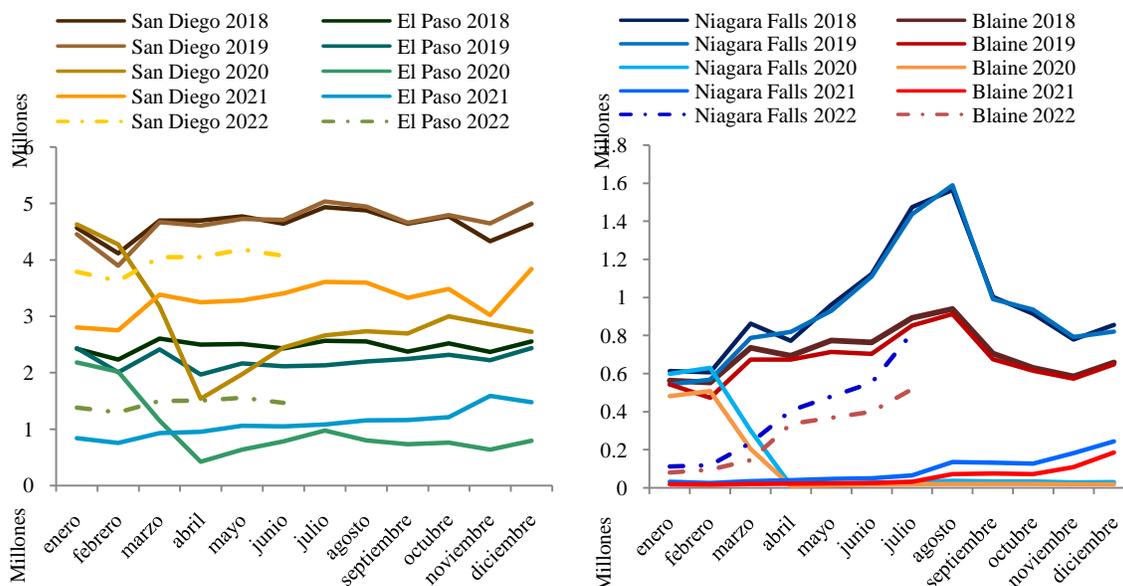
unifamiliares para ciudades seleccionadas con una metodología de ventas repetidas, así como un índice compuesto basado en once ciudades principales. Los aumentos más significativos de precios los ha tenido Vancouver con una variación de 9.03 puntos en el segundo trimestre de 2006 y de 8.08 en el mismo periodo de 2016. Las caídas más significativas, además del periodo de la crisis inmobiliaria, ocurrió a finales de 2017 e inicios de 2018 con caídas de 1.05 a 5.67 para todas las ciudades, con excepción de Vancouver, que se mantuvo con un índice positivo (ver gráfica 6).

Finalmente, a la luz de la coyuntura ocasionada por la pandemia de COVID-19, así como debido a las medidas aplicadas para contener esta crisis sanitaria, las ciudades fronterizas experimentaron procesos de cambio que podrían potencialmente generar secuelas en los mercados de vivienda. Los efectos generales de la pandemia serán más visibles conforme la economía y empleo de las ciudades se estabilice, no obstante es posible identificar algunos cambios importantes.

Cabe resaltar el efecto del cierre de los cruces fronterizos no esenciales hacia Estados Unidos, instaurados desde el 21 de marzo de 2020 y que tuvieron efecto hasta el 7 de noviembre del presente año 2021, en el volumen de cruces de personas. Debido a que la medida de restricción de viajes por puertos terrestres excluyó a los ciudadanos y residentes legales estadounidenses, es decir, sólo a ellos les era permitido cruzar la frontera hacia Estados Unidos, es posible distinguir a partir de esta coyuntura, el tamaño de la interacción por nacionalidad o lugar de residencia.

Por ejemplo, destaca que para los puertos de entrada de entre San Diego y Tijuana, la caída en los cruces de marzo a abril de 2020 fue de alrededor de 1.6 millones, menor del promedio de más de 4 millones en 2018 y 2019. Sin embargo, el cruce de estadounidenses y residentes legales se mantuvo en el orden de los 3 millones de cruces, es decir que se muestra el tamaño de la interacción fronteriza procedente de los Estados Unidos, compuesto en su mayoría por *commuters* o personas que se desplazan para trabajar, en este caso atravesando la frontera, por tanto constituyen un caso de trabajadores transmigrantes. Un comportamiento simétrico a este es el que ocurrió entre la frontera de Ciudad Juárez y el Paso, con una reducción de los cruces en el orden de los 0.7 millones entre marzo y abril de 2020 y alrededor de 1 millón de viajes realizados por *commuters* (ver gráfica 6).

**Gráfica 6. Cruce mensual de personas a Estados Unidos provenientes de México y Canadá por puertos de entrada seleccionados, 2018-2021.**



Notas: San Diego incluye los puertos de entrada de San Ysidro, Cross Border Xpress, Otay Mesa y Tecate.  
Fuente: Elaboración propia con datos del *Bureau of Transportation Statistics* (BTS), 2021.

En la frontera con Canadá, los cruces en el área del Niagara tienen un comportamiento estacional con aumentos en el número de viajes durante la temporada de verano, atribuible al turismo, los cuales tuvieron una reducción superior a 0.5 millones durante las medidas restrictivas de 2020 y 2021. En cuanto al cruce de Blaine en el Condado de Whatcom con el Distrito Metropolitano de Vancouver, se redujo en más de 100 mil cruces entre marzo y abril de 2020, lo que representa 600 mil cruces menos en abril de 2020 en comparación con el año anterior (ver gráfica 6). Por tanto, casi la totalidad de los cruces fronterizos en estos dos puertos de entrada los realizan canadienses o personas de otras nacionalidades, a diferencia de lo que ocurre en las ciudades mexicanas.

A partir de la coyuntura ocasionada por la pandemia de COVID19 y las medidas de control sanitario y fronterizo instauradas por los Estados Unidos, se puede observar la naturaleza que distingue a cada región fronteriza y sus puertos binacionales por la composición de la población que de forma mayoritaria cruza por vía terrestre. En términos generales destacan los pares de ciudades en la frontera México-Estados Unidos con flujos

mayores y continuos de población *commuter*, frente a un comportamiento más estacional en la frontera Estados Unidos-Canadá.

En resumen, las trece aglomeraciones en los límites territoriales de Norteamérica, atraviesan un entorno natural, procesos históricos y condiciones económicas que permiten la comparación de sus mercados de vivienda, tomando en cuenta sus similitudes y diferencias. Por tanto, el área de estudio exhibe una gran diversidad en sus características tanto a nivel intraregional o dentro de cada par de ciudades, como interregional o entre grupos. Al mismo tiempo, esta heterogeneidad justifica la aplicación de enfoques teóricos y metodologías que dan cuenta de esta pluralidad, como lo son los submercados de vivienda y la información asimétrica.

## II. MARCO TEÓRICO

---

El presente apartado establece el marco epistemológico que sustenta esta investigación, la cual se inscribe dentro del estudio de la economía regional. Primeramente, se presenta el estado de la cuestión del campo de estudio de la economía urbana, la metodología de precios hedónicos y el concepto de mercados multijurisdiccionales. Posteriormente, la revisión crítica de la bibliografía para cada uno de los dos temas que componen esta investigación: submercados de vivienda e información asimétrica en el mercado de venta.

Respecto a las pautas principales, la economía urbana es una rama de conocimiento que conjuga la economía y el urbanismo con el objetivo de analizar fenómenos de áreas urbanas como la vivienda, el transporte, entre otras. Esta línea de la economía regional permite incluir al espacio de forma explícita en los análisis, y con ello, dar cuenta de la heterogeneidad del entorno urbano. La metodología de los precios hedónicos constituye una herramienta de valoración de no mercado de las cualidades de bienes y servicios. A partir de esta técnica se estudia a la vivienda como un cúmulo indivisible de atributos intrínsecos, extrínsecos y sociales. Por su parte, los mercados bajo múltiples jurisdicciones reconocen a las áreas que comparten productos, compradores y vendedores, comprendidas entre diferentes regímenes político-administrativos.

En cuanto a la revisión de las líneas de estudio y sus contribuciones, en primer lugar, los submercados de vivienda se distinguen por la similitud y sustituibilidad de las unidades dentro de cada grupo, donde la oferta y demanda para cada submercado es distinta. Entre más parecidas sean unas viviendas entre sí, a la vez que muestren una mayor diferencia con respecto a otras, formarán parte de un mismo submercado. Por tanto, el mercado de vivienda está constituido por estos subgrupos y presenta un rango de precios con equilibrios múltiples.

En segundo lugar, la asimetría de la información da cuenta de las fricciones en el mercado inmobiliario residencial, donde la información es escasa y costosa, lo que lleva a ineficiencias. La información imperfecta es una condición presente en muchos mercados, sin embargo en el sector de bienes raíces se agudiza la asimetría entre aquellos que conoce

de manera más completa los bienes, en este caso los oferentes directos o intermediarios inmobiliarios, frente a los demandantes, en particular si desconocen ese mercado local. Ante esto, las empresas de tecnologías de la información en la industria inmobiliaria han experimentado un auge, producto de la introducción de plataformas digitales, y su irrupción ha tenido efectos en el acceso y costo de la información del mercado de vivienda en general.

## **2.1. Estado de la cuestión**

Los mercados de vivienda son el resultado de un conjunto de condiciones geográficas, políticas, sociales y económicas particulares que se desarrollan a lo largo de procesos históricos. Además, intervienen agentes que configuran cada urbe a partir de sus interacciones, por lo cual las ciudades y sus residencias son heterogéneas. Como resultado, los mercados de vivienda presentan diferentes atributos que pueden ser comparados, con el fin de identificar las variables más significativas que intervienen en su precio y a partir de ello, abstraer un modelo explicativo para un mercado en particular y posteriormente ser replicado en el estudio de otros mercados con características similares.

Esta investigación tiene como propósito generar explicaciones sobre las condiciones de los mercados de vivienda en las ciudades fronterizas entre México, Estados Unidos y Canadá, de modo tal que sea posible analizarlos dentro de un contexto común. A partir de los aportes teóricos conceptuales sobre la identificación de los submercados residenciales y las asimetrías de la información, se diseñaron las pruebas empíricas mediante las cuales se evalúa el efecto de la contigüidad e interacción fronteriza en la oferta, demanda y precios en el área de estudio.

### *2.1.1. Economía urbana*

El análisis de los precios en el mercado residencial se sitúa en el campo de investigación de la economía urbana, que a su vez se inscribe dentro de la economía regional. Este marco epistemológico pone énfasis en el espacio geográfico como un factor principal para el análisis económico con el fin de responder a preguntas relacionadas con la elección de la localización de las actividades productivas y los hogares, así como sus efectos en el territorio.

La economía regional es una disciplina que analiza los fenómenos económicos enmarcados en un espacio geográfico delimitado, por medio de métodos y técnicas que incorporan la dimensión espacial al análisis económico de manera explícita. Tiene como antecedente a los estudios de la localización agrícola e industrial (Thünen, 1826; Launhardt, 1882; Weber, 1929; Hotelling, 1929). Igualmente, incluye los modelos de estructura urbana como el modelo de zonas concéntricas (Burguess, 1925), *Central Place* (Christaller, 1933); modelo sectorial (Hoyt, 1939), equilibrio general espacial (Lösch, 1940), teoría general de la renta de la tierra con *Central Business District* (Alonso, 1964), núcleos múltiples (Harris y Ullman, 1945), *Pôles de Croissance* (Perroux, 1950), patrones de uso residencial (Muth, 1969) y modelo de estructura urbana (Mills, 1972).

Con el fin de brindar reunir este cuerpo de conocimiento, Walter Isard (1956) buscó una teoría general que relacionara la localización de las actividades económicas, el uso del suelo y la estructura urbana, tomando en cuenta factores como el patrón geográfico de la ciudad, el desarrollo de nodos, subcentros y ciudades satélite. De este modo, Isard (1975) propuso la ciencia regional como una disciplina enfocada en los problemas sociales con dimensión regional o espacial empleando una diversa combinación de investigación analítica y empírica (Isard, 1967).

En cuanto a la economía urbana, es una rama que integra la exploración de regularidades y su modelación a través del análisis económico con la búsqueda de las características únicas que distinguen a los sitios desde la perspectiva geográfica (Capello y Nijkamp, 2004). Esta unión de economía y geografía permite indagar la ubicación de las actividades económicas y el sector residencial a partir de una función de utilidad donde las empresas maximizan sus ganancias y los hogares su utilidad (O'Sullivan, 2012).

Este campo de estudio, también denominado en la década de 1970 como nueva economía urbana (Mills y MacKinnon, 1973), busca integrar la economía del bienestar con un modelo de equilibrio general utilizando métodos matemáticos de análisis (Richardson, 1976, p. 137). Para ello, parte del modelo monocéntrico de Alonso –*Central Business District*, CBD–, e incorpora fundamentos microeconómicos como las funciones de utilidad y costes de transporte (Ramírez Franco, 2008). Los trabajos posteriores extendieron el modelo básico e incorporaron centros múltiples, efectos de aglomeración y dispersión.

A su vez, la denominada nueva geografía económica (Fujita, Krugman y Venables, 1999), presenta una modificación a la teoría económica neoclásica desde un enfoque de comercio internacional con fuerzas centrífugas que permiten la aglomeración y las economías de escala. Los trabajos inscritos dentro de esta línea formalizan los retornos crecientes a escala para modelos de centro industrial con periferia agrícola y costes de transporte (Krugman, 1991 y 1995; Fujita, 1989; Fujita y Thisse, 2002).

En suma, el espacio constituye una dimensión que permite incorporar la complejidad y heterogeneidad de las ubicaciones geográficas y las interacciones sociales que ahí se desarrollan como una conjunción de paisaje y sociedad (Santos, 2009). En los aportes arriba señalados se distinguen nociones diferentes sobre el espacio y el equilibrio, que tendrán eco en las propuestas metodológicas posteriores. Por un lado, las que parten de un equilibrio que después de sufrir perturbaciones tenderá a éste nuevamente, y por el otro, las que percibe al desequilibrio como una condición natural que no tiende a contrarrestarse con el tiempo.

### *2.1.2. Precios hedónicos*

El modelo de precios hedónicos desagrega la valoración de los atributos que componen un bien en particular, a partir del cálculo de la disposición a pagar de los agentes económicos por dichas características. Esta metodología se fundamenta en el principio de bienes complejos y heterogéneos que son valorados subjetivamente por los agentes económicos. Existen diferentes tipos de productos que se componen de múltiples características cualitativas, es decir, que son bienes complejos, donde dichos atributos se presentan en conjunto y no puede separarse.

Los bienes raíces cumplen perfectamente con esta definición al ser un conjunto de atributos intrínsecos o relacionados con la propiedad, extrínsecos o propios a su ubicación y demográficos, mismos que no se pueden desasociar. Por lo cual, la aplicación de la metodología de precios hedónicos es una técnica idónea para identificar la valoración subjetiva que hacen los agentes económicos sobre los atributos objetivos del mercado residencial arriba mencionados, mismos que se presentan en un paquete o conjunto indisoluble.

En 1928 Frederick V. Waugh publicó el primer artículo con este enfoque, en el cual calculó con una regresión múltiple los coeficientes parciales, para analizar el efecto de las características físicas de los espárragos en sus precios (Berndt, 1991). La denominación de esta técnica obedece al artículo “*Hedonic price indexes with automotive examples*” de Andrew Court (1939), en el cual creó un índice de preferencias expresadas en unidades monetarias, para valuar las especificaciones de los automóviles y su calidad a través del tiempo. No obstante, la difusión de esta técnica para captar los denominados precios implícitos, precios no revelados, precios ocultos o precios sombra ocurrió hasta dos décadas después de su aplicación bajo el nombre de precios hedónicos (Goodman, 1998).

El trabajo de Zvi Griliches (1961) logró que resurgiera este método relativamente sencillo y directo para ajustar precios ante cambios en la calidad, a través de su análisis de los índices de precios de automóviles mediante una función semi-logarítmica, en el cual los coeficientes estimados de las especificaciones de los vehículos, representan los precios implícitos. Posteriormente, Kelvin Lancaster (1966) propuso la incorporación de esta técnica en el análisis del comportamiento del consumidor, con el fin de dar cuenta de la relevancia de las propiedades intrínsecas de los bienes, a partir de la estimación de los precios como una combinación lineal de sus atributos.

A su vez, Sherwin Rosen (1974), tomó como punto de partida el supuesto de indivisibilidad en el trabajo de Lancaster y su extensión del modelo de Houthakker para la teoría del consumidor y la variación de la calidad de los productos. Además, Rosen impuso condiciones de competencia perfecta y ausencia de arbitraje en su modelo. Los aportes de Griliches y Lancaster también fueron estudiados por Robert Lucas (1975), quien realizó tres aplicaciones del modelo a partir de la teoría de elección del consumidor, la maximización de ganancias por firmas competitivas y la determinación de precios competitivos, respectivamente.

Años más tarde aparecieron los primeros artículos que aplicaron la metodología de los precios hedónicos a los bienes inmuebles. Allen Goodman (1978) extendió el modelo hedónico para la formación de un índice de precios del mercado residencial en un área metropolitana, en el cual comparó el enfoque de equilibrio en los valores de los componentes de la vivienda en el largo plazo, con un modelo de equilibrio en el corto plazo

para submercados separados en el tiempo y el espacio. Dentro de esta línea se inscriben los aportes de Ellickson (1981), Barnett (1985), Shonkwiler y Reynolds (1986), entre otros.

En los últimos años, la técnica de precios hedónicos ha sido ampliamente utilizada en el mercado inmobiliario, sin embargo, subsiste la necesidad de una base teórica más consistente (Malpezzi, 2003). Por tanto, se propone una discusión para asentar la aplicación de los conceptos de bien complejo y heterogéneo de la vivienda, desde un enfoque metodológico que vincule los resultados empíricos del modelo de precios hedónicos, con la noción de desequilibrio en el mercado residencial y en la mejora de los índices de precios.

### *2.1.3. Mercados multijurisdiccionales*

Los mercados de vivienda en conurbaciones fronterizas representan un caso especial de mercados bajo diferentes jurisdicciones, porque constituyen un espacio que comparte características topográficas y medioambientales en una localización contigua, pero que se diferencia por sus instituciones político-administrativas y sociales, así como por sus fundamentales macroeconómicos. Es el caso del área de estudio de esta investigación que comprende las ciudades en la región limítrofe entre México, Estados Unidos y Canadá. Aunado a lo anterior, esta delimitación representa una oportunidad para comparar el efecto de las interacciones espaciales en los mercados de vivienda entre, por un lado, un país en vías de desarrollo y uno desarrollado y por el otro, entre dos países desarrollados.

Referente a los estudios sobre mercados de vivienda en contextos multijurisdiccionales, se encuentran metodologías de precios hedónicos (Black, 1999) matrices espaciales (Helgers y Buyst, 2016) y estadística inferencial (Freemark, 2019) para evaluar el efecto en los precios de los límites entre distritos escolares, las fronteras lingüísticas y las zonas sujetas a diferentes políticas de vivienda, respectivamente. De igual modo, algunos otros trabajos versaron específicamente sobre mercados en fronteras internacionales (Stevenson, 2004; Rérat, Moine, Gertsch, & Signoret, 2011; Zhou & Guo, 2015; Fereidouni, Al-MMulali, Lee, & Mohammed, 2016; Honkanen, Pitkänen, & Hall, 2016; Fullerton & Bujanda, 2017; Micheli, Rouwendal, & Dekkers, 2019).

Respecto al área de estudio en general, el enfoque de jurisdicciones múltiples ha sido poco explorado y se han privilegiado otros abordajes. En particular, sobre la frontera México-Estados Unidos, Rubin-Kurtzman, Ham-Chande y Arsdol (1996), Norman, Feller y

Guertin (2009) y Zhao, Jensen y Zhan (2017), estudiaron el crecimiento poblacional y urbano en el Sur de California-Baja California, Nogales Arizona-Nogales, Sonora y Laredo-Nuevo Laredo, respectivamente. Kincal, Fullerton, Holcomb y Barraza de Anda (2010) analizaron el impacto del ciclo de negocios en el mercado de vivienda de El Paso, Texas, mientras que Herzog y Sohn (2017) revisaron las dinámicas fronterizas en la región San Diego-Tijuana en el contexto de la globalización.

Sobre la frontera Estados Unidos-Canadá, se encuentra un número reducido de investigaciones que de igual modo se abocan a otros temas. Wexler (1996) elaboró una comparación de las políticas de vivienda de Canadá con Estados Unidos, Lucas y Purkayastha (2007) estudiaron las experiencias transnacionales de migrantes entre Canadá-Estados Unidos, y más recientemente Butsic, Hanak y Valletta (2011) analizaron el impacto del cambio climático en los precios inmobiliarios cercanos a *ski resorts* en Estados Unidos y Canadá.

Los estudios señalados al inicio de esta sección, muestran cuan adecuada es la aplicación del marco analítico de las jurisdicciones múltiples cuando se analizan fenómenos que trascienden los límites administrativos, como es el caso de las fronteras internacionales. A pesar de ello, subsiste una carencia en la bibliografía sobre los mercados de vivienda de las fronteras entre México, Estados Unidos y Canadá bajo este enfoque. Por tanto, es pertinente emplear la categoría de análisis de múltiples jurisdicciones en las ciudades fronterizas del área de estudio para examinar el comportamiento de los precios en mercados de vivienda más allá de las fronteras. Además, el carácter multijurisdiccional del área de estudio permite no sólo identificar diferencias y similitudes entre las ciudades estudiadas, sino considerar el producto de este análisis en el abordaje de otras regiones con condiciones similares, aún cuando no se traten de áreas fronterizas.

## **2.2. Revisión crítica de la bibliografía**

En esta sección se presenta la revisión de la bibliografía, a partir de la cual se identifican los principales aportes al estudio de los dos ejes temáticos de esta investigación. Por un lado, los submercados de vivienda y su identificación bajo los tres tipos de segmentación: topográfica, cualitativa y anidada. Por el otro, la información asimétrica y la introducción de fuentes de información especializada en internet, que lleva a la reducción de los costos de búsqueda y la dispersión de los precios. Asimismo, se destacan aspectos relevantes que no han sido suficientemente abordados en trabajos previos, en los cuales se inserta este estudio.

### *2.2.1. Submercados*

La teoría de segmentación de mercados fue planteada desde el campo de las finanzas para analizar la estructura de las tasas de interés para bonos gubernamentales con diferentes plazos, expresados en la curva de rendimientos (Culbertson, 1957). Tiene como supuestos la no sustitución entre bienes de segmentos distintos y la especialización de los inversionistas en una parte específica del mercado. En microeconomía, la segmentación es empleada por la empresa monopolista discriminadora de precios para maximizar sus beneficios en función de la elasticidad precio de la demanda de los consumidores. De igual modo, este concepto se utiliza en el diseño de campañas publicitarias asumiendo una demanda heterogénea (Claycamp & Massy, 1968).

Los productos inmobiliarios son bienes complejos, es decir, compuestos por múltiples atributos, los cuales se caracterizan por tres dimensiones: inmovilidad espacial, durabilidad y heterogeneidad. Esta multidimensión de la vivienda propicia la segmentación por el lado de la oferta, a la vez que devela las diferencias entre los hogares y permite su subdivisión por el lado de la demanda (Tu, 2003). Bajo este enfoque, cada submercado exhibe una oferta y demanda distinta, por lo tanto, existen múltiples equilibrios –o desequilibrios– en el mercado de vivienda con más de un nivel de precios que equipara la oferta y demanda, contrario a los modelos de un sólo mercado de vivienda con equilibrio en el corto plazo.

La existencia de varios equilibrios puede ser constatada ante la presencia de variaciones en los precios implícitos por características individuales o atributos entre los diferentes submercados (Dale-Johnson, 1982). Asimismo, los equilibrios múltiples se

asocian a las amenidades endógenas en una distribución espacial específica (S. Lee & Lin, 2018), es decir, la heterogeneidad individual del espacio. De forma análoga, los modelos de regímenes espaciales dan cuenta de diferentes parámetros que se oponen a la homogeneidad y constancia de la varianza a través de diferentes regiones geográficas (Fischer et al., 1998), por tanto expresan más de un equilibrio.

Por definición, los submercados son conjuntos de bienes que se segmentan en función de cuan similares y sustituibles son éstos entre sí, a la vez que se diferencian de los otros bienes agrupados en el resto de los conjuntos (Galster, 1996; Watkins, 2001; Bourassa, Hoesli y Peng, 2003; Durlauf, 2004; Bates, 2006). Los submercados son resultado de inelasticidad de la demanda causada por costos de transacción, fallos de información y preferencias, así como por una oferta inelástica debido al marco regulatorio o restricciones tecnológicas, entre otras (Ellen, 2012).

Los criterios para su identificación son: 1) la *similitud* expresada como el grado de homogeneidad y sustituibilidad al interior del conjunto; 2) la *simplicidad* en el sentido de un número reducido de submercados que faciliten su interpretación; y 3) la *compactibilidad* es decir, la cercanía entre las unidades de vivienda de un mismo submercado (Tu, 2003).

En general existen tres tipos de submercados de vivienda. En primer lugar se distinguen los submercados de base topográfica o geográfica, los cuales se delimitan por la ubicación espacial de las residencias. Bajo esta clasificación, las propiedades se agrupan con otras contiguas, conformando conglomerados que constituyen submercados (Tu, 2003). Este tipo de submercado ha sido estudiado abundantemente en la bibliografía aplicada y refiere a los atributos de ubicación (Galster, 1996). Los principios bajo los cuales se rige esta categorización son fundamentalmente dos: proximidad, entendida como la concentración de las viviendas más cercanas entre sí en un mismo grupo; y diferencia, es decir que a la vez éstas mantenga la mayor distancia con el resto de los grupos.

En segundo lugar, los submercados con base cualitativa o por características se definen como un conjunto de viviendas que son percibidas por los demandantes y oferentes como productos con una calidad similar, y por lo tanto son sustitutos entre sí (Galster, 1996; Galster & Rothenberg, 1991). A esta característica se le conoce también como el principio de Grigsby de segmentación del mercado de vivienda, por la definición de submercados hecha por William Grigsby en 1963 (Galster, 1996). Bajo esta acepción, el

mercado de vivienda está constituido por un espectro de submercados que agrupan a las propiedades, las cuales se distribuyen a lo largo de una dimensión de calidad (Tu, 2003). Estos grupos de viviendas diferenciados por características han sido relacionados en algunas ocasiones con el filtrado (Galster & Rothenberg, 1991), proceso que se explica más adelante.

En tercer lugar, los submercados anidados son un enfoque híbrido propuesto por Young Tu (1997), en el cual la estructura de submercados está dividida entre submercados topográficos y dentro de cada uno de estos, se encuentran agrupadas las viviendas por calidad (Tu, 2003). De este modo se toman en cuenta los dos tipos de submercados anteriores sin prescindir de ninguna información. El modelo básico consiste en dos niveles, donde la raíz es el mercado de vivienda local, el primer nivel consiste en los submercados topográficos a modo de vecindarios, el segundo nivel contiene los submercados por calidad dentro de cada submercado geográfico y dentro de cada rama se presentan las viviendas (Tu, 1997).

Debido a lo anterior, no existe un sólo precio de mercado sino diferentes precios que varían a lo largo de los submercados de vivienda. Es decir, los submercados persisten como parte de la estructura del mercado residencial debido a las disparidades significativas en los precios asociados a los diferentes grupos de propiedades con características similares. Asimismo, los hogares varían en su comportamiento, características socioeconómicas y preferencias. Dichos consumidores pueden clasificarse por conjuntos, los cuales participan en un proceso de asignación o emparejamiento entre dichos grupos de hogares y los submercados de vivienda (Tu, 1997). Por tanto, el componente idiosincrático es uno de los factores que contribuyen conformación de los submercados, tanto en su fundamentación teórica como en su identificación empírica.

La delimitación de submercados de vivienda es un aspecto relevante a considerar dentro del análisis de precios, ya que diversos estudios empíricos (Bourassa, Hoesli, & Peng, 2003; Calanog, Lynford, & Yang, 2017), encontraron diferencias significativas entre los coeficientes de propiedades residenciales agrupadas por ubicaciones, así como por niveles de calidad, los cuales representan a los atributos en el modelo de precios hedónicos. Por tanto, es importante establecer dichos submercados para calcular estimadores más confiables.

En cuanto a la metodología empleada para caracterizar dichos segmentos, se emplean diferentes modelos y técnicas estadísticas. Para identificar los submercados topográficos se utiliza el análisis de componentes principales, para encontrar la estructura subyacente del entorno urbano. En segundo lugar, para distinguir a los submercados por calidad se realiza el análisis de conglomerados por medio del agrupamiento por K-medias, vecinos más próximos –*K-Nearest Neighbors*, K-NN–, o los conglomerados jerárquicos –HC–. En tercer lugar, para diferenciar a los submercados anidados se emplean técnicas de econometría espacial como la aplicación de matrices de pesos espaciales, modelos de cambios de régimen y de regímenes espaciales (Belasco, Farmer, & Lipscomb, 2012; Wilson, White, Dunse, Cheong, & Zurbruegg, 2011; Wu & Sharma, 2012).

La explicación de las técnicas estadísticas y econométricas antes mencionadas se presenta en extenso en la sección de metodología. En general, los trabajos aplicados emplearon los procedimientos antes señalados, así como otros análisis de tipo espacial y temporal. Sólo en casos especiales se aplicaron métodos cualitativos, por ejemplo, la identificación de patrones por agentes expertos en el mercado inmobiliario (Keskin & Watkins, 2017).

Otro fenómeno que se asocia a los submercados residenciales es el conocido como filtrado –*filtering*–, el cual se mencionó anteriormente. Se trata del proceso mediante el cual los hogares cambian de vivienda según modifican sus necesidades e ingresos, por lo cual una misma propiedad puede ser ocupada por hogares de diferentes características a lo largo de su vida útil, generando un proceso dinámico (Gibb, 2003). A su vez, el filtrado se asocia al ciclo de vida de los vecindarios y a temas como la asequibilidad de la vivienda (Roberts, 1991).

Existen diferentes definiciones de filtrado en el mercado residencial, no obstante, es posible identificar cuatro tipos básicos. Primero, el filtrado de precio, que se refiere a la caída relativa de los precios en un submercado determinado, lo que permite que hogares que se encontraban previamente en viviendas de submercados de menor calidad, se muden a las casas que experimentan el filtrado. Segundo, el filtrado de vivienda, mismo que ocurre cuando los dueños cambian la calidad de sus propiedades con el fin de obtener mejores tasas de retorno (Gibb, 2003).

Tercero, el filtrado pasivo, proceso que tiene lugar cuando las casas que han sido modificadas se encuentran concentradas espacialmente, lo que ocasiona el filtrado para el resto de las propiedades circundantes, producto de la externalidad espacial. Cuarto y último, la dinámica similar de filtrado, cuando los cambios inducidos por los precios conducen en el largo plazo a cambios en los precios y calidad de las viviendas en submercados relacionados así como en la movilidad de los hogares (Gibb, 2003). En suma, el filtrado es un proceso dinámico de los precios de las viviendas, los cambios en la calidad de las mismas y los movimientos de los hogares asociados a estos (Galster, 1996). La relación entre el filtrado y los submercados obedece a que los procesos que ocurren en un subgrupo de viviendas en particular tienen repercusiones de filtrado en otros subconjuntos de propiedades.

A pesar de que las investigaciones especializadas en el mercado de vivienda resaltan la importancia de incluir la categoría de submercados en sus análisis (Bates, 2006; Bourassa et al., 2003; Goodman & Thibodeau, 2003; Tu, Sun, & Yu, 2007; Watkins, 2001), no existe un consenso respecto a cuál metodología ofrece una mejor segmentación que cumpla con los tres criterios de los submercados: simplicidad, similitud y compactibilidad (Bourassa, Hamelink, Hoesli y Macgregor, 1999; Tu, 2003), así como con las propiedades estadísticas de los resultados insesgados, eficientes, consistentes y suficientes.

De este modo, existe una clara necesidad en la bibliografía de submercados de contar con una metodología precisa que pruebe cumplir con los requisitos antes señalados en mercados con características similares. Es por ello que esta investigación sobre los mercados de las conurbaciones entre las fronteras de México, Estados Unidos y Canadá, constituye un aporte a la conciliación en la identificación de submercados de calidad bajo el principio de similitud, con la identificación geográfica, por contigüidad.

Es importante hacer hincapié en el tercer criterio de los submercados, la compactibilidad, y su relación con las aglomeraciones fronterizas. Las comunidades establecidas en la franja fronteriza cumplen con los principios de proximidad y contigüidad, por lo cual conviene considerarlas como un espacio común para la aplicación del análisis de submercados, un enfoque que no ha sido propuesto en otros estudios del área. Por ello, esta investigación contribuye con la metodología de identificación de submercados de vivienda fronterizas a través de pruebas empíricas que permiten fortalecer esta técnica.

**Tabla 6 Definición operacional de las variables e hipótesis de submercados**

Conceptos	Variables	Principio	Categorías	Unidad de análisis	Indicadores	Hipótesis
Submercados topográficos	Ubicación	Proximidad	Conglomerados	Vivienda	Distancia euclidiana	$J = S_1 + S_2 + \dots S_n$
Submercados por calidad	Características intrínsecas	Similitud	Factores comunes	Vivienda	Factores subyacentes	$\sigma p_{ij} < \sigma p_i$
Submercados anidados	Ubicación y características intrínsecas	Proximidad y similitud	Regímenes espaciales	Vivienda	Subconjuntos	$B_{MIG} \neq 0$

Notas: Variación de precio en estimación con submercados, menor a estimación sin submercados:  $\sigma p_{ij} < \sigma p_i$

Submercado  $j \subset J$  donde  $j = 1, 2, 3 \dots n$

Mercado de vivienda  $J = S_1 + S_2 + \dots S_n$

Estimadores  $B_{ij}$  para el submercado de vivienda  $S_1$

Contexto fronterizo captado por variables migratorias  $B_{MIG} \neq 0$

Fuente: Elaboración propia.

### 2.2.2. Información asimétrica

Existe una asimetría de la información cuando el precio, las características de los productos o las condiciones del contrato no son plenamente conocidas por al menos una de las partes en la transacción, situación que lleva a fallas de mercado e ineficiencias. En economía, el fenómeno de la información asimétrica se estudia en el marco de la toma de decisiones y bajo el problema del agente y el principal, con casos como la selección adversa (Akerlof, 1970) y el riesgo moral (Arrow, 1963). Asimismo, la búsqueda (Stigler, 1961) y la coordinación, son gastos en los que se incurre en los procesos de compra-venta, entre otros costos de transacción (O. E. Williamson, 1979; S. D. Williamson, 1986).

En economía de la información, la búsqueda se refiere al proceso mediante el cual un comprador o vendedor que desee averiguar el precio más favorable, debe sondear a varios vendedores o compradores. Este fenómeno de consulta está determinado por el rango de precios de venta, el número de vendedores, el precio mínimo esperado por el comprador, el costo de búsqueda por unidad, la cantidad fija de compras y los gastos totales. Los mecanismos de identificación de compradores y vendedores, por ejemplo a través de la publicidad, reducen drásticamente los costos de búsqueda (Stigler, 1961). Mientras que los costos de coordinación son ocasionados por la supervisión, control y gestión de las transacciones (Young, 2013).

Entre las estrategias para responder a esta incertidumbre, producto de la asimetría de información, destacan la señalización y el escaneo. La señalización –*market signaling*– fue propuesta por Michael Spence (Spence, 1973, 1981), para explicar el proceso mediante el cual una de las partes envía una pieza de información, denominada señal, que devela algo relevante a la otra parte. De este modo, las señales proporcionan información más allá del mensaje formal, con el fin de inducir algún comportamiento. Éstas se envían entre los agentes económicos incluyendo competidores, clientes, proveedores, accionistas, gobierno y sociedad en general. Por su parte, la selección o escaneo –*screening*– consiste en clasificar el bien o servicio según sus cualidades (Rothschild & Stiglitz, 1976; Stiglitz, 1975). Este dispositivo de selección permite identificar la calidad de los productos y los etiqueta, para así resolver el problema de la selección adversa.

En el mercado de vivienda, así como en el mercado de bienes raíces en general, la información suele ser escasa y costosa (Ball, 2006). El mercado inmobiliario es poco líquido, por tanto el mecanismo de precios se transmite más lentamente a los participantes del mercado, quienes se enfrentan a fricciones. Los vendedores destinan recursos para anunciar sus propiedades y descubrir los precios de otros; al mismo tiempo, los compradores potenciales tienen costos de búsqueda para conocer la oferta y precios de diferentes viviendas, mientras que algunas firmas recopilan información sobre los precios para venderla a las partes interesadas (Fallis, 1985). Todo ello deja de manifiesto el importante papel que desempeña la información asimétrica en el mercado residencial, sin embargo, existen relativamente pocos estudios empíricos que prueben la teoría.

Los trabajos aplicados al mercado de vivienda pueden agruparse bajo tres enfoques. En primer lugar, los que examinan la asimetría de información entre diferentes tipos de individuos. Es el caso de Chau y Choy (2011), quienes estudiaron a los compradores y vendedores al medir el impacto de las regulaciones legales bajo las doctrinas *caveat emptor* en la cual comprador asume el riesgo frente a *caveat venditor* en donde vendedor asume el riesgo, en los precios de las viviendas. En investigaciones independientes Levitt y Syverson (2008) así como Xie (2018), analizaron desde un enfoque microeconómico la asimetría de la información entre clientes y agentes inmobiliarios, bajo el modelo del problema del agente y el principal.

A su vez, los artículos de Chincó y Mayer (2016), Holmes y Xie (2018), y Zhou, Gibler y Zahirovic-Herbert (2015), compararon los niveles de información de compradores locales frente a compradores foráneos. Todos los artículos mencionados, encontraron diferencias significativas en cuanto informados están los grupos de agentes a partir de los niveles de precios y las características de compra-venta, lo que constituye evidencia de asimetría en el mercado de vivienda.

En segundo lugar, se distinguen los trabajos que identifican efectos espaciales. Garmaiser y Moskowitz (2004), analizaron la proximidad de origen de los compradores con respecto a las propiedades a partir de la identificación de variaciones exógenas en el ambiente de información. Alti, Engelberg y Parsons (2012), probaron el efecto del desbordamiento *–spillover–* de información de un mercado residencial hacia los mercados cercanos a través del mecanismo de revelación de precios.

Por su parte, Badarinza (2018) revisó el desbordamiento de información a partir de la dependencia espacial de la migración, las transferencias de capital y los flujos de información, formas de retroalimentación que general aprendizaje espacial. Algunas formas de mitigar la asimetría de información por parte de los agentes menos informados incluye la compra de propiedades cercanas a su residencia pasada o presente, la preferencia por aquellas que contienen más historial disponible y que no son representadas por agencias de bienes raíces (Garmaise & Moskowitz, 2004).

En tercer lugar, se identifican las obras que estudian el acceso a nuevas fuentes de información pública disponible, en particular, la publicación de sitios web especializados sobre el mercado de vivienda y sus efectos en el mismo. Eerola y Lyytikäinen (2015), analizaron el impacto de la publicación de una página web gubernamental en el funcionamiento del mercado de vivienda. La evaluación se estimó a partir de un modelo de diferencia en diferencias *–DiD–* con el cual se compararon los precios de venta para transacciones de viviendas con características similares antes y después de que la página web fue puesta en línea. Los resultados arrojaron que la información detallada sobre el mercado reduce el tiempo de venta y es más valiosa para los vendedores que subestiman el valor de sus propiedades que para aquellos que lo sobreestiman.

El trabajo de Lee y Sasaki (2018), examinó cuán sensibles son las transacciones de casas a las estimaciones de precios de la página de bienes raíces *Zillow*, conocido como

*Zestimate*. El efecto fue medido a partir de un modelo de precios hedónicos que incorpora el efecto de la estimación en el precio final de la transacción, donde se encontró que el *Zestimate* y el precio final están altamente correlacionados.

Destaca la investigación de Katrine Kandlbinder (2018) sobre la disponibilidad, demanda y oferta de información y sus efectos en el mercado inmobiliario, de la cual se desprenden otros artículos (Kandlbinder, 2017; Kandlbinder, Miller, & Sklarz, 2019), en los cuales se comparó el *premium* que pagan los compradores foráneos frente a los locales antes y después del lanzamiento de las páginas de búsqueda de vivienda Zillow, Trulia y Redfin. Entre los resultados se confirmó que los compradores fuera del mercado local pagan mayores precios, experimentan mayores costos de búsqueda y presentan un sesgo cognitivo conocido como efecto de anclaje, en el cual los compradores provenientes de mercados con precios relativos mayores al mercado en el cual buscan adquirir propiedades, toman como referencia los precios de su lugar de origen.

En general, los estudios consultados sobre información asimétrica y páginas especializadas en internet encontraron que la liberación de detalles como el precio, las características de las viviendas y las condiciones de las transacciones, benefician a los oferentes ya que tienden a reducir el tiempo de venta de una propiedad. Asimismo, ayuda a los agentes con menos información sobre el mercado de vivienda, como los comparadores foráneos o los dueños de propiedades que no son profesionales inmobiliarios, ya que permite subsanar algunos costos de búsqueda. De este modo se acota la incertidumbre asociada a la asimetría entre los agentes.

A pesar de que el acceso a la información disponible para todo público por medio de páginas y aplicaciones web sobre el mercado de vivienda es un fenómeno relativamente reciente, los trabajos en la materia que se mencionaron anteriormente apuntan a la relación entre la disponibilidad en internet de datos sobre el mercado y los precios. Esta línea de investigación representa una nueva veta en la bibliografía aplicada al análisis de la información asimétrica en el mercado de vivienda, camino que sigue este trabajo.

La información imperfecta representa una condición del mercado que enriquece el análisis económico, al considerar factores como la toma de decisiones, los costos de transacción, la selección adversa y la búsqueda o sondeo. En el mercado inmobiliario

residencial, se encuentra evidencia de este fenómeno, como dan cuenta las obras aquí mencionadas.

Por tanto, ante el creciente acceso a la información en internet sobre precios y características de las viviendas, aunado al extendido uso de este recurso, el presente estudio se inserta en esta incipiente línea de investigación, a partir de la comparación entre el acceso, cantidad y calidad de la información sobre los mercados de vivienda en el área de estudio. Este enfoque enfatiza la importancia de la asimetría de la información en los mercados de vivienda e indaga los efectos de las tecnologías de la información y la comunicación, cada vez más presentes en la dinámica de precios del mercado residencial.

**Tabla 7 Definición operacional de las variables e hipótesis de información asimétrica**

Conceptos	Variables	Principio	Categorías	Unidad de análisis	Indicadores	Hipótesis
Selección adversa	Precios	Dispersión	Unidades tratadas	Viviendas enlistadas en páginas web	Nivel de precios y dispersión	$p_{t_0} > p_{t_1}$
Búsqueda	Información web	Nivel	Cantidad y calidad		Índice de información	Mayor info. $\sigma < \sigma$ menor info.

Fuente: Elaboración propia.

En resumen, la presente investigación busca responder a las interrogantes sobre la oferta, demanda y valoración de los precios de los mercados de vivienda en las regiones limítrofes entre México, Estados Unidos y Canadá, bajo el concepto de mercados multijurisdiccionales. Se enmarca bajo el enfoque teórico-metodológico de la economía urbana, a partir de la cual se estructuran las dos líneas temáticas centrales: de los submercados de vivienda y la información asimétrica. El diseño metodológico parte del modelo de precios hedónicos, y la aplicación de técnicas de estadística multivariante y herramientas econométricas.

### III. METODOLOGÍA

---

La estrategia de investigación del presente trabajo está estructurada en un nivel general y dos específicos, estos últimos obedecen a cada uno de los temas eje de este estudio: los submercados de vivienda y la información asimétrica. El diseño general de esta investigación está sustentado en un enfoque cuantitativo que opera bajo los conceptos de la economía urbana y de la economía de la información como ejes rectores. De igual modo, la metodología se fundamenta en la implementación de herramientas de la estadística multivariante y la econometría espacial. A continuación, se delinea la estrategia base y en los apartados subsiguientes de esta sección, se presentan los datos y la metodología particular de cada uno de los dos ejes temáticos antes señalados.

#### **3.1. Diseño de investigación general**

El estudio general de los mercados de vivienda en las aglomeraciones urbanas entre México, Estados Unidos y Canadá se estableció en dos niveles de desagregación como unidades de análisis: la vivienda y el área geostatística básica. De este modo, la dimensión espacial comprende la información disponible de forma pública por unidad individual, así como aquella que es generada de forma agregada pero que a su vez mantiene una escala geográfica pequeña, por tanto contiene un mayor nivel de detalle. Respecto a la dimensión temporal, se establecen diferentes cortes de tiempo durante el periodo comprendido del año 2000 al 2021.

##### *3.1.1. Fuentes de datos*

Las fuentes de información que se utilizaron para esta investigación se dividen en dos tipos. El primero proviene de fuentes oficiales de información que incluyen censos, encuestas y capas de datos espaciales elaboradas las autoridades estadísticas de cada país: el Instituto Nacional de Estadística y Geografía (INEGI) para México, *United States Census Bureau* (US Census) para Estados Unidos y *Statistics Canada* (StatCan) para Canadá.

La comparación entre estos tres países a partir de la información estadística es posible gracias a la armonización de bases de datos elaborada por IPUMS. El proyecto IPUMS (*Integrated Public Use Microdata Series*) de la Universidad de Minnesota, aplica una metodología de armonización a las bases de datos de censos y encuestas de varios países, mediante la cual empatan las variables que describen las mismas categorías. Después les agrega un código y una etiqueta común a las variables integradas. Dichos códigos se mantienen a lo largo de diferentes muestras, por lo tanto permiten mantener la consistencia a través del tiempo. Adicionalmente, incluyen notas para explicar cuestiones que no son suficientemente explícitas en los códigos o etiquetas (IPUMS, 2021).

De este modo, fue posible trabajar con diferentes fuentes de datos de las agencias estadísticas de Estados Unidos, México y Canadá, las cuales forman parte de los proyectos IPUMS USA e IPUMS International, el repositorio más grande de su tipo en el mundo (Sobek, 2020; Sobek & Cleveland, 2017). Aunado a lo anterior, la armonización de estas bases de datos permitió asegurar la compatibilidad, a la vez que dio acceso a la consulta de los valores originales de las variables, es decir, sin desperdiciar ninguna información.

El segundo tipo de fuente comprendió a las bases de datos que se integrarán expresamente para esta investigación con información de oferta de vivienda en venta. Esta información se recopiló de páginas de internet especializadas en bienes raíces, sitios web de mercados en línea –*online marketplace*– de propósito general, así como aplicaciones. El proceso de recopilación se denomina *web scraping*, es decir, extracción masiva de datos web.

En este respecto, las técnicas de recolección y manejo de datos para el conjunto de información de oferta residencial, incluyen la extracción de precios de lista y características de los productos inmobiliarios. La muestra de propiedades en venta se tomó de los portales web de empresas digitales especializadas, las cuales albergan la mayor oferta para las ciudades incluidas en el estudio, como se describe en el subapartado de submercados de esta misma sección.

Una vez identificadas las fuentes de información, se procedió a establecer la estructura y diseño de la base de datos en la cual se concentraron los registros, con el fin de convertir los datos no estructurados en información estructurada. Posteriormente se

programó el código con las instrucciones para la recolección mediante la técnica de extracción de datos web.

El siguiente paso en el preprocesamiento consistió en la extracción e importación de los datos, repitiéndose este proceso en repetidas ocasiones hasta que se consiguió una muestra representativa de cada mercado. Después de coleccionar la información, se realizó la limpieza, verificación y consolidación del conjunto de datos. Finalmente, ya que se contó con la estructura de la base de datos, el código de extracción, la recolección, la limpieza y consolidación de los datos, se aplicó su procesamiento y análisis, siguiendo la guía de la estrategia metodológica que aquí se describe.

### *3.1.2. Modelo base*

La operacionalización de las variables y formalización de los modelos, tuvo como punto de partida la fundamentación teórica y el examen estadístico, a partir del análisis exploratorio de datos en su versión convencional y espacial (Anselin, 2010). Con ello se buscó conocer la distribución y propiedades de las variables, así como las relaciones existentes entre las mismas. Consecutivamente, se realizó la estimación del modelo de precios hedónicos, el cual constituye el modelo base o restringido. Por último, se aplicaron técnicas de econometría espacial para controlar los efectos de contigüidad e interacción (Anselin, Florax, & Rey, 2004).

El cálculo del modelo de precios hedónicos permitió determinar la contribución de las características estructurales, de ubicación, del vecindario y de los ocupantes, en el precio de los bienes raíces, así como la identificación de las relaciones entre las mismas. El modelo básico de precios hedónicos es el siguiente:

$$P_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_n X_{ni} + \mu_i \quad (1)$$

En esta ecuación la variable dependiente  $P_i$  es la valoración del precio del bien raíz, cada  $X_i$  corresponde a las características objetivas del inmueble  $i$ , el coeficiente  $\beta$  corresponde a la valoración de la cualidad  $n$  y el término de error  $\mu_i$ . Con una forma funcional en logaritmos, los coeficientes manifiestan la propensión marginal a pagar por un producto inmobiliario dados sus atributos. Cabe resaltar que la valoración del precio por

calidades puede diferir del precio de mercado, el cual es producto de la interacción entre la oferta y la demanda.

La segunda etapa de esta estimación correspondió a la recuperación de la información geográfica a través de un modelo espacial, atendiendo a su vez a la identificación de submercados con características particulares dentro de cada mercado regional. Para la estimación del modelo espacial se utilizó un modelo que permitió identificar y en su caso corregir el comportamiento de rezago espacial del siguiente modo:

$$y = \rho W y + X \beta + \varepsilon \quad (2)$$

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

Donde  $\rho$  representa el coeficiente de la variable dependiente rezagada,  $W$  es la matriz de contigüidad espacial,  $X$  es la matriz de datos que contiene las variables explicativas por un parámetro  $\beta$  que refleja la influencia de las variables explicativas en el cambio de la variable dependiente más  $\varepsilon$  el residuo o término de error aleatorio. Para detectar y en su caso corregir la dependencia espacial en los errores, también conocido como error espacial, se emplea el estadístico de los multiplicadores de Lagrange representado por  $\lambda$ , la variable dependiente  $y$  está en función de una matriz de variables explicativas  $X$  por un parámetro  $\beta$ .

$$y = X \beta + u \quad (3)$$

$$u = \lambda W u + \varepsilon$$

$$\varepsilon \sim N(0, \sigma^2 I_n)$$

La diferencia respecto al modelo anterior estriba en que la dependencia está en el término de perturbación donde el residuo es igual a un parámetro  $\lambda$  que representa el coeficiente de los errores correlacionados espacialmente multiplicado por la matriz de contigüidad espacial  $W$  y  $\varepsilon$  término de error aleatorio.

## 3.2. Metodología y datos por eje de investigación

### 3.2.1. Submercados

La metodología que se aplicó para identificar a los submercados comprende tres etapas, la primera con un ejercicio de clasificación para la unidad de análisis agregada, la segunda etapa con otro ejercicio de clasificación para las unidades individuales, y una tercera etapa en la cual se consideran en conjunto las dos clasificaciones anteriores. El primer ejercicio se realizó con un nivel de desagregación por unidad geográfica, para el caso de México por Área Geoestadística Básica (AGEB), por *Block Group* (BG) para Estados Unidos y por *Census Subdivision* (CSD) para Canadá. El segundo ejercicio de clasificación se efectuó en un nivel de análisis por vivienda, a partir de su localización por coordenadas geográficas y las variables que representan los atributos intrínsecos de cada mercado inmobiliario residencial para las veintiséis ciudades de México, Estados Unidos y Canadá.

Respecto al primer ejercicio de clasificación de submercados, las características del vecindario se recuperaron de diversas fuentes para cada país, las cuales se enuncian a continuación. Fueron elaboradas con fines distintos a la identificación de cualidades del entorno urbano. A pesar de ello, los repositorios consultados cubren las ciudades del área de estudio, presentan un nivel de desagregación geográfica en unidades de análisis pequeñas lo suficientemente específicas y capturan atributos similares.

Las bases de datos empleadas fueron para México, el *Inventario Nacional de Vivienda* (INV) 2016 de INEGI y el Índice de Desigualdad Urbana 2021 del *World Resources Institute*. Para Canadá se utilizó el conjunto de bases del entorno de datos abiertos –*Linkable Open Data Environment* (LODE) –, proyecto estadístico elaborado por StatCan con el fin de contribuir en el análisis de los bienes y servicios en las comunidades, en un esfuerzo por armonizar y hacer disponible dicha información, el cual se vio impulsado bajo el contexto del confinamiento por la pandemia de COVID19. Dichas fuentes son el *Open Database Cultural and Art Facilities* (ODCAF) del año 2020, *Open Database Educational Facilities* (ODEF) de 2021, *Open Database Healthcare Facilities* (ODHF) de 2020, y el *Open Database Recreational and Sport Facilities* (ODRSF) de 2021.

En el caso de Estados Unidos, no se encontró un proyecto estadístico agregador que comprenda todas los atributos a evaluar, por lo cual se recurrió a diversas fuentes con el fin de integrar sus mediciones. Estas fueron el *Location Affordability Index* 2017 del

*Department of Housing and Urban Development (HUD)*, el *Public Libraries Survey 2020* del *Institute of Museum and Library Services (IMLS)*, el *National Walkability Index 2019* del *U.S. Environmental Protection Agency (EPA)*, y el conjunto de datos *U.S. ParkServe 2022* de la organización no gubernamental *Trust for Public Land (TPL)*.

Cabe resaltar que las condiciones que representan las características de los vecindarios, experimentan en general cambios más lentos y menos considerables. Es por ello que la diferencia en la fecha de captura y publicación de las bases de datos empleadas arriba mencionadas, no representa ningún obstáculo para la identificación de los submercados geográficos.

En cuanto el segundo ejercicio de identificación de submercados, éste se realizó en nivel de desagregación por vivienda. Para este examen se empleó información procedente de páginas de internet con oferta de bienes raíces residenciales en las áreas de estudio, las cuales contienen las características de las viviendas y sus precios de lista. Los conjuntos de datos fueron recabados mediante *web scraping* y las fuentes fueron los principales sitios especializados con cobertura en los mercados fronterizos de México, Estados Unidos y Canadá. Las variables a considerar se dividieron en tipo de vivienda y características estructurales, atributos internos y amenidades externas.

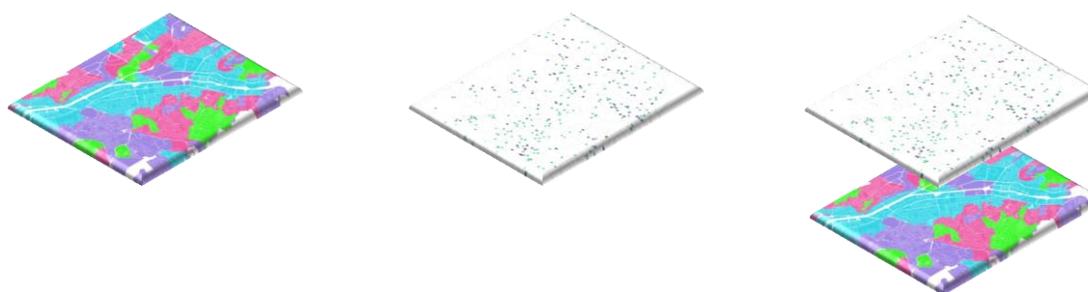
En concreto, se extrajo información de los listados de oferta de vivienda que se encuentran disponible en los sitios web especializados *Century21* en su versión para Estados Unidos, Canadá y Global, *Realtor*, *Point2Homes* y *Vivanuncios*. Por otra parte, para la caracterización de los mercados y análisis exploratorio, se consultaron instrumentos estadísticos como la Encuesta Nacional de Ingreso-Gasto de los Hogares (ENIGH) con frecuencia bianual de 2000 a 2018, la *American Community Survey (ACS)* anual de 2000 a 2019 y del *Census of Population Program* de Canadá con frecuencia quinquenal del 2001 al 2021.

Las variables demográficas se obtuvieron de los proyectos estadísticos con mayor cobertura para cada país. En México, los Censos de Población y Vivienda 200, 2010 y 2020, Conteo2005 y Encuesta Intercensal 2015. Para los Estados Unidos, los proyectos *Decennial Census of Population and Housing* 2000, 2010 y 2020, y el *American Community Survey* 2005 y 2015. En el caso de Canadá, los programas *Census of*

*Population* 2001, 2006, 2011, 2016 y 2021. Todas las bases de datos referidas son elaboradas por las autoridades estadísticas de cada país: INEGI, US Census y StatCan.

Como se mencionó anteriormente, se aplicaron tres etapas de identificación de submercados. Éstas obedecen a los tres tipos de submercados definidos en el marco teórico: submercados topográficos o geográficos, con base cualitativa o por características y anidados. A continuación se describe la metodología aplicada para la delimitación de los mismos.

### Esquema 1. Resumen de métodos de clasificación por tipo de submercado



*Submercados:*

Geográficos

Por características

Anidados

*Nivel de desagregación:*

Vecindarios

Viviendas

Viviendas en vecindarios

*Técnica de clasificación:*

Análisis de Componentes  
Principales (PCA)

K-Medias,  
Conglomerados  
Jerárquicos (HC)

Regímenes espaciales

Fuente: Elaboración propia.

En primer lugar, los submercados topográficos obedecen al principio de proximidad geográfica. Para su clasificación se utilizó el método de análisis de componentes principales –*Principal Component Analysis, PCA*–. Esta técnica permite encontrar los factores subyacentes que comparten las observaciones de una muestra, a partir de un análisis estadístico que identifica la varianza de las variables originales y produce unas nuevas variables que concentran la información de las anteriores.

Con este método estadístico se logra la aproximación a los factores subyacentes del conjunto de datos, por medio de los cuales se realiza la agrupación en submercados que comparten la presencia de dichos factores. Una técnica estadística cercana al análisis de componentes principales es el análisis factorial, el cual adiciona el supuesto de la presencia

de una estructura causal subyacente, la cual se muestra por la matriz de covarianzas. La formalización del modelo factorial se presenta a continuación:

$$x_{ij} = a_{i1} \cdot F_{1j} + a_{i2} \cdot F_{2j} + a_{i3} \cdot F_{3j} + \dots + a_{ij} \cdot F_{ij} + d_i \cdot U_{ij} \quad (4)$$

donde se obtiene el valor normalizado de la variable  $i$  para la observación  $j$ , que es igual a la relación entre la variable y el factor expresada por el parámetro  $a$ , multiplicado por el valor del factor  $F$  más un componente aleatorio  $d_j \cdot U_{ij}$ .

La segunda etapa de identificación pertenece a los submercados por características, los cuales cumplen con el principio de similitud de atributos. La clasificación de este submercado se efectuó por medio de distintas técnicas de análisis de conglomerados o *clustering*, con el objetivo de agrupar a las viviendas que se asocian por compartir características intrínsecas similares.

Dentro de esta metodología se encuentran en primer lugar, las técnicas de conglomerados por K-Medias, K-Medianas o K-Medoids, donde  $K$  representa el número de clústeres o grupos a partir de los cuales se clasificarán las observaciones a partir de la distancia entre dichos puntos y el centroide de cada K. En segundo lugar, el clasificador de K-vecinos más cercanos –*K-Nearests Neighbor*, K-NN–, donde K indica el número de vecinos a considerar en la clasificación bajo el principio de mayoría. En tercer lugar, se encuentra la clasificación por conglomerados jerárquicos –*Hierarchical Clustering*, HC–, el cual puede realizarse a partir de la agregación o subdivisión de las observaciones una a una en forma de árbol. Las técnicas de análisis de datos antes mencionadas constituyen formas de clasificación no supervisada en el caso de K-Medias y HC, y supervisada con la técnica de K-NN. En todos los casos, se calcula la distancia entre todas las unidades de observación, uno de los cálculos de más utilizados es la distancia euclidiana, la cual entre dos puntos  $i$  y  $j$  con coordenadas  $X$  se define como:

$$D(i, j) = \sqrt{A^2 + B^2} = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2} \quad (5)$$

Dicha distancia es recalculada durante un gran número definido de iteraciones, con el fin de encontrar para cada observación, el conglomerado correspondiente que minimizar su distancia, a la vez que maximice la distancia entre grupos. Es decir, se busca reducir la distancia intra-clúster con el fin de agrupar unidades más homogéneas en su interior, al

mismo tiempo que se aumenta la distancia inter-clúster, para lograr grupos más heterogéneos entre sí.

En particular, la técnica de K-Medias calcula la distancia euclidiana de cada unidad de observación hasta un centroide, que corresponde al centro de cada clúster y representa el promedio de cada conglomerado. En el caso de K-Medias, el centro del clúster corresponde a la mediana, la cual tiene la ventaja de ser menos sensible a datos atípicos. El punto en el cual se ubicará un centroide  $C_i$ , se calcula del siguiente modo:

$$C_i(x_1, x_2, \dots, x_k)^n = \left( \frac{\sum_{i=1}^k x_{1i}}{k}, \frac{\sum_{i=1}^k x_{2i}}{k}, \dots, \frac{\sum_{i=1}^k x_{ni}}{k} \right) \quad (6)$$

Las distancias se calcula a lo largo de numerosas iteraciones hasta lograr la minimización intra-clúster y la maximización inter-clúster. Pero en este caso, también se calculan los centroides tras cada iteración hasta que éstos no cambien, es decir, sean estables. Por tanto, la formalización del método de K-Medias representa la minimización de la suma de los cuadrados al interior de cada conglomerado –WCSS, *Within-Cluster Sum of Squares*–, y se expresa como sigue:

$$WCSS = \sum_{i=1}^K \sum_{x \in C_i} \text{dist}^2(m_i, x) \quad (7)$$

Para esta investigación, el número de conglomerados utilizado en cada submercado dependió de los tres principios de los submercados establecidos en el marco teórico: similitud, simplicidad y compactibilidad, así como de los resultados de la optimización matemática conocida como método del codo –*Elbow Curve*–. A partir de este proceso, se calculó el valor óptimo conglomerados, el cual es igual al punto de inflexión a partir del cual se obtienen rendimientos decrecientes. Es decir, se establece el número de clústeres a partir del cual ya no se obtiene una mejor clasificación ni se aporta información. De igual modo, se aplicaron diferentes pruebas estadísticas para corroborar la validez del rango de distancia empleado y de los conglomerados resultantes.

La tercera y última etapa de la identificación consistió en los submercados anidados. Esta clasificación retoma los principios de los dos agrupamientos anteriores: proximidad y similitud. Primeramente, se evaluó si la estructura del conjunto de variables presenta heterogeneidad espacial, dependencia espacial o ambas simultáneamente. La identificación

y corrección de dichos problemas asociados a las observaciones geográficas se realizó con el fin de comprobar la correcta especificación del modelo y asegurar la calidad de los estimadores obtenidos.

Para la delimitación de los submercados anidados, existen distintas técnicas econométricas que permitan estimar la dependencia de interacciones no observadas. Una especificación que fundamenta esta relación encuentra en el modelo de cambio de régimen—*switching regression*—. Esta modelación permitió lidiar con casos en los cuales la pendiente de los coeficientes así como del intercepto, cambia bajo diferentes regímenes, es decir, afectan a toda la función (Páez & Scott, 2004).

Una de los métodos de econometría espacial que forma parte de este grupo de modelos de cambio de régimen, son los regímenes espaciales —*spatial regimes*—. La aplicación del modelo de regímenes espaciales permitió dar cuenta de la heterogeneidad por grupos, a partir de la cual se puso a prueba la estabilidad estructural del modelo y se estimaron diferentes regresiones para cada conjunto (Anselin, Florax, & Rey, 2004). Por tanto, con este método se logra detectar los cambios en la valoración de los atributos en diferentes submercados. Formalmente se expresa del siguiente modo:

$$y_{ij} = \alpha_j + X'_{ij} \beta_j + \epsilon_{ij}, \quad i = 1, \dots, n \quad (8)$$

Las tres etapas aquí descritas se aplicaron para la identificación de los submercados geográficos o por características del vecindario, por atributos intrínsecos de las viviendas y de manera conjunta en los submercados anidados. A partir de estos ejercicios se obtuvieron los submercados que cumplan con los tres principios de proximidad geográfica, similitud de atributos y simplicidad. Con ello se busca responder a la pregunta de investigación a partir de una metodología parsimoniosa, que concilió los beneficios de los submercados geográficos con los submercados por atributos. Los submercados anidados resultantes brindan un panorama de la heterogeneidad de los mercados de vivienda en las conurbaciones fronterizas entre México, Estados Unidos y Canadá, a la vez que son útiles como categoría de análisis para las partes subsecuentes de esta investigación.

### 3.2.2. Información asimétrica

El diseño de la investigación sobre la asimetría de la información en el mercado de vivienda se estructura a partir de dos fases. La primera de ellas consistió en la evaluación de la cantidad y calidad de la información disponible en las páginas especializadas en oferta inmobiliaria. La segunda fase consistió en medir la dispersión de los precios de vivienda y comparar la desviación estándar de los niveles de precios publicados de sitios web especializados.

La principal fuente de información para este análisis está compuesta por las páginas de internet de oferta residencial. Los sitios web que presentan oferta para los tres países son *Century21* y *Point2*. En particular para los mercados de vivienda en Estados Unidos y Canadá se encuentran los sitios web *Zillow*, *Realtor*, *RE/MAX* y *Redfin*, mientras que las páginas especializadas en México son *Metros Cúbicos* y *Vivanuncios*. La extracción y procesamiento de datos de estas fuentes se generó con los métodos de *web scraping*, descritos en las secciones anteriores.

La primera fase del estudio se realizó a partir de los criterios definidos por Chiu, Hsieh y Kao (2005) para la evaluación de la calidad y cantidad de la información en páginas de internet a partir de un índice. La calidad de las páginas de internet se evaluó con respecto a cinco factores: la conectividad, calidad de la información, la interactividad, la experiencia y el aprendizaje. La formalización del índice se expresa como sigue:

$$P_i = f(X_{ij}) + \mu_i \quad (9)$$

Dentro de la clasificación del índice, la oferta online del mercado de vivienda, se ubican dentro de la categoría de bienes de búsqueda, al igual que el resto de productos inmobiliarios. Una vez aplicada la evaluación de las páginas de internet especializadas, se estimó el papel de dicha calidad en la reducción de las fricciones de búsqueda de información. En cuanto a los signos esperados, se tiene el supuesto de relación negativa entre el índice y la dispersión de los precios de oferta.

El análisis de la información publicada en sitios especializados y la formulación de índices, permite poner en contexto a las conurbaciones fronterizas entre México, Estados Unidos y Canadá, mismas que presentan diferencias en la calidad y cantidad de información disponible. La dispersión de precios es una medida de la ignorancia en el

mercado, con una distribución de frecuencias de los precios cotizados por los vendedores. Un comprador comprará el bien si está conforme con el precio del primer vendedor consultado, de lo contrario continuará la búsqueda (Stigler, 1961). Adicionalmente, la perspectiva de la información imperfecta está presente en toda la investigación, al destacar el efecto de la selección adversa y los costos de búsqueda en los cuales incurren los agentes participantes en los mercados de vivienda estudiados.

## IV. MODELO DE PRECIOS HEDÓNICOS

---

La valoración del precio de las viviendas por cualidades a partir del modelo de precios hedónicos, es una técnica que permite encontrar los valores implícitos de un bien complejo, en el sentido de que está compuesto por distintos atributos. Se trata de una estimación de la valoración de no mercado de las viviendas a partir de sus características. En esta investigación, el modelo de precios hedónicos constituye el paso previo a la identificación de los submercados de vivienda. A partir del modelo base, se identificó la estructura de las relaciones entre las variables y se especificó el modelo en cuanto a su forma funcional y variables a incluir. Asimismo, se aplicaron las pruebas estadísticas de significancia y especificación y se realizaron los ajustes necesarios en modelos sucesivos.

En este apartado se presentan dos secciones, en primer lugar, el preprocesamiento de las bases de datos para su consolidación, seguido del análisis exploratorio de datos –EDA– y el análisis exploratorio espacial –ESDA–. En segundo lugar, el modelo de precios hedónicos base estimado por mínimos cuadrados ordinarios –OLS–, el modelo semilogarítmico y semilog con ruido blanco –HC– con errores estándar robustos.

### 4.1. Preprocesamiento y análisis exploratorio

#### 4.1.1. Preparación de datos: recopilación, limpieza y transformación

El trabajo de preprocesamiento, en particular cuando no se cuenta con conjuntos de datos estructurados, representa una etapa fundamental para la consolidación de las fuentes de información previa al análisis. Esta labor toma mayor relevancia en casos en los cuales, por la naturaleza o novedad del tipo de estudio, no se cuenta con bases de datos prediseñadas y se requiere recopilar la información, misma que en algunos casos es de difícil acceso o restringido, se encuentra dispersa y presenta diferentes estructuras y formatos. Este es el caso del presente estudio, en el cual se contó con dos tipos de fuentes, la primera proveniente de los repositorios de las agencias estadísticas de cada país, y el segundo, creado expresamente para esta investigación.

Este segundo conjuntos de datos los conforma la oferta de vivienda publicada en páginas web, información que fue extraída por medio de técnicas de *web scraping* durante

el periodo de septiembre a noviembre de 2021. Las bases de datos se integraron a partir de la fuentes de la empresas especializadas en el mercado de bienes raíces con oferta en el área de estudio: *Century21*, *Century21 Canada*, *Century21 Global*, *Realtor*, *Point2Homes* y *Vivanuncios*. Al tratarse de sitios de internet con diferente estructura en su protocolo de transferencia de hipertexto –*http*–, algunas de ellas con aplicaciones web dinámicas, existen diferencias entre la información mostrada. Algunos sitios tienen un registro más detallado de las características en categorías definidas de los inmuebles expresadas en forma numérica, mientras que otros tienen menos campos definidos y la mayor parte de la información se encuentra en la descripción con un formato de cadena o texto.

Por tanto, una vez realizada la extracción de la información bajo un diseño para datos no estructurados, se inició con el primer proceso de limpieza de datos, en el cual se eliminaron los registros repetidos, los que presentaban un gran número de columnas faltantes y fuera del área de estudio. El segundo proceso consistió en la manipulación de los datos con el objetivo de darles una estructura de filas que concentraran toda la información de cada vivienda u observación, y columnas que incluyeran por separado la mayor cantidad de características intrínsecas sobre las mismas. En este paso también se extrajo información del campo de descripción, con el fin de completar los registros con campos faltantes y en la medida de lo posible, evitar perder información, bajo el principio de integridad de la información.

**Tabla 8 Descripción de variables**

Clave	Descripción	Tipo <sup>a</sup>	Signo esperado
<i>Variable dependiente</i>			
USD_PRICE	Precio de lista ofrecido a la venta en dólares estadounidenses.	NC	
<i>Variables independientes</i>			
<i>Localización</i>			
LAT	Latitud en la cual se localiza la propiedad, eje vertical.		+/-
LONG	Longitud en la cual se localiza la propiedad, eje horizontal.		+/-
<i>Atributos del inmueble</i>			
SQM	Extensión en metros cuadrados.	NC	+
ROOMS	Número de habitaciones separadas.	ND	+
BATHS	Número de baños, incluyendo medios baños.	NC	+
CARS	Cantidad de plazas de estacionamiento.	ND	+

<sup>a)</sup> Claves de tipo de variable: NC – Numérico continuo      ND – Numérico discreto

Fuente: Elaboración propia.

Como se mencionó, la estructura de cada fuente de información es diferente, por lo tanto se requirió homologarla. Esta fase incluyó la conversión de unidades de medida y moneda, así como la integración de campos con información dispersa sobre las mismas características. De esta forma, todos los conjuntos de datos se consolidaron con las variables: número de registro, precio en dólares estadounidenses, metros cuadrados, número de habitaciones, número de baños incluyendo medios baños, latitud y longitud para todas las ciudades y plazas de estacionamiento para las ciudades mexicanas.

Cabe resaltar que se extrajo información sobre otras características de las viviendas, tales como amenidades externas las cuales incluyen caseta de seguridad, estacionamiento para visitas, albercas, área de juegos infantiles, terrazas, asadores, entre otras áreas comunes; así como atributos internos como aire acondicionado y calefacción, cuarto de servicio, cocina integral, balcones, patios y jardines, entre otros. No obstante, debido a que al menos la mitad de las observaciones no contaban con dicha información, se optó por no incluir estas características en el estudio. Finalmente, se procedió a la revisión de los valores de dichas observaciones a través del análisis exploratorio de datos.

#### *4.1.2. Análisis exploratorio de datos (EDA)*

En primer lugar, se realizó el análisis exploratorio de datos convencional y espacial, por medio de los cuales se identificó la distribución de las variables y la incidencia de datos atípicos. Posteriormente se realizaron gráficas de dispersión y se calcularon los coeficientes de correlación para las variables de interés.

A continuación se muestra los estadísticos de resumen por pares de ciudades para la franja fronteriza México-Estados Unidos seguidas de las comunidades entre Estados Unidos-Canadá. Asimismo, las gráficas de dispersión entre precios en dólares estadounidenses y metros cuadrados de las viviendas en orden antes señalado.

**Tabla 9 Estadística descriptiva, ciudades fronterizas México-Estados Unidos**

Clave	MX01 Tijuana			US01 San Diego		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	211426.9	179551.1	159122.6	800384.4	725000.0	417008.6
SQM	186.8189	125.0000	231.9752	144.5512	133.8268	57.35555
ROOMS	2.559592	3.000000	0.798951	2.910737	3.000000	0.950438
BATHS	2.208537	2.000000	0.860928	2.134831	2.000000	0.660464
LAT	32.45522	32.48492	0.072419	32.86204	32.79809	0.196031
LONG	-117.013	-117.0298	0.080527	-117.0815	-117.0856	0.148055
CARS	1.885287	2.000000	0.729558			
<i>Número total de observaciones</i>			7157			1602

Clave	MX02 Mexicali			US02 Imperial		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	135633.0	99993.77	126081.3	316847.9	281500.0	155392.3
SQM	225.6283	160.0000	202.3426	164.3591	140.3300	85.86054
ROOMS	2.798961	3.000000	0.721151	3.875000	3.000000	2.089742
BATHS	2.222181	2.000000	0.956052	2.385417	2.000000	1.345756
LAT	32.57531	32.61914	0.261388	32.85753	32.79748	0.202222
LONG	-115.4104	-115.4224	0.114444	-115.5403	-115.5465	0.199778
CARS	2.076409	2.000000	0.763461			
<i>Número total de observaciones</i>			1348			48

Clave	MX03 Nogales			US03 Nogales		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	81724.56	30199.50	96344.29	363814.7	316250.0	167845.5
SQM	138.8433	70.00000	138.8334	201.2206	174.0538	89.34934
ROOMS	2.694030	2.000000	1.012906	2.988636	3.000000	0.702938
BATHS	1.854478	1.500000	1.070494	2.329545	2.000000	0.605940
LAT	31.29046	31.29270	0.022638	31.51283	31.53902	0.093951
LONG	-110.9485	-110.9552	0.018684	-110.9404	-110.9869	0.136484
CARS	1.305970	1.000000	0.768042			
<i>Número total de observaciones</i>			134			88

Clave	MX04 Ciudad Juárez			US04 El Paso		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	145072.1	94264.34	159137.2	255141.5	244450.0	98489.27
SQM	230.7808	152.5000	208.1073	184.3594	175.7726	57.45238
ROOMS	3.412607	3.000000	1.334499	3.656479	4.000000	0.729238
BATHS	2.342765	2.000000	1.056550	2.371638	2.500000	0.619474
LAT	31.67786	31.69036	0.046738	31.78488	31.78253	0.083811
LONG	-106.4079	-106.4079	0.034996	-106.3451	-106.2954	0.131620
CARS	2.007880	2.000000	0.835122			
<i>Número total de observaciones</i>			1396			818

Clave	MX05 Piedras Negras			US05 Eagle Pass		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	107943.9	88778.05	79419.67	211232.7	196750.0	99111.22
SQM	307.9184	160.0000	392.4934	169.4129	148.5055	58.46268
ROOMS	3.061224	3.000000	1.048971	3.340909	3.000000	0.584992
BATHS	2.346939	2.000000	0.995846	2.164773	2.000000	0.623863
LAT	28.68261	28.68585	0.023727	28.70904	28.71025	0.073603
LONG	-100.5487	-100.5452	0.020385	-100.4744	-100.4865	0.052070
CARS	1.122449	1.000000	0.331201			
<i>Número total de observaciones</i>			49			

88

Clave	MX06 Nuevo Laredo			US06 Laredo		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	12742.06	3411.471	16670.06	323095.3	314121.5	82180.55
SQM	92.58451	90.00000	25.09450	220.5115	218.4150	53.43956
ROOMS	1.338028	1.000000	0.569774	3.933333	4.000000	0.583292
BATHS	1.098592	1.000000	0.353500	2.616667	2.000000	0.784366
LAT	27.49343	27.50894	0.029193	27.55224	27.57835	0.074459
LONG	-99.57363	-99.57843	0.021407	-99.47581	-99.45655	0.041311
CARS	1.007042	1.000000	0.083918			
<i>Número total de observaciones</i>			142			

30

Clave	MX07 Reynosa			US07 McAllen		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	51214.36	27930.17	54941.69	288423.2	239000.0	157074.5
SQM	126.1867	91.00000	109.0828	211.0811	180.6035	98.40570
ROOMS	2.455614	2.000000	0.664694	3.485043	3.000000	0.791950
BATHS	1.593995	1.000000	0.828206	2.681624	2.500000	0.964645
LAT	26.05723	26.06782	0.031784	26.23496	26.23640	0.062760
LONG	-98.31268	-98.33269	0.073110	-98.20221	-98.221	0.117678
CARS	1.091384	1.000000	0.288342			
<i>Número total de observaciones</i>			766			

468

Clave	MX08 Matamoros			US08 Brownsville		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	59826.70	35910.22	61613.46	713504.2	355000.0	759079.9
SQM	197.5113	120.0000	282.1112	204.2695	187.5712	95.03349
ROOMS	2.334842	2.000000	0.760448	3.280632	3.000000	1.037157
BATHS	1.730769	1.000000	0.924983	2.911067	2.500000	1.131702
LAT	25.84765	25.85138	0.031130	26.09953	26.11320	0.074803
LONG	-97.51442	-97.51112	0.055387	-97.35886	-97.2961	0.212530
CARS	1.140271	1.000000	0.385248			
<i>Número total de observaciones</i>			221			

253

**Tabla 10 Estadística descriptiva, ciudades fronterizas Estados Unidos-Canadá**

Clave	US09 Whatcom			CA09 Metro Vancouver		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	664059.5	549500.0	329802.0	1155537.	959200.0	837171.6
SQM	174.3376	151.8500	63.63972	184.1837	133.9662	123.4683
ROOMS	3.023810	3.000000	0.680319	3.453831	3.000000	1.891842
BATHS	2.107143	2.000000	0.568864	2.866405	2.000000	1.603647
LAT	48.82696	48.81101	0.087375	49.16732	49.16283	0.084963
LONG	-122.5446	-122.4968	0.201026	-122.9315	-122.9123	0.156586
<i>Número total de observaciones</i>			42			509

Clave	US10 Sault St. Marie			CA10 Sault Saint Marie		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	172277.1	181200.0	59416.97	220791.0	180647.0	112859.1
SQM	139.6819	136.1959	40.92947	131.4619	120.7740	42.90105
ROOMS	3.028571	3.000000	0.837316	3.074074	3.000000	0.780824
BATHS	1.604762	2.000000	0.544621	1.666667	2.000000	0.620174
LAT	45.37462	45.01163	0.636267	46.52462	46.52280	0.013612
LONG	-89.57715	-91.2629	3.018175	-84.33367	-84.33873	0.029635
<i>Número total de observaciones</i>			105			27

Clave	US11 Detroit			CA11 Windsor		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	215233.9	165000.0	180512.7	711898.6	540686.5	423499.5
SQM	140.1920	118.4514	65.09701	215.6341	188.3145	95.45503
ROOMS	3.107896	3.000000	0.809208	3.760000	4.000000	1.134973
BATHS	1.797770	1.500000	0.824424	2.760000	2.000000	1.079682
LAT	42.34018	42.35570	0.083200	42.24074	42.26130	0.084531
LONG	-83.23177	-83.2427	0.172332	-83.02516	-83.03019	0.068113
<i>Número total de observaciones</i>			1659			50

Clave	US12 Port Huron			CA12 Point Edward		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	321062.8	325000.0	160920.8	456990.4	461375.5	155063.8
SQM	182.4292	179.5816	74.11380	150.4140	142.2810	64.19174
ROOMS	3.023256	3.000000	0.830616	3.400000	3.000000	0.753937
BATHS	2.523256	2.500000	1.112448	2.350000	2.000000	0.812728
LAT	42.83735	42.84030	0.015984	42.82197	42.95099	0.260573
LONG	-82.48928	-82.4865	0.018622	-81.88363	-81.79972	0.276312
<i>Número total de observaciones</i>			43			20

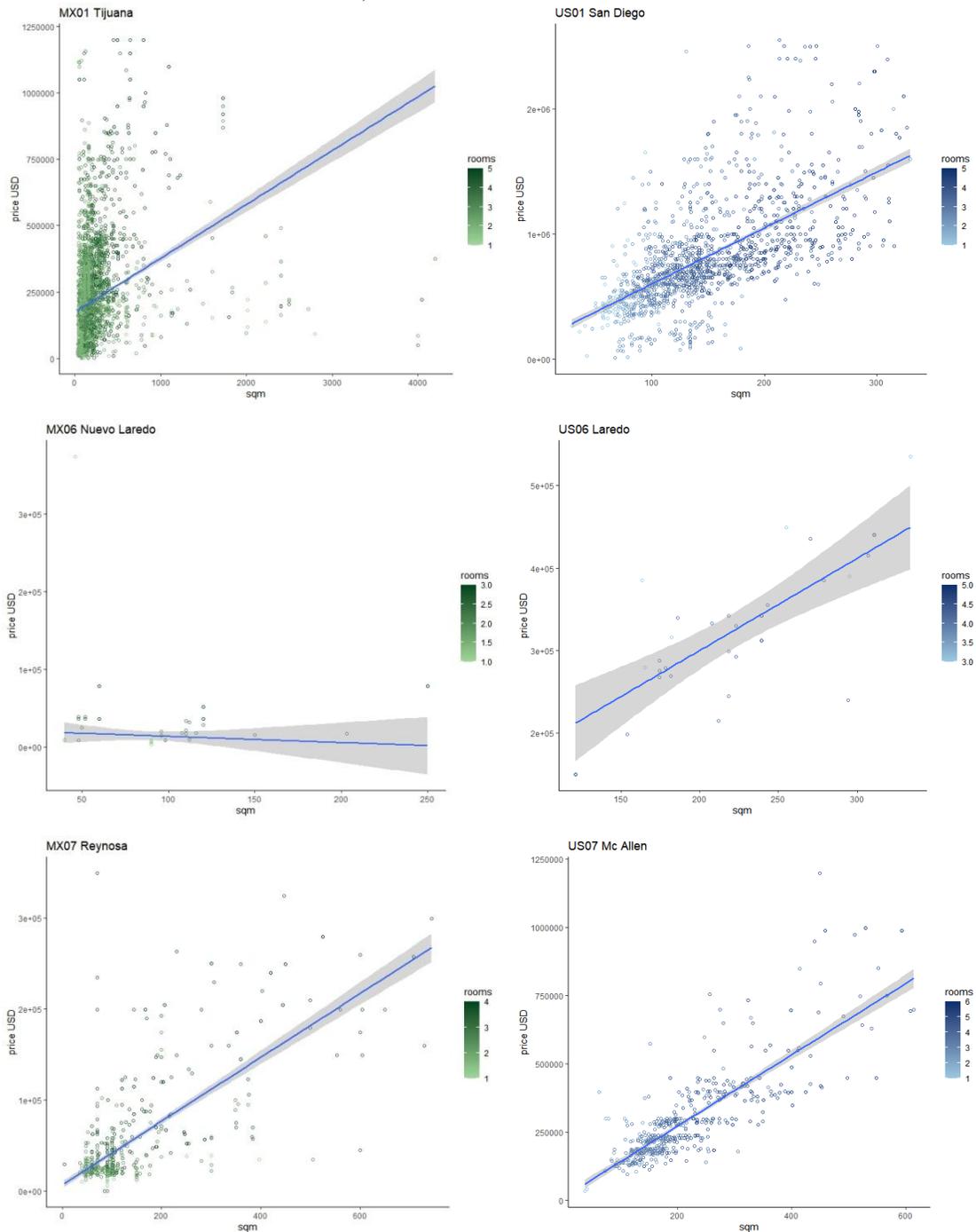
Clave	US13 Buffalo-Niagara Falls			CA13 Great Golden Horseshoe		
	Media	Mediana	Desv. estándar	Media	Mediana	Desv. estándar
USD_PRICE	234634.9	179900.0	153137.3	1335685.	802429.0	1206165.
SQM	137.7910	127.2539	43.12036	246.6725	185.8061	170.6450
ROOMS	3.269384	3.000000	0.886604	3.785714	4.000000	1.201730
BATHS	1.759940	1.500000	0.686961	3.196429	3.000000	1.577335
LAT	42.95015	42.95515	0.150282	43.25811	43.24378	0.181514
LONG	-78.80386	-78.81233	0.118319	-79.56042	-79.6967	0.357584
<i>Número total de observaciones</i>			1006			

Como se observa, la mayoría de las variables presenta asimetría positiva con una distribución que se aproxima a la normal. Entre los pares de ciudades se aprecian contrastes notables, a continuación se mencionan los casos más extremos. En primer lugar, destacan las diferencias de precios de lista, en la frontera México-Estados Unidos la mediana de San Diego es \$545,449 USD mayor a la de Tijuana; en la frontera con Canadá la mediana del área conocida como Great Golden Horseshoe está \$622,529 USD por encima de la zona metropolitana de Buffalo-Niagara Falls.

En cuanto a metros cuadrados de las viviendas, la mediana de Laredo es 128.42 m<sup>2</sup> superior a la de su ciudad gemela Nuevo Laredo; mientras que entre Estados Unidos y Canadá la mayor diferencia es de 69.86 m<sup>2</sup> más en Windsor que en la ciudad de Detroit. Respecto al número de habitaciones, las medianas más altas son para dos ciudades estadounidenses y dos canadienses respectivamente: El Paso, Laredo, Windsor y Great Golden Horseshoe con 4 recámaras, la mediana más pequeña es de 1 sola habitación en la ciudad mexicana de Nuevo Laredo. La media y mediana de número de cuartos de baño es más homogéneo a lo largo de toda la muestra.

Para las ciudades mexicanas se incluye el número de espacios de estacionamiento, ya que la fuente de la cual se extrajeron los datos contaba con dicho campo de información en formato numérico. Para Estados Unidos y Canadá, la estructura de las fuentes consultadas no tenía un espacio destinado a esta característica, en algunos casos se mencionaba en la descripción o sólo se indicaba la presencia del atributo sin especificar su número. Es por ello que, con el fin de evitar contar con un número considerable de registros no válidos o sin información para las ciudades estadounidenses y canadienses, no se incluyó la variable del número de plazas de automóviles para dichos países.

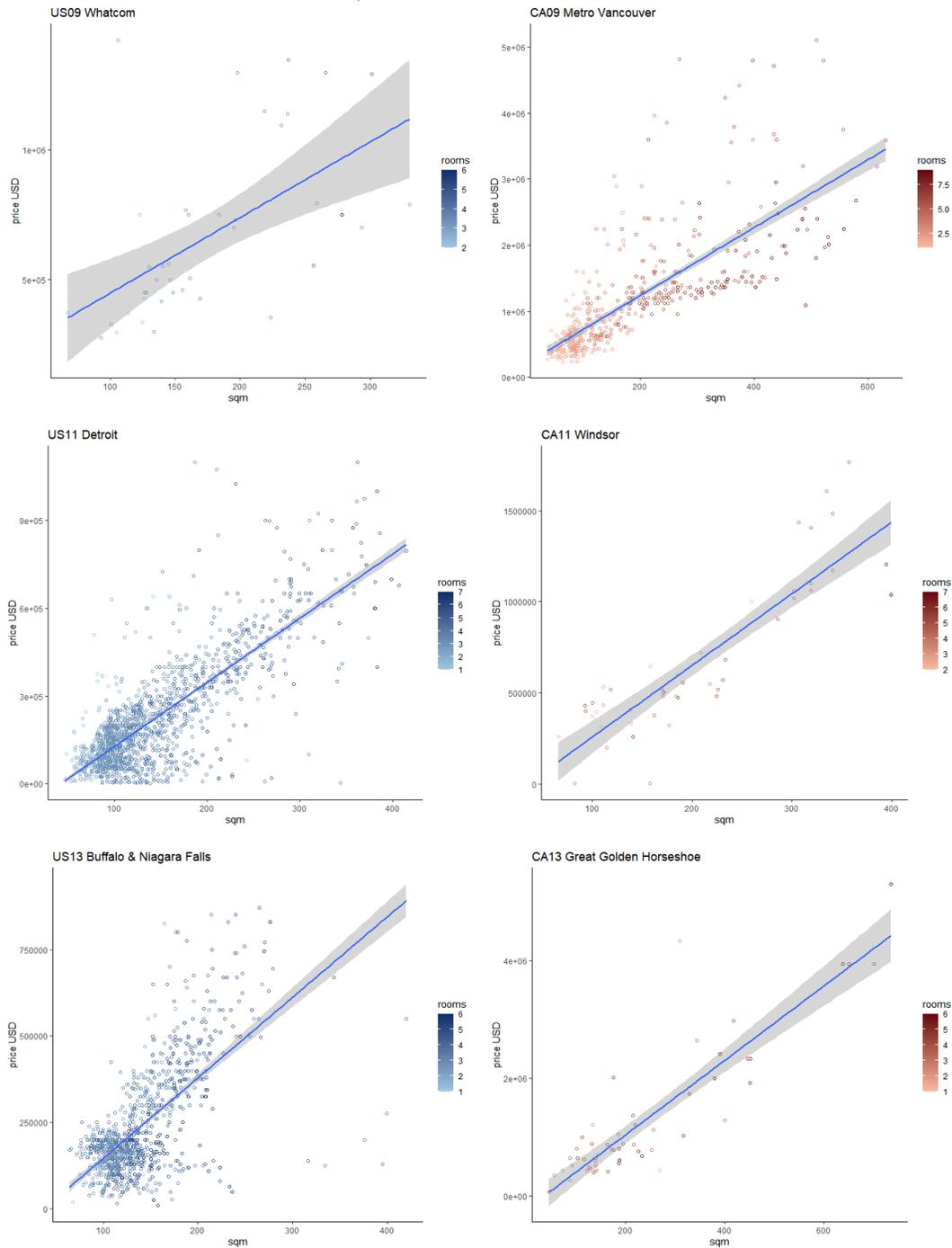
**Gráfica 7. Dispersión de precios en dólares y metros cuadrados de las viviendas de la frontera México-Estados Unidos, ciudades seleccionadas**



A partir de la revisión de las gráficas de dispersión que aquí se presentan, en conjunto con las gráficas de correlación se aprecia en qué casos la variable de metros cuadrados se encuentra mayor o menormente relacionada con la variable dependiente de

precios hedónicos o de no mercado. Para las ciudades mexicanas, la relación más clara se aprecia para Reynosa, y la menos evidente para Nuevo Laredo. Las gráficas de dispersión para las ciudades estadounidenses son más consistentes.

**Gráfica 8. Dispersión de precios en dólares y metros cuadrados de las viviendas de la frontera Estados Unidos-Canadá, ciudades seleccionadas**

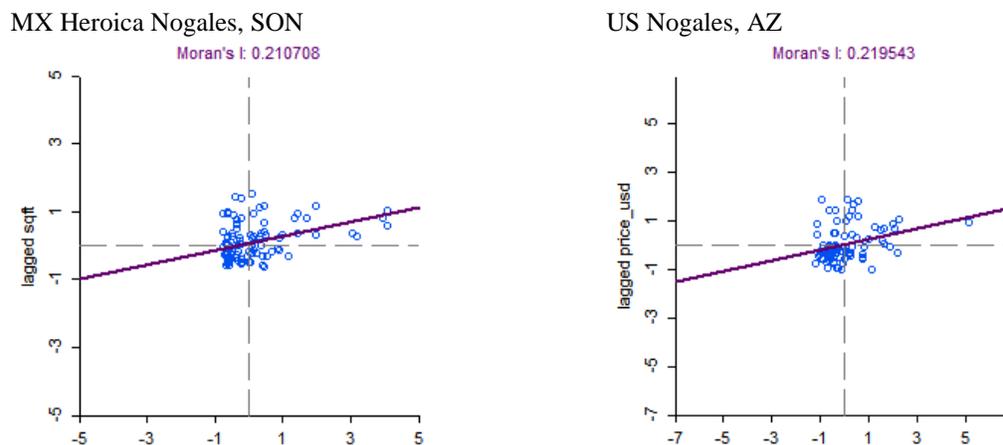


Además de las variables de precio de lista y metros cuadrados, a partir de tablas de contingencias se examinó la relación entre otras variables de interés presentes en los conjuntos de datos, con el fin de identificar las características más comunes. Destacan el número de habitaciones y baños, en varias de las ciudades estudiadas, la mayoría de las viviendas ofrecidas son de 3 recámaras y 2 cuartos de baño, por ejemplo en Tijuana (21.09%), San Diego (23.91%) y Whatcom (30.95). Destaca el caso de Windsor (28.0%) con cuatro habitaciones e igual número de baños. Una vez realizado el análisis exploratorio convencional, se exploraron los datos en su vertiente espacial, donde se calculó el Índice de Moran para capturar la asociación espacial entre las variables de pies cuadrados y el precio rezagado.

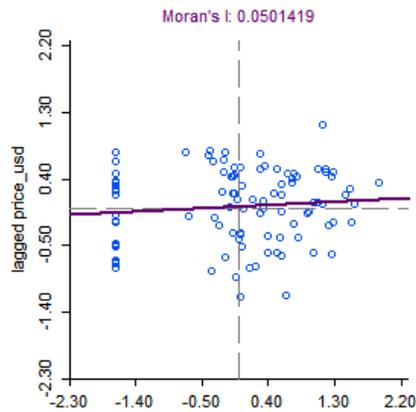
#### 4.1.3. Análisis exploratorio de datos espaciales (ESDA)

Una vez efectuado el análisis exploratorio convencional, se realizó el análisis de datos en su vertiente espacial, con el objetivo de identificar las características espaciales de las variables de estudio, así como de aplicar las pruebas que permiten probar la existencia de asociación espacial entre las mismas. A continuación se presenta una muestra de los Índices de Moran bivariados de asociación espacial entre la variable metros cuadrados con el rezago de la valoración del precio.

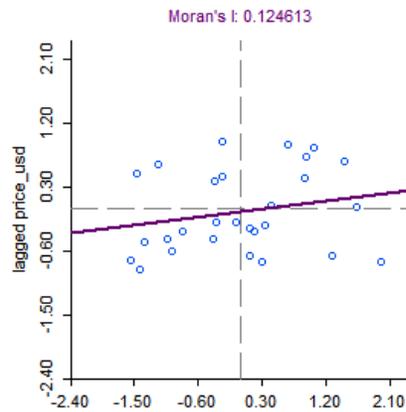
**Gráfica 9. Índice de Moran para las ciudades fronterizas entre México, Estados Unidos y Canadá, ciudades seleccionadas**



US Sault St. Marie, MI



CA Sault Saint Marie, ONT



#### 4.4. Modelo de precios hedónicos

##### 4.4.1. Modelo base

Con relación a la valoración de no mercado de las cualidades de las viviendas, se realizó la estimación del modelo base de forma lineal incluyendo únicamente las características intrínsecas de la vivienda y la ubicación como variables independientes, y la valoración del precio expresado en dólares estadounidenses como la variable independiente. Posteriormente se evaluó la forma funcional semilogarítmica donde la variable dependiente se transforma en logaritmo y la parte derecha de la ecuación se mantiene lineal. Dicha transformación responde tanto a las características de los datos como a la convención en la bibliografía académica, la cual sugiere esta especificación para modelos de valoración de precios.

Seguidamente, se estimó el modelo semilog con ruido blanco –Huber-White, HC– el cual es un modelo que brinda errores estándar robustos adecuando cambios en el cálculo de la covarianza ante la violación del supuesto de homocedasticidad. A continuación se presenta el resumen de los resultados de la estimación del modelo de precios hedónicos para cada una de las veintiséis ciudades del área de estudio, los cuales incluyen los resultados del modelo base de mínimos cuadrados ordinarios, modelo semilogarítmico y semilog con ruido blanco (HC).

**Tabla 11 Resumen de resultados modelos OLS, semilog y semilog con ruido blanco HC para las ciudades fronterizas México-Estados Unidos**

Modelo	MX01 Tijuana			US01 San Diego		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	<sup>a)</sup> HC Semilog
C	-47456281*** (2544067)	-367.1062*** (13.92052)	-367.1062*** (20.49400)	-1.10E+08*** (5765892)	-152.0360*** (8.140213)	-152.0360*** (13.24606)
SQM	137.6646*** (6.443176)	0.000676*** (0.000035)	0.000676*** (0.000058)	4731.966*** (221.3586)	0.005165*** (0.000313)	0.005165*** (0.000309)
ROOMS	-12059.63*** (2325.413)	-0.088596*** (0.012724)	-0.088596*** (0.012918)	-8981.823 (11278.89)	0.074133*** (0.015923)	0.074133*** (0.016093)
BATHS	79854.91*** (2281.324)	0.438558*** (0.012483)	0.438558*** (0.013223)	22410.22 (17187)	0.016309 (0.024264)	0.016309 (0.021770)
LAT	55379.85*** (20917.79)	0.043357 (0.114457)	0.043357 (0.117464)	-380976.6*** (38738.59)	-0.476493*** (0.054691)	-0.476493*** (0.071232)
LONG	-389976.4*** (19474.07)	-3.216725*** (0.106557)	-3.216725*** (0.159535)	-1043177*** (50560.69)	-1.538682*** (0.071381)	-1.538682*** (0.123962)
CARS	35437.04*** (2235.751)	0.20159*** (0.012233)	0.20159*** (0.013988)			
R <sup>2</sup> Ajustada	0.398751	0.430698	0.430698	0.515818	0.487558	0.487558
<sup>b)</sup> Estadístico F	791.9831***	903.2972***	903.2972***	342.1222***	305.6515***	305.6515***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	708.0358***	183.3530***		135.3214***	70.77086***	

Notas generales: Errores estándar entre paréntesis. \*, \*\* y \*\*\* representan el nivel de significancia  $p < 10, 5$  y  $1$  por ciento respectivamente.

<sup>a)</sup>Modelo HC White con errores estándar robustos a la heterocedasticidad.

<sup>b)</sup>Estadístico de F y probabilidad expresada en asteriscos.

<sup>c)</sup>Prueba de heterocedasticidad de Breusch-Pagan-Godfrey muestra las observaciones por R<sup>2</sup>, probabilidad de X<sup>2</sup> expresada con asteriscos.

Modelos Semilog y HC Semilog obtienen los mismos coeficientes, R<sup>2</sup> y F. Cambian los errores estándar, estadístico de t y p-valores.

Modelo	MX02 Mexicali			US02 Imperial		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	<sup>a)</sup> HC Semilog
C	15929642*** (4114527)	191.1504*** (26.21697)	191.1504*** (45.27414)	-8699606 (8321128)	2.734438 (36.29759)	2.734438 (29.27874)
SQM	153.5854*** (-14.49207)	0.000683*** (0.000092)	0.000683*** (0.000134)	1742.443*** (385.5373)	0.005798*** (0.001682)	0.005798*** (0.001578)
ROOMS	8382.538* (-4584.372)	0.204926*** (0.029211)	0.204926*** (0.034551)	-6969.169 (14177.91)	-0.009982 (0.061845)	-0.009982 (0.068358)
BATHS	58612.91*** (-3655.640)	0.441287*** (0.023293)	0.441287*** (0.026713)	-7090.197 (18466.66)	-0.052021 (0.080553)	-0.052021 (0.050952)
LAT	67325.02*** (-17244.35)	0.683865*** (0.109878)	0.683865*** (0.174439)	13845.14 (84259.34)	-0.078193 (0.367548)	-0.078193 (0.445730)
LONG	157630.6*** (-39488.82)	1.765225*** (0.251615)	1.765225*** (0.436303)	-72001.48 (84994.17)	-0.100190 (0.370753)	-0.100190 (0.306887)
CARS	8041.769** (-3983.992)	0.021683 (0.025385)	0.021683 (0.026725)			
R <sup>2</sup> Ajustada	0.420248	0.506407	0.506407	0.688665	0.537347	0.537347
<sup>b)</sup> Estadístico F	163.7345***	231.3285***	231.3285***	21.79257***	11.91761***	11.91761***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	118.1761***	21.64314***		5.79334	2.90164	

Modelo	MX03 Nogales			US03 Nogales		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	<sup>a)</sup> HC Semilog
C	-43817168 (29909111)	-477.5604 (290.0954)	-477.5604 (319.4586)	-12092020 (8790808)	-53.72205** (23.37728)	-53.72205** (22.25350)
SQM	312.2605*** (56.62231)	0.002275*** (0.000549)	0.002275*** (0.000677)	990.6365*** (150.5515)	0.002643*** (0.000400)	0.002643*** (0.000560)
ROOMS	8786.695 (7219.365)	0.177988** (0.070022)	0.177988** (0.086388)	-9918.010 (15243.48)	0.010468 (0.040537)	0.010468 (0.039864)
BATHS	9886.335 (6604.869)	0.238199*** (0.064062)	0.238199*** (0.067282)	48722.04** (23706.3)	0.075317 (0.063042)	0.075317 (0.076456)
LAT	-546554.2** (228535.5)	-5.509319** (2.216619)	-5.509319 (3.337841)	720732.9*** (107174.5)	2.245275*** (0.285008)	2.245275*** (0.350396)
LONG	-548541.2** (274416.4)	-5.940673** (2.661628)	-5.940673** (2.837648)	95003.26 (69738.8)	0.045646 (0.185455)	0.045646 (0.193832)
CARS	42602.62*** (8018.644)	0.30475*** (0.077775)	0.30475*** (0.104422)			
R <sup>2</sup> Ajustada	0.653123	0.669843	0.669843	0.760308	0.752732	0.752732
<sup>b)</sup> Estadístico F	42.7368***	45.9732***	45.9732***	56.19308***	53.96889***	53.96889***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	28.94248***	2.695118		20.61114***	15.88885***	

Modelo	MX04 Ciudad Juárez			US04 El Paso		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	<sup>a)</sup> HC Semilog
C	26362016*** (7373402)	370.3398*** (45.96967)	370.3398*** (49.51459)	-3183742* (1698403)	14.23835** (6.443905)	14.23835** (6.689671)
SQM	426.3914*** (14.002600)	0.001526*** (0.000087)	0.001526*** (0.000119)	1123.928*** (51.19405)	0.004032*** (0.000194)	0.004032*** (0.000273)
ROOMS	-5236.082*** (1793.942)	-0.06772*** (0.011184)	-0.06772*** (0.013773)	4416.609 (3508.485)	0.049911*** (0.013312)	0.049911*** (0.017193)
BATHS	47849.33*** (2960.196)	0.412038*** (0.018455)	0.412038*** (0.021926)	29452.88*** (4680.630)	0.114552*** (0.017759)	0.114552*** (0.022443)
LAT	292323.8*** (58430.26)	7.390916*** (0.364285)	7.390916*** (0.430945)	-38497.60 (31652.29)	0.094525 (0.120092)	0.094525 (0.125965)
LONG	335410.8*** (73467.99)	5.586111*** (0.458038)	5.586111*** (0.496623)	-41086.23** (20342.91)	0.056986 (0.077183)	0.056986 (0.079707)
CARS	10247.91*** (3264.574)	0.109121*** (0.020353)	0.109121*** (0.020177)			
R <sup>2</sup> Ajustada	0.695500	0.732217	0.732217	0.675172	0.676825	0.676825
<sup>b)</sup> Estadístico F	532.0465***	636.7416***	636.7416***	340.6348***	343.2083***	343.2083***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	470.2820***	308.0966***		123.2732***	33.61339***	

Modelo	MX05 Piedras Negras			US05 Eagle Pass		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	<sup>a)</sup> HC Semilog
C	-42200143 (68684048)	-300.8363 (657.0518)	-300.8363 (560.3626)	58157431* (29315361)	194.0915 (139.9947)	194.0915 (140.5884)
SQM	31.43014 (34.64946)	0.000260 (0.000331)	0.000260 (0.000314)	722.6268*** (191.3218)	0.002703*** (0.000914)	0.002703*** (0.000864)
ROOMS	-12941.09 (15218.15)	-0.107318 (0.145581)	-0.107318 (0.136017)	-44130.02*** (15312.07)	-0.220082*** (0.073122)	-0.220082*** (0.065436)
BATHS	39554.46*** (14435.38)	0.469703*** (0.138093)	0.469703*** (0.096089)	53759.88*** (19113.69)	0.291608*** (0.091277)	0.291608*** (0.099477)
LAT	309189.3 (540561.1)	4.481762 (5.171166)	4.481762 (4.169429)	587639.0** (251047.4)	1.966391 (1.198870)	1.966391 (1.395274)
LONG	-331953 (613061.6)	-1.816644 (5.864727)	-1.816644 (5.228287)	745544.6** (354897.0)	2.376118 (1.694801)	2.376118 (1.753484)
CARS	-527.1893 (35245.70)	0.058740 (0.337171)	0.058740 (0.271104)			
R <sup>2</sup> Ajustada	0.177268	0.2753710	0.2753710	0.440490	0.393462	0.393463
<sup>b)</sup> Estadístico F	2.723705**	4.040129***	4.040129***	14.69864***	12.28742***	12.28742***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	0.579677	6.733717		13.21114**	4.105457	

Modelo	MX06 Nuevo Laredo			US06 Laredo		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	<sup>a)</sup> HC Semilog
C	-2363033 (2728435)	-551.6434*** (180.3206)	-551.6434** (257.8582)	39331840 (26399720)	174.7906** (83.63356)	174.7906 (169.4122)
SQM	-24.61710 (24.38164)	-0.003023* (0.001611)	-0.003023** (0.001429)	841.2706*** (262.1680)	0.002299** (0.000831)	0.002299** (0.000873)
ROOMS	23106.57*** (1343.681)	1.396349*** (0.088803)	1.396349*** (0.090701)	-34780.98* (18073.29)	-0.103208* (0.057256)	-0.103208* (0.056098)
BATHS	2979.349 (2057.284)	0.192119 (0.135965)	0.192119* (0.109112)	-1094.071 (15387.59)	-0.010583 (0.048747)	-0.010583 (0.039026)
LAT	-1084.305 (27530.17)	-5.907522*** (1.819451)	-5.907522** (2.593840)	287603.4 (180442.2)	1.330159** (0.571636)	1.330159*** (0.475027)
LONG	-23222.64 (30720.13)	-7.240564*** (2.030274)	-7.240564** (2.815303)	472262.6 (287216.8)	1.999074** (0.909895)	1.999074 (1.759740)
CARS	60883.74*** (6953.411)	0.19397 (0.459547)	0.19397 (0.157417)			
R <sup>2</sup> Ajustada	0.864219	0.824279	0.824279	0.595773	0.608097	0.608097
<sup>b)</sup> Estadístico F	150.5725***	111.2346***	111.2346***	9.548379***	9.999579***	9.999579***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	55.07669***	12.33584**		19.76731***	20.90438***	

Modelo	MX07 Reynosa			US07 McAllen		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	<sup>a)</sup> HC Semilog
C	4574237*** (1661804)	54.23794** (21.95385)	54.23794** (24.05913)	-7087795** (3599810)	-23.16445** (10.85392)	-23.16445** (11.95284)
SQM	225.4722*** (13.07847)	0.002368*** (0.000173)	0.002368*** (0.000206)	1171.39*** (70.63804)	0.003098*** (0.000213)	0.003098*** (0.000257)
ROOMS	4613.647* (2404.330)	0.197381*** (0.031763)	0.197381*** (0.032448)	-22935.12*** (7153.654)	-0.021196 (0.021569)	-0.021196 (0.029424)
BATHS	26813.53*** (2076.103)	0.358929*** (0.027427)	0.358929*** (0.032098)	33498.59*** (6888.639)	0.104879*** (0.020770)	0.104879*** (0.023532)
LAT	53035.70 (44618.94)	0.988478* (0.589454)	0.988478* (0.546916)	206628.6*** (65574.03)	0.923030*** (0.197715)	0.923030*** (0.205139)
LONG	60933.12*** (19400.63)	0.721115*** (0.256299)	0.721115*** (0.249918)	-17292.55 (35158.35)	-0.107384 (0.106007)	-0.107384 (0.123751)
CARS	2739.858 (4296.497)	0.048469 (0.056760)	0.048469 (0.066476)			
R <sup>2</sup> Ajustada	0.634544	0.640828	0.640828	0.688595	0.676068	0.676068
<sup>b)</sup> Estadístico F	222.3795***	228.4832***	228.4832***	207.5311***	195.9318***	195.9318***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	50.64933***	27.65049***		62.58616***	15.57765***	

Modelo	MX08 Matamoros			US08 Brownsville		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	<sup>a)</sup> HC Semilog
C	-5094575 (7131421)	-69.25125 (73.81626)	-69.25125 (71.64537)	94277513*** (17895533)	167.6801*** (17.39662)	167.6801*** (17.59820)
SQM	72.32840*** (10.36124)	0.000705*** (0.000107)	0.000705*** (0.000221)	1485.920** (642.7575)	0.003164*** (0.000625)	0.003164*** (0.000772)
ROOMS	12439.47** (4962.657)	0.284839*** (0.051368)	0.284839*** (0.061618)	-30244.40 (46910.08)	-0.029330 (0.045602)	-0.029330 (0.069674)
BATHS	20669.76*** (4472.755)	0.267620*** (0.046297)	0.267620*** (0.053326)	325047.2*** (55320.28)	0.310683*** (0.053778)	0.310683*** (0.062520)
LAT	197472.0* (100851.8)	2.644968** (1.043902)	2.644968** (1.081314)	1537749*** (417979.8)	1.426630*** (0.406327)	1.426630*** (0.396035)
LONG	731.7762 (56917.40)	-0.101724 (0.589143)	-0.101724 (0.512126)	1385072*** (163646.1)	1.985692*** (0.159084)	1.985692*** (0.166204)
CARS	37245.07*** (8046.137)	0.334696*** (0.083284)	0.334696*** (0.090309)			
R <sup>2</sup> Ajustada	0.561428	0.665247	0.665247	0.589688	0.720244	0.720244
<sup>b)</sup> Estadístico F	47.93798***	73.86688***	73.86688***	73.43323***	130.7571***	130.7571***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	49.54934***	44.31706***		32.38617***	1.384871	

**Tabla 12 Resumen de resultados modelos OLS, semilog y semilog con ruido blanco HC para las ciudades fronterizas Estados Unidos-Canadá**

Modelo	US09 Whatcom			CA09 Metro Vancouver		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	HC Semilog
C	21116467 (33820032)	41.05996 (44.97106)	41.05996 (39.14740)	-16800000*** (21205848)	-120.5410*** (15.26730)	-120.5410*** (17.99297)
SQM	2673.975*** (951.6988)	0.004067*** (0.001265)	0.004067*** (0.001269)	6439.754*** (414.7187)	0.003567*** (0.000299)	0.003567*** (0.000303)
ROOMS	37208.25 (83721.62)	0.040030 (0.111326)	0.040030 (0.123607)	-22744.33 (26604.81)	0.096082*** (0.019154)	0.096082*** (0.019105)
BATHS	-43967.46 (99202.20)	-0.034730 (0.131911)	-0.034730 (0.175888)	-72207.38** (30430.54)	-0.049906** (0.021909)	-0.049906** (0.024789)
LAT	-593147.2 (590815.1)	-1.080033 (0.785617)	-1.080033 (0.733844)	873932*** (290151.8)	0.447309** (0.208897)	0.447309** (0.222342)
LONG	-65471.13 (245675.8)	-0.197597 (0.326679)	-0.197597 (0.317328)	-1021054*** (154384.6)	-0.906577*** (0.11115)	-0.906577*** (0.126153)
R <sup>2</sup> Ajustada	0.250766	0.336288	0.336288	0.633256	0.682145	0.682145
<sup>b)</sup> Estadístico F	3.744512***	5.154755***	5.154755***	176.4327***	219.0426***	219.0426***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	4.241301	4.096962		54.46853***	23.75742***	

Notas: Errores estándar entre paréntesis. \*, \*\* y \*\*\* representan el nivel de significancia  $p < 10$ , 5 y 1 % respectivamente.

<sup>a)</sup>Modelo HC White con errores estándar robustos a la heterocedasticidad.

<sup>b)</sup>Estadístico de F y probabilidad expresada en asteriscos.

<sup>c)</sup>Prueba de heterocedasticidad de Breusch-Pagan-Godfrey muestra las observaciones por R<sup>2</sup>, probabilidad de X<sup>2</sup> expresada con asteriscos.

Modelos Semilog y HC Semilog obtienen los mismos coeficientes, R<sup>2</sup> y F. Cambian los errores estándar, estadístico de t y p-valores.

Modelo	US10 Sault St. Marie			CA10 Sault Saint Marie		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	HC Semilog
C	-2641186 (2823889)	-8.338477 (19.26142)	-8.338477 (17.96390)	-94865805* (50990884)	-281.8310 (272.8197)	-281.831 (249.4213)
SQM	647.5659*** (168.3003)	0.004007*** (0.001148)	0.004007*** (0.001013)	740.0635 (441.4693)	0.002865 (0.002362)	0.002865 (0.001703)
ROOMS	-17806.32** (7295.254)	-0.092833* (0.049760)	-0.092833* (0.050089)	-39020.12** (17588.67)	-0.161583 (0.094106)	-0.161583 (0.107424)
BATHS	25725.84** (11003.61)	0.159175** (0.075054)	0.159175** (0.070025)	80372.15*** (26551.19)	0.324899** (0.142059)	0.324899** (0.147546)
LAT	35196.20 (44068.84)	0.242251 (0.300588)	0.242251 (0.280278)	3531789** (1531883)	14.77042* (8.196129)	14.77042** (7.041209)
LONG	-12711.27 (9303.086)	-0.098195 (0.063455)	-0.098195 (0.059196)	822208.3 (600943.4)	4.667116 (3.215265)	4.667116 (2.826862)
R <sup>2</sup> Ajustada	0.320702	0.336683	0.336683	0.671149	0.524482	0.524482
<sup>b)</sup> Estadístico F	10.81984***	11.55756***	11.55756***	11.61265***	6.735442***	6.735442***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	7.383954	9.023126		9.554071**	9.080549	

Modelo	US11 Detroit			CA11 Windsor		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	HC Semilog
C	-12172474*** (2064375)	-26.52994* (14.77054)	-26.52994** (11.20467)	-67956312 (70225712)	33.06587 (426.2762)	33.06587 (319.1151)
SQM	1675.277*** (67.54516)	0.003180*** (0.000483)	0.003180*** (0.000699)	3064.721*** (735.7689)	0.003547 (0.004466)	0.003547 (0.004762)
ROOMS	-32053.06*** (3911.197)	-0.053701* (0.027984)	-0.053701* (0.030657)	-58161.93** (26435.71)	-0.00436 (0.160467)	-0.00436 (0.090142)
BATHS	73145.53*** (4941.108)	0.539434*** (0.035353)	0.539434*** (0.045819)	133489.5** (58374.91)	0.398574 (0.354341)	0.398574 (0.475814)
LAT	58330.75* (30772.19)	-0.996873*** (0.220174)	-0.996873*** (0.143706)	655816.5 (503576.3)	-0.15219 (3.056752)	-0.15219 (1.481187)
LONG	-115956.1*** (15227.09)	-0.953964*** (0.108949)	-0.953964*** (0.105984)	-483653.9 (656550.1)	0.184917 (3.985317)	0.184917 (3.175006)
R <sup>2</sup> Ajustada	0.697195	0.464030	0.464030	0.811780	0.256661	0.256661
<sup>b)</sup> Estadístico F	764.4926***	288.0915***	288.0915***	43.26671***	4.383757***	4.383757***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	170.0724***	143.9791***		12.73533**	3.370734	

Modelo	US12 Port Huron			CA12 Point Edward		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	HC Semilog
C	12549919 (46365819)	0.474042 (219.6070)	0.474042 (147.7720)	-5949353 (10814309)	-11.96908 (30.98486)	-11.96908 (30.98486)
SQM	1322.597*** (278.1554)	0.002919** (0.001317)	0.002919*** (0.000827)	418.2555 (440.6997)	0.000745 (0.001263)	0.000745 (0.001263)
ROOMS	8364.695 (14050.72)	0.070491 (0.066550)	0.070491 (0.047462)	-46311.34 (31618.07)	-0.097368 (0.090591)	-0.097368* (0.090591)
BATHS	28383.32 (17182.98)	0.165043** (0.081385)	0.165043* (0.086592)	155212.5*** (34777.21)	0.390957*** (0.099643)	0.390957*** (0.099643)
LAT	2171250*** (672132.7)	8.286829** (3.183488)	8.286829** (4.045561)	217219.4** (95376.19)	0.624978** (0.273269)	0.624978* (0.273269)
LONG	1279895** (600642.1)	8.286829 (2.844881)	8.286829 (2.519276)	38660.01 (96965.75)	0.030868 (0.277824)	0.030868 (0.277824)
R <sup>2</sup> Ajustada	0.856492	0.735573	0.735573	0.694759	0.618818	0.618818
<sup>b)</sup> Estadístico F	51.13316***	24.36682***	24.36682***	9.649198***	7.168997***	7.168997***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	6.873947	2.052103		1.422242	6.043872	

Modelo	US13 Buffalo Niagara Falls			CA13 Great Golden Horseshoe		
	OLS	Semilog	<sup>a)</sup> HC Semilog	OLS	Semilog	HC Semilog
C	5324437** (2443509)	48.80297*** (11.16005)	48.80297*** (12.23210)	-32849277* (18389036)	-29.39952* (15.33541)	-29.39952** (15.33541)
SQM	1655.795*** (95.47600)	0.004913*** (0.000436)	0.004913*** (0.000732)	4572.608*** (765.9093)	0.002464*** (0.000639)	0.002464*** (0.000639)
ROOMS	-28190.26*** (4094.587)	-0.088822*** (0.018701)	-0.088822*** (0.022372)	-188326.1** (79393.26)	-0.040878 (0.066209)	-0.040878 (0.066209)
BATHS	96083.32*** (5922.705)	0.385988*** (0.027050)	0.385988*** (0.033597)	258941.0*** (95794.23)	0.178243** (0.079887)	0.178243** (0.079887)
LAT	-35512.08* (21479.06)	-0.135896 (0.098100)	-0.135896 (0.105464)	969446.0 (617464.1)	0.469426 (0.514930)	0.469426 (0.514930)
LONG	49104.73* (27973.56)	0.404217*** (0.127761)	0.404217*** (0.139994)	113047.6 (284348.4)	-0.274537 (0.237130)	-0.274537 (0.237130)
R <sup>2</sup> Ajustada	0.556001	0.433121	0.433121	0.829680	0.744848	0.744848
<sup>b)</sup> Estadístico F	252.7041***	154.5728***	154.5728***	54.58413***	33.11161***	33.11161***
<sup>c)</sup> Het. Obs*R <sup>2</sup>	230.9740***	24.29618***		8.304831	7.948886	

Como se observa en los resultados, existen distintos factores significativos para cada par de ciudades así como para cada ciudad en particular, lo cual indica la disímil valoración del precio de los bienes inmuebles residenciales entre mercados de vivienda contiguos así como a lo largo de la franja fronteriza, a la vez que apunta a la presencia de otros factores que influyen en dicha valoración.

Respecto a la especificación del modelo, la elección de la forma funcional responde a la estructura de los datos de cada ciudad, así como aquella que es señalada por la bibliografía académica como la forma más adecuada para los modelos de precios hedónicos de vivienda, la cual es la función semilogarítmica. En cuanto a la elección de las variables, después de estimar individualmente los modelos y aplicar las pruebas de variables redundantes a aquellas que resultaron no significativas individualmente, en su mayoría se obtuvo como resultado la no aceptación de la hipótesis de no significatividad. Este resultado, aunado al obtenido de la prueba F significancia conjunta, permitió especificar el modelo conservando las variables independientes antes señaladas para todas las ciudades. Además se aplicó la prueba del error de especificación de la ecuación de regresión – *Regression Specification Error Test, RESET*–, pruebas de variables omitidas y diagnóstico de estabilidad.

Asimismo, para la mayoría de las ciudades se detectó que la varianza no es consistente, lo cual es un resultado esperado dada la naturaleza transversal de los conjuntos de datos, aunado a su comportamiento espacialmente determinado. Es por ello que ante la presencia de heterocedasticidad, se estimó el modelo HC con errores estándar y covarianza consistentes. Sin embargo, la corrección de la varianza de los errores se realizó después mediante la estimación del modelo espacial con los submercados anidados, con el cual se incorporaron los submercados por condiciones del vecindario y características de la vivienda.

## V. IDENTIFICACIÓN DE SUBMERCADOS

---

Los submercados de vivienda constituyen una categoría de análisis teórico del mercado inmobiliario que a su vez representa un fenómeno para el estudio empírico. La identificación de los submercados enriquece el análisis del mercado residencial, en particular, ayuda a explicar las variaciones en los precios de propiedades contiguas o con atributos similares, producto de la subagrupación simultánea de una misma vivienda en dos submercados de diferente naturaleza: uno por amenidades del vecindario y otro por atributos intrínsecos de la misma. La conjugación de ambos submercados da origen a los submercados anidados, los cuales consideran el entorno urbano, las características de las propiedades, su ubicación geográfica y la interacción de la misma con respecto a las viviendas circundantes. Todo ello corresponde a la identificación de submercados, eje principal de esta investigación.

Los trece pares de ciudades que conforman el área de estudio de esta investigación constituyen ejemplos de pares de localidades contiguas, ciudades gemelas, comunidades hermanas y conurbaciones transnacionales. La diversidad que representan las mismas permite hacer un análisis comparativo intra-frontera, con lo cual es posible destacar las características comunes y disímiles presentes en este espacio, sometido a un contexto multijurisdiccional. De este modo se contribuye en general al estudio de los submercados de vivienda desde una metodología robusta, y en particular al análisis de los bienes raíces residenciales en las fronteras de Norteamérica.

En las secciones subsecuentes de este capítulo se presentan los resultados del estudio de submercados de vivienda, los cuales se componen de dos tipos de submercados. En primer lugar, la identificación de los submercados por características del vecindario a partir de la aplicación del análisis de componentes principales –PCA–. Con esta metodología el objetivo es obtener índices sintéticos de calidad del entorno urbano a partir de los cuales se clasifica a los vecindarios. En segundo lugar, la clasificación de submercados por características intrínsecas de las viviendas mediante el análisis de conglomerados –K-Means y HC–. La aplicación de esta técnica tiene como propósito distinguir grupos de viviendas al interior de las ciudades de estudio las cuales presenten características similares.

## **5.1. Submercados por características del vecindario**

### *5.1.1. Análisis de Componentes Principales (PCA)*

Los submercados de vivienda por características del vecindario identifican los elementos del entorno urbano que diferencian a los espacios dentro de una ciudad, a la vez que agrupan a las viviendas ubicadas en entornos que comparten condiciones similares. A partir del análisis de la estructura subyacente de las variables que describen el espacio exterior a las viviendas, se encuentran los factores que explican la mayor variación de dichas amenidades. De este modo se clasifican a las viviendas según sean las características del entorno urbano, lo que permite agrupar vecindarios con espacios similares aunque se encuentren dispersos geográficamente. La identificación de este tipo de submercados permite identificar la diversidad de espacios urbanos al interior de una misma ciudad, así como incorporar las condiciones del vecindario en la valoración de las viviendas.

El análisis de componentes principales es una técnica que permite identificar la estructura subyacente de un conjunto de datos a partir de las interrelaciones entre las variables y los elementos que explican la mayor varianza (Hotelling, 1933). En otras palabras, este análisis permite transformar un conjunto de indicadores simples, en este caso de calidad del entorno urbano, a un nuevo conjunto de indicadores compuestos correlacionados o componentes principales.

Entre los criterios de selección del número de componentes óptimos destacan tres. En primer lugar la regla de Kaiser-Guttman (Kaiser, 1991), con la cual se busca retener las dimensiones con un eigenvalor mayor o cercano a uno. En segundo lugar, la elección de los componentes que concentran la mayor varianza acumulada, para lo cual se establecen umbrales. En tercer lugar, el gráfico de sedimentación –*scree plot*– (Cattell, 1996), el cual muestra la contribución de cada componente en la varianza total explicada, con base en el cual se mantienen los componentes anteriores al punto de inflexión a partir del cual la contribución es menos significativa.

Como resultado de la consideración de las pautas antes señaladas y con el fin de aplicar un criterio unificado para la identificación de los submercados por características del vecindario en todas las ciudades de los tres países del área de estudio, se estableció retener los dos primeros componentes principales, los cuales capturan la mayor información de cada entorno urbano en una forma simplificada. Con estos dos componentes principales

resultantes para cada país, es posible dar cuenta de la diversidad de los submercados atendiendo el principio de parsimonia, ya que la conformación de los mismos varía en cada ciudad a la vez que esta acotado dentro de los dos primeros componentes, por lo cual permiten la comparación entre las áreas de estudio.

Una vez identificadas las dos dimensiones subyacentes que explican en su mayoría las diferencias en el entorno urbano, se elaboró una clasificación de las unidades geográficas, asumidas como vecindarios, según la combinación de ambos factores. Dicha categorización muestra si una unidad geográfica tiene un valor propio por arriba o por debajo de la media para el primer y segundo componente. El resultado combinado para las dos dimensiones genera cuatro combinaciones o regímenes espaciales, los cuales se integran en la estimación del modelo de precios hedónicos ampliado o espacial, el cual se explicará más ampliamente en la sección correspondiente.

**Tabla 13 Clasificación vecindarios, combinación de los dos primeros factores principales resultantes del PCA**

Regimen	Factor 1	Factor 2	Regla
1	$\lambda > \bar{x}$ ① ↑ media	$\lambda > \bar{x}$ ② ↑ media	Eigenvalor de los factores 1 y 2 por arriba de la media.
2	$\lambda > \bar{x}$ ① ↑ media	$\lambda < \bar{x}$ ② ↓ media	Eigenvalor del factor 1 por arriba de la media. Eigenvalor del factor 2 por debajo de la media.
3	$\lambda < \bar{x}$ ① ↓ media	$\lambda > \bar{x}$ ② ↑ media	Eigenvalor del factor 1 por debajo de la media. Eigenvalor del factor 2 por arriba de la media.
4	$\lambda < \bar{x}$ ① ↓ media	$\lambda < \bar{x}$ ② ↓ media	Eigenvalor de los factores 1 y 2 por debajo de la media.

Fuente: Elaboración propia.

### 5.1.2 Descripción de los indicadores del entorno urbano

A continuación se presenta la descripción de los conjuntos de datos seleccionados para la clasificación de submercados por características del vecindario, los cuales incorporan diferentes características de la infraestructura, servicios y amenidades en las ciudades. Cabe resaltar que los atributos considerados varían por país, por ejemplo para el caso de México se distingue el nivel de atención de servicios de salud, mientras que para Canadá sólo se

incluye un indicador general para la salud y otro para la cercanía a farmacias, mientras que para Estados Unidos no se encontró una base de datos concentrada sino que se realizó la búsqueda y consolidación de indicadores individuales.

### 1. México

- **Índice de Desigualdad Urbana WRI 2021 (IDU2021):** Elaborado para el *World Resources Institute* WRI a partir de datos del Marco Geoestadístico, DENUE y Censos Económicos de INEGI, Secretaría de Salud y del propio WRI (Brito, Macias, Ramirez, Jacquin, & Zubicaray, 2021). Índices de variables numéricas y categóricas.

**Tabla 14 Variables del conjunto de datos IDU2021**

Clave	Descripción	Tipo <sup>a</sup>
gmu	Grado de Marginación Urbana (CONAPO).	CO
iisu_sun	Índice de Inclusión Social Urbana (Nivel Nacional).	CO
iisu_cd	Índice de Inclusión Social Urbana (Nivel Ciudad).	CO
Empleo	Número de empleos formales accesibles en 30 minutos <sup>b</sup> .	ND
E_basica	Número de escuelas públicas de educación básica accesibles en 15 minutos <sup>b</sup> .	ND
E_media	Número de escuelas públicas de educación media superior accesibles en 30 minutos <sup>b</sup> .	ND
E_superior	Número de escuelas públicas de educación superior accesibles en 30 minutos <sup>b</sup> .	ND
Salud_cama	Número de camas de hospital público accesibles en 30 minutos <sup>b</sup> .	ND
Salud_cons	Número de consultorios médicos públicos accesibles en 30 minutos <sup>b</sup> .	ND
Abasto	Número de unidades económicas minoristas de alimentos accesibles en 20 minutos <sup>b</sup> .	ND
Espacio_ab	Metros cuadrados de espacio abierto accesibles en 20 minutos <sup>b</sup> .	ND
Cultura	Número de instalaciones culturales públicas accesibles en 20 minutos <sup>b</sup> .	ND
Est_Tpte <sup>c</sup>	Número de estaciones del sistema de transporte público masivo accesibles en 15 minutos <sup>b</sup> .	ND

<sup>a)</sup> Claves de tipo de variable:  
 NC – Numérico continuo      CO – Categórico ordinal  
 ND – Numérico discreto      CN – Categórico nominal

<sup>b)</sup> Accesibilidad peatonal, en transporte público masivo o una combinación de ambos.

<sup>c)</sup> Sólo Tijuana y Ciudad Juárez cuentan con la variable “Est\_Tpte”, elaborada por el WRI México en 2020, la cual da cuenta del sistema de transporte público masivo.

Fuente: Elaborado con base en el Índice de Desigualdad Urbana WRI México, 2021.

### 2. Estados Unidos

Recopilación e integración de índices y datos del entorno urbano disponibles para las ciudades de estudio. Índices y variables numéricas.

- **Local Affordability Index (LAI17):** Elaborado por el *Department of Housing and Urban Development* (HUD) y el *Department of Transportation* (DOT), versión 2 del año 2017. Acceso a trabajos y comercios.

- **National Walkability Index (NWI21):** Elaborado por la *Environmental Protection Agency* (EPA) para el año 2021. Clasificación del acceso por manzana a empleos, comercios y presencia de senderos orientados al paso peatonal. Puntajes y clasificaciones numéricas.
- **Park Serve (PKS22):** Elaborado por el *Trust for Public Lands* (TPL) para el año 2022. Número de parques y metros cuadrados de extensión de los mismos.
- **Public Libraries (PLS20):** Elaborado por el *Institute of Museum and Library Services* (IMLS) en el año 2020. Número de bibliotecas y área de cobertura.

**Tabla 15 Variables de los conjuntos de datos LAI17, NWI21, PKS22 y PLS20**

Clave	Descripción	Tipo <sup>a</sup>
<i>Local Affordability Index (LAI17)</i>		
median_com	Distancia mediana de traslado entre los centros de trabajo y las viviendas.	NC
employment	Índice de acceso a empleos. Cociente del número de empleos entre la distancia al cuadrado.	NC
local_job_	Densidad de empleos locales. Cociente del número de empleos entre el área total.	NC
local_reta	Densidad de empleos en comercios locales. Cociente del número de empleos en el sector comercial entre el área total.	NC
retail_acc	Acceso a comercios minoristas. Cociente del número de comercios locales entre la distancia al cuadrado.	NC
hh_type1_t	Gasto estimado en transporte como porcentaje del ingreso del hogar.	NC
hh_type1_2	Viajes estimados por hogar.	NC
<i>National Walkability Index (NWI21)</i>		
D2B_E8MIXA	Mezcla de diferentes tipos de empleos (industrial, administrativo, comercial) en la unidad censal <i>block group</i> <sup>*</sup> o grupo de bloques.	NC
D2A_EPHHM	Mezcla de diferentes tipos de empleo y viviendas ocupadas en la unidad censal <i>block group</i> .	NC
D3B	Densidad de intersecciones peatonales.	NC
D4A	Proximidad a paradas del sistema de transporte.	NC
NatWalkInd	Índice de cuan amigable para el peatón es el <i>block group</i> .	NC
D2A_Ranked	Puntaje de clasificación del criterio mezcla de tipos de empleos.	ND
D2B_Ranked	Puntaje de clasificación del criterio mezcla de empleos y viviendas.	ND
D3B_Ranked	Puntaje de clasificación del criterio de senderos peatonales.	ND
D4A_Ranked	Puntaje de clasificación del criterio servicio de transporte.	ND
<i>Park Serve (PKS22)</i>		
sumarea	Sumatoria de metros cuadrados totales de extensión de parques en la unidad censal <i>block group</i> .	ND
<i>Public Libraries (PLS20)</i>		
linum	Sumatoria del número de bibliotecas ubicadas en el <i>block group</i> o en un radio de distancia accesible.	ND

<sup>a)</sup> Claves de tipo de variable:

NC – Numérico continuo      CO – Categórico ordinal  
 ND – Numérico discreto      CN – Categórico nominal

<sup>\*</sup>*Block Group* es una unidad geográfica definida por el *U.S. Census Bureau*, con un nivel de desagregación menor que un *Census Tract* y mayor a un *Block* o manzana.

Fuente: Elaborado con base en información de HUD, EPA, TPL y el IMLS, 2017-2022.

### 3. Canadá

- **Proximity Measure Database (PMD2020):** Elaborado por la agencia estadística gubernamental *Statistics Canada*. Índices numéricos, en su mayoría continuos. Una tienda de comestibles, una farmacia y una parada de transporte público a menos de 1 km a pie; una guardería, una escuela primaria y una biblioteca a menos de 1,5 km a pie; un centro de salud a menos de 3 km en coche; y un empleo a menos de 10 km en coche.

**Tabla 16 Variables del conjunto de datos PMD2020**

Clave	Descripción	Tipo <sup>a</sup>
prox_idx_emp	Índice de cercanía a centros de empleo a una distancia menor de 10km en vehículo particular.	NC
prox_idx_health	Índice de cercanía a centros de salud a una distancia menor de 3km en vehículo particular.	NC
prox_idx_childcare	Índice de cercanía guarderías y centros de cuidado infantil a una distancia menor de 1.5km caminando.	NC
prox_idx_educpri	Índice de cercanía a centros de educación primaria a una distancia menor de 1.5km caminando.	NC
prox_idx_educsec	Índice de cercanía a centros de educación secundaria a una distancia menor de 1.5km caminando.	NC
prox_idx_lib	Índice de cercanía a bibliotecas a una distancia menor de 1.5km caminando.	NC
prox_idx_parks	Índice de cercanía a parques urbanos a una distancia menor de 1km caminando.	NC
prox_idx_grocery	Índice de cercanía a tiendas minoristas de alimentos a una distancia menor de 1km caminando.	NC
prox_idx_pharma	Índice de cercanía a farmacias o boticas a una distancia menor de 1km caminando.	NC
prox_idx_transit	Índice de cercanía a servicios de transporte público a una distancia menor de 1km caminando.	NC
amenity_dense	Indicador de densidad de equipamientos y servicios en el <i>dissemination block</i> o bloque de manzanas.	ND

<sup>a)</sup> Claves de tipo de variable:

NC – Numérico continuo

CO – Categórico ordinal

ND – Numérico discreto

CN – Categórico nominal

\* Medidas calculadas con base en distancias de red entre centroides de los *Dissemination Blocks* o bloques de manzanas.

Fuente: Elaborado con base en *Proximity Measure Database*, Statistics Canada, 2020.

#### 5.1.3 Resultados del análisis de componentes principales

A continuación se presentan los principales resultados del análisis de componentes principales, divididos por país y por ciudad. En general, para todos los casos se aplicó la misma técnica estadística PCA convencional, pero para los casos de observaciones con registros faltantes, se empleó un método avanzado de análisis de componentes que permite

imputar valores más consistentes que la media o la mediana, al calcular previamente el PCA y aplicar la imputación con los resultados obtenidos (Josse & Husson, 2016). Para el caso de México, se aplicó primero un método avanzado que permite el tratamiento adecuado a la conjunción de variables cuantitativas y cualitativas, conocido como análisis de factores para datos mixtos –*Factor Analysis of Mixed Data*, FAMD–(Pagès, 2004), posteriormente se aplicó el PCA convencional.

### 1. México

El primer paso consistió en aplicar el método de tratamiento para valores faltantes, específicos para la clasificación de PCA llamado missMDA para R. Como se mencionó arriba, esta herramienta permite identificar los valores más probables que tendrían los registros faltantes a partir del cálculo de sus componentes principales, previo al análisis PCA en sí. Su aplicación puede darse tanto en el caso de variables cuantitativas como cualitativas, sin embargo, debe establecerse previamente el tipo de datos a tratar así como el número de componentes a extraer. Posteriormente, se realizó el método avanzado de análisis de componentes para variables cuantitativas y cualitativas FAMD, cuyos principales resultados se muestran a continuación.

**Tabla 17 Resumen FAMD: número de factores y varianza acumulada**

	01TIJ	02MXL	03NOG	04CJZ	05PDN	06NLD	07REY	08MTM
<i>Número de factores*</i>	3	2	3	3	2	2	2	2
<i>% Varianza acumulada 2 dimensiones</i>	31.7786	35.5056	37.4021	38.3478	45.2495	36.7615	35.7536	39.7132

Nota: \*Gráfica de sedimentación, método de agrupamiento del codo o *Elbowmethod*.

Fuente: Elaboración propia.

**Tabla 18 Contribución de las variables originales en los primeros dos componentes principales por ciudad, FAMD**

	MX01 Tijuana				MX02 Mexicali			
	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>
1° var.	iisu_cd	0.5544	iisu_sun	0.9218	Empleo	0.6549	gmu	0.9376
2° var.	gmu	0.5382	gmu	0.9218	iisu_sun	0.5714	iisu_sun	0.9376
3° var.	iisu_sun	0.5382	iisu_cd	0.8766	gmu	0.5714	iisu_cd	0.8230
4° var.	Empleo	0.5333	Abasto	0.2102	iisu_cd	0.5122	Espacio_ab	0.0615
5° var.	Salud_cons	0.4009	E_basica	0.1686	Salud_cons	0.4413	Abasto	0.0438

MX03 Nogales					MX04 Ciudad Juárez			
	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>
1° var.	Abasto	0.7620	iisu_sun	0.8319	Empleo	0.5981	gmu	0.8015
2° var.	Salud_cons	0.6919	iisu_cd	0.7261	Salud_cons	0.5976	iisu_sun	0.8015
3° var.	Salud_cama	0.6522	gmu	0.6921	gmu	0.5287	iisu_cd	0.6657
4° var.	Empleo	0.6479	Espacio_ab	0.2014	iisu_sun	0.5287	Abasto	0.4159
5° var.	E_basica	0.6213	E_superior	0.1607	E_media	0.4571	E_basica	0.3711

MX05 Piedras Negras					MX06 Nuevo Laredo			
	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>
1° var.	Empleo	0.7372	gmu	0.9622	Empleo	0.7740	iisu_cd	0.8279
2° var.	E_media	0.6457	iisu_sun	0.9622	E_basica	0.6345	iisu_sun	0.7789
3° var.	gmu	0.6177	iisu_cd	0.8468	Salud_cons	0.5657	gmu	0.6984
4° var.	iisu_sun	0.6177	Abasto	0.0236	iisu_cd	0.4958	Abasto	0.2076
5° var.	Salud_cons	0.5893	E_basica	0.0215	Salud_cama	0.4857	E_basica	0.1053

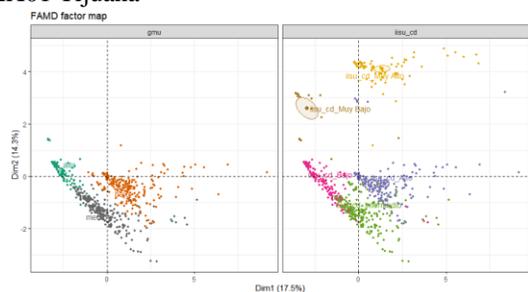
MX07 Reynosa					MX08 Matamoros			
	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>
1° var.	Empleo	0.7624	iisu_sun	0.8655	Empleo	0.7451	iisu_sun	0.8747
2° var.	E_media	0.6526	iisu_cd	0.8474	Salud_cons	0.7062	iisu_cd	0.8587
3° var.	Salud_cons	0.5501	gmu	0.6092	Salud_cama	0.6566	gmu	0.6420
4° var.	iisu_cd	0.5284	Abasto	0.1895	E_media	0.6462	Abasto	0.0795
5° var.	iisu_sun	0.4822	E_basica	0.1202	E_superior	0.5892	E_basica	0.0603

<sup>a)</sup> Varianza de las coordenadas principales correspondientes al valor propio de la dimensión.

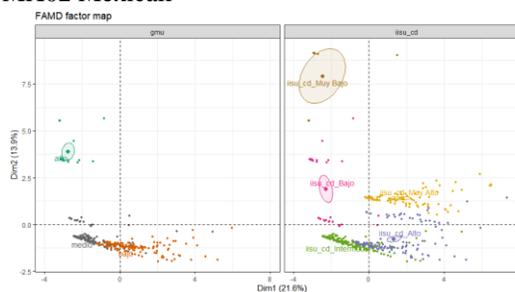
Fuente: Elaboración propia.

## Gráfica 10. Observaciones agrupadas por variables cualitativas del entorno urbano, FAMD México

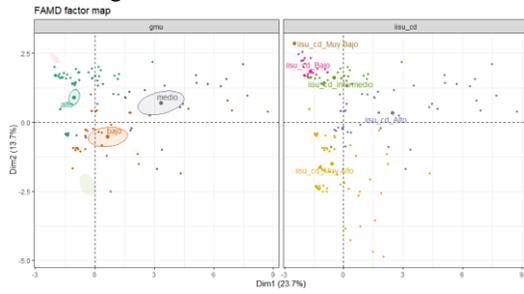
MX01 Tijuana



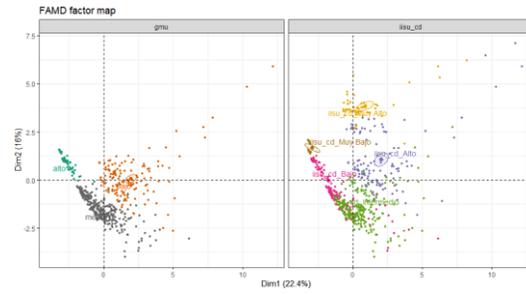
MX02 Mexicali



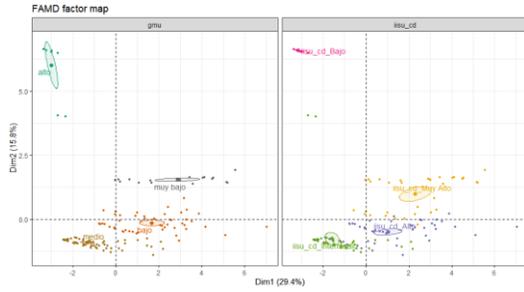
MX03 Nogales



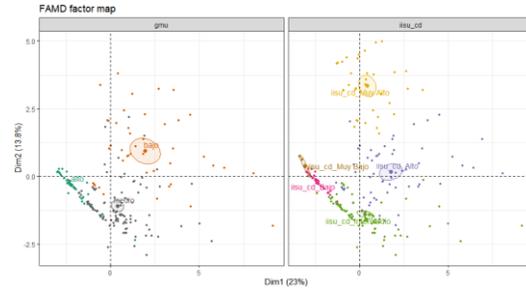
MX04 Ciudad Juárez



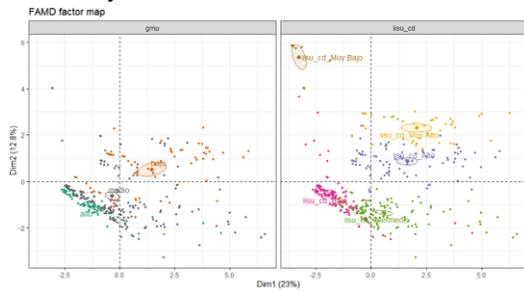
MX05 Piedras Negras



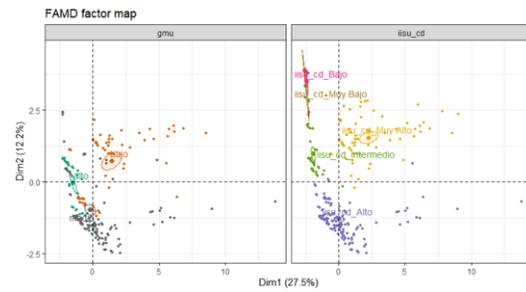
MX06 Nuevo Laredo



MX07 Reynosa



MX08 Matamoros



Fuente: Elaboración propia.

Una vez revisados los resultados del FAMD, se llegó a la conclusión de que las dimensiones resultantes son difíciles de interpretar, debido a que incluyen entre las características más importantes a variables que son resultado del agregado de otras o que incluyen en su construcción diferentes a las características del vecindario, las cuales se busca clasificar. Es por ello que se realizó un nuevo análisis en su vertiente convencional de PCA incluyendo solamente las variables numéricas del entorno urbano.

**Tabla 19 Resumen PCA: número de factores y varianza acumulada**

	<b>01TIJ</b>	<b>02MXL</b>	<b>03NOG</b>	<b>04CJZ</b>	<b>05PDN</b>	<b>06NLD</b>	<b>07REY</b>	<b>08MTM</b>
<i>Varianza acumulada en dos dimensiones (%):</i>								
	46.3036	58.1383	67.787	62.5872	60.7412	59.1825	60.0713	72.1212
<i>Número de factores óptimos*:</i>								
	2	2	3	3	2	2	2	2

Nota: \*Gráfica de sedimentación, método de agrupamiento del codo o *Elbowmethod*.

Fuente: Elaboración propia.

**Tabla 20 Contribución de las variables originales en los primeros dos componentes principales por ciudad, PCA**

		<b>MX01 Tijuana</b>				<b>MX02 Mexicali</b>			
	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	
1° var.	Empleo	0.8318	E_basica	0.7268	Empleo	0.8240	Salud_cama	0.5984	
2° var.	Salud_cons	0.7547	Abasto	0.6761	Salud_cons	0.7602	Salud_cons	0.5223	
3° var.	Salud_cama	0.6011	E_media	0.1130	E_basica	0.7237	Cultura	0.4605	
4° var.	E_media	0.5869	E_superior	0.0385	Abasto	0.7152	E_superior	0.2328	
5° var.	Abasto	0.5862	Cultura	-0.0241	E_media	0.6820	Espacio_ab	-0.0074	

		<b>MX03 Nogales</b>				<b>MX04 Ciudad Juárez</b>			
	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	
1° var.	Abasto	0.8876	E_superior	0.6726	Salud_cons	0.8725	Abasto	0.7048	
2° var.	Empleo	0.8622	E_media	0.6248	E_superior	0.7457	E_basica	0.6699	
3° var.	Salud_cons	0.8566	Espacio_ab	0.6169	Empleo	0.7136	Est_Tpte	0.5110	
4° var.	Salud_cama	0.8311	Empleo	0.1184	E_media	0.7041	E_media	0.3115	
5° var.	E_basica	0.8175	E_basica	0.0656	Cultura	0.7012	Empleo	0.2546	

		<b>MX05 Piedras Negras</b>				<b>MX06 Nuevo Laredo</b>			
	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	
1° var.	Empleo	0.8737	Cultura	0.6816	Empleo	0.9165	E_media	0.7383	
2° var.	E_media	0.8565	Espacio_ab	0.4621	Salud_cons	0.8721	Abasto	0.5262	
3° var.	Salud_cons	0.8322	E_basica	0.2835	E_basica	0.8164	E_basica	0.3794	
4° var.	E_basica	0.8083	Abasto	0.2285	Salud_cama	0.8135	Espacio_ab	0.1298	
5° var.	Abasto	0.7357	E_media	0.1156	E_superior	0.6631	E_superior	0.0566	

		<b>MX07 Reynosa</b>				<b>MX08 Matamoros</b>			
	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	
1° var.	Empleo	0.8989	Abasto	0.5521	Salud_cons	0.9031	Abasto	0.7385	
2° var.	E_media	0.8490	Espacio_ab	0.5362	Salud_cama	0.8714	E_basica	0.6278	
3° var.	Salud_cons	0.8072	E_basica	0.3767	Empleo	0.8569	Empleo	0.1587	
4° var.	E_basica	0.7236	Cultura	0.3576	E_media	0.8503	Espacio_ab	0.0894	
5° var.	Abasto	0.6226	E_media	-0.0854	Cultura	0.8083	E_media	-0.0618	

Nota: <sup>a)</sup> Varianza de las coordenadas principales correspondientes al valor propio de la dimensión.

Fuente: Elaboración propia.

Los resultados obtenidos por medio del PCA fueron satisfactorios, ya que permiten identificar claramente las variables que contribuyen a la mayor varianza de las primeras donde dimensiones, como se muestra en las tablas anteriores. A pesar de que las dimensiones cambian para cada ciudad, se encuentra que las variables empleo y salud son

las que más contribuyen en la composición de los dos primeros factores, los cuales permiten clasificar a los vecindarios entre aquellos que cuentan con una mayor dotación y acceso a estos servicios y los que presentan una distribución por debajo de la media en dichas características del entorno urbano.

## 2. Estados Unidos

Para el caso de las ciudades estadounidenses, se recopiló información de diferentes atributos, con el fin de acercarse a las clasificaciones del entorno urbano con las que se cuentan en México y Canadá. Los datos seleccionados fueron de tipo numérico, por lo cual se aplicó directamente la técnica convencional de *Principal Component Analysis* –PCA–. Asimismo, debido a que algunas observaciones presentaban algunos valores faltantes, se empleó el método missPCA antes de la clasificación. Los resultados obtenidos se detallan a continuación.

**Tabla 21 Resumen PCA: número de factores y varianza acumulada**

<b>01SD</b>	<b>02IMP</b>	<b>03NOG</b>	<b>04ELP</b>	<b>05EGP</b>	<b>06LRD</b>	<b>07MCA</b>
<i>Varianza acumulada en dos dimensiones (%):</i>						
50.4484	60.8023	72.1708	53.1139	69.6742	59.8384	57.8871
<i>Número de factores óptimos*:</i>						
2	2	3	3	2	3	2
<b>08BRW</b>	<b>09WHA</b>	<b>10SSM</b>	<b>11DET</b>	<b>12PTH</b>	<b>13NIA</b>	
<i>Varianza acumulada en dos dimensiones (%):</i>						
57.9753	64.7272	74.0325	44.1674	60.1906	55.8519	
<i>Número de factores óptimos*:</i>						
3	2	3	3	2	3	

\*Gráfica de sedimentación, método de agrupamiento del codo o *Elbowmethod*.

Fuente: Elaboración propia.

**Tabla 22 Contribución de las variables originales en los primeros dos componentes principales por ciudad, PCA**

		<b>US01 San Diego</b>				<b>US02 Imperial</b>			
		<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>
1° var.	NatWalkInd	0.8346	D2A_EPHHM	0.8259	NatWalkInd	0.9368	D2A_Ranked	0.8356	
2° var.	D3B_Ranked	0.7578	D2A_Ranked	0.8184	local_job_	0.8681	D2A_EPHHM	0.8265	
3° var.	D4A_Ranked	0.7429	D2B_Ranked	0.7709	employment	0.8385	D2B_E8MIXA	0.6553	
4° var.	retail_acc	0.7271	D2B_E8MIXA	0.7701	D3B_Ranked	0.7534	D2B_Ranked	0.6455	
5° var.	employment	0.7258	NatWalkInd	0.2897	retail_acc	0.7477	hh_type1_t	0.1855	

<b>US03 Nogales</b>				<b>US04 El Paso</b>				
<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	
1° var.	retail_acc	0.8894	D2B_E8MIXA	0.8286	NatWalkInd	0.8889	D2A_EPHHM	0.8313
2° var.	NatWalkInd	0.8801	D2B_Ranked	0.8280	employment	0.8245	D2A_Ranked	0.8196
3° var.	employment	0.8626	D2A_EPHHM	0.6477	D4A_Ranked	0.7510	D2B_Ranked	0.7305
4° var.	local_reta	0.8599	D2A_Ranked	0.6370	D3B_Ranked	0.7147	D2B_E8MIXA	0.7174
5° var.	local_job_	0.8596	median_com	0.5033	local_job_	0.6981	D4A	0.2793

<b>US05 Eagle Pass</b>				<b>US06 Laredo</b>				
<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	
1° var.	NatWalkInd	0.9460	sumarea	0.8737	NatWalkInd	0.9120	D2B_E8MIXA	0.7402
2° var.	employment	0.9256	D2A_EPHHM	0.8445	employment	0.9030	D2A_EPHHM	0.7389
3° var.	retail_acc	0.9045	D2A_Ranked	0.8419	retail_acc	0.8465	D2A_Ranked	0.7367
4° var.	D3B_Ranked	0.8515	hh_type1_t	0.2623	D3B_Ranked	0.7618	D2B_Ranked	0.7167
5° var.	local_job_	0.8356	local_reta	0.2354	local_job_	0.7309	hh_type1_t	0.3399

<b>US07 McAllen</b>				<b>US08 Brownsville</b>				
<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	
1° var.	NatWalkInd	0.9352	D2B_E8MIXA	0.6483	retail_acc	0.8725	D2B_E8MIXA	0.7764
2° var.	employment	0.8678	D2B_Ranked	0.6352	employment	0.8673	D2B_Ranked	0.7710
3° var.	retail_acc	0.8116	D2A_EPHHM	0.5139	NatWalkInd	0.8467	D2A_EPHHM	0.7246
4° var.	D3B_Ranked	0.8077	D2A_Ranked	0.4987	D3B_Ranked	0.8025	D2A_Ranked	0.7194
5° var.	local_job_	0.8057	hh_type1_t	0.2999	D3B	0.7287	NatWalkInd	0.3762

<b>US09 Whatcom</b>				<b>US10 Sault St. Marie</b>				
<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	
1° var.	employment	0.9108	D2A_EPHHM	0.8380	NatWalkInd	0.9700	D2A_EPHHM	0.7878
2° var.	NatWalkInd	0.9067	D2A_Ranked	0.8274	local_reta	0.9269	D2A_Ranked	0.7581
3° var.	retail_acc	0.8857	D2B_E8MIXA	0.4367	employment	0.9042	D2B_E8MIXA	0.7214
4° var.	D3B_Ranked	0.8686	D2B_Ranked	0.4152	retail_acc	0.9035	D2B_Ranked	0.6965
5° var.	local_reta	0.8460	median_com	0.3762	D3B_Ranked	0.8711	hh_type1_t	0.3293

<b>US11 Detroit</b>				<b>US12 Point Edward</b>				
<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	
1° var.	NatWalkInd	0.7495	D2A_EPHHM	0.8100	NatWalkInd	0.8985	D2B_Ranked	0.8323
2° var.	D4A_Ranked	0.7003	D2A_Ranked	0.7961	D3B_Ranked	0.8725	D2B_E8MIXA	0.8194
3° var.	D3B_Ranked	0.6773	D2B_E8MIXA	0.7951	D3B	0.8384	D2A_EPHHM	0.5954
4° var.	employment	0.6750	D2B_Ranked	0.7894	employment	0.8374	D2A_Ranked	0.5696
5° var.	D3B	0.6731	local_reta	0.4914	local_job_	0.7806	NatWalkInd	0.3041

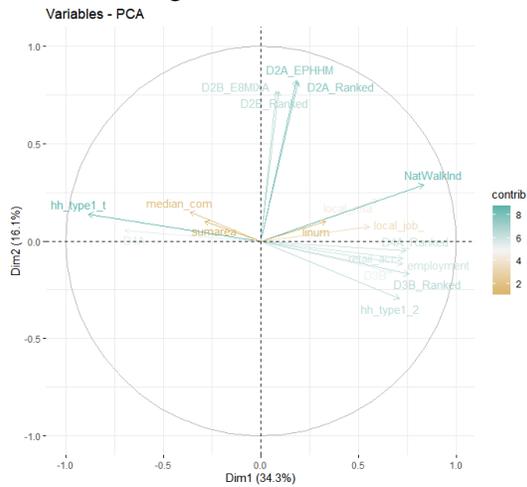
US13 Buffalo-Niagara				
	Dim. 1	Coord. <sup>a</sup>	Dim. 2	Coord. <sup>a</sup>
1° var.	employment	0.8529	D2A_EPHHM	0.8654
2° var.	NatWalkInd	0.8385	D2A_Ranked	0.8587
3° var.	D4A_Ranked	0.8302	D2B_Ranked	0.8092
4° var.	D3B_Ranked	0.8069	D2B_E8MIXA	0.8077
5° var.	hh_type1_2	0.7882	NatWalkInd	0.3691

<sup>a)</sup> Varianza de las coordenadas principales correspondientes al valor propio de la dimensión.

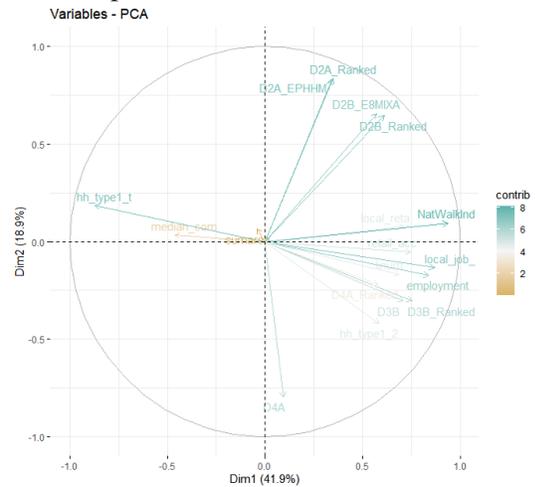
Fuente: Elaboración propia.

### Gráfica 11. Círculo de correlación, contribución de las variables en las dimensiones 1 y 2 del entorno urbano, PCA Estados Unidos

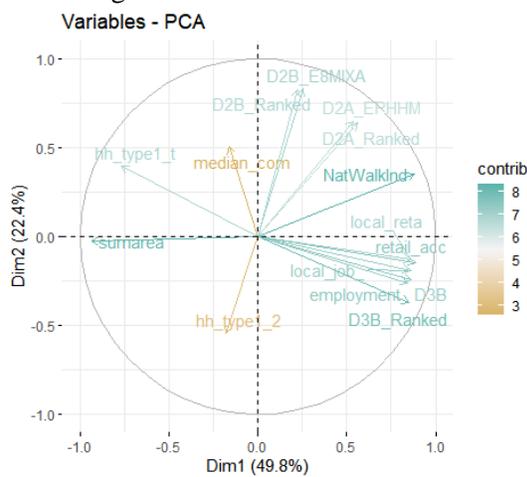
US01 San Diego



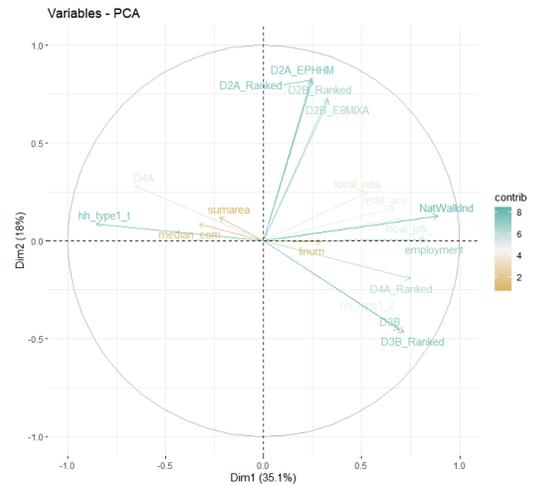
US02 Imperial



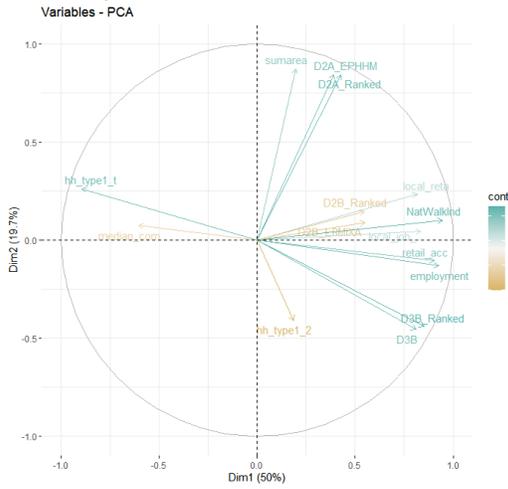
US03 Nogales



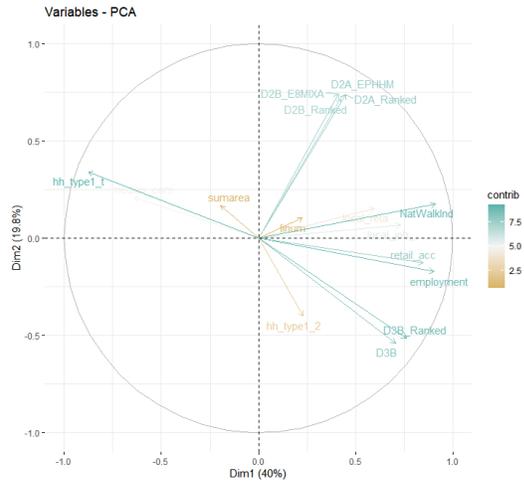
US04 El Paso



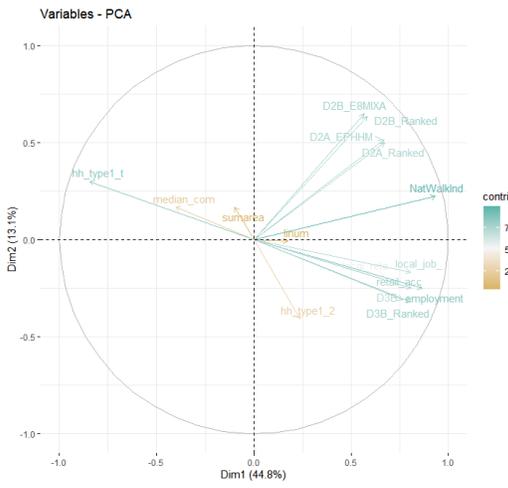
### US05 Eagle Pass



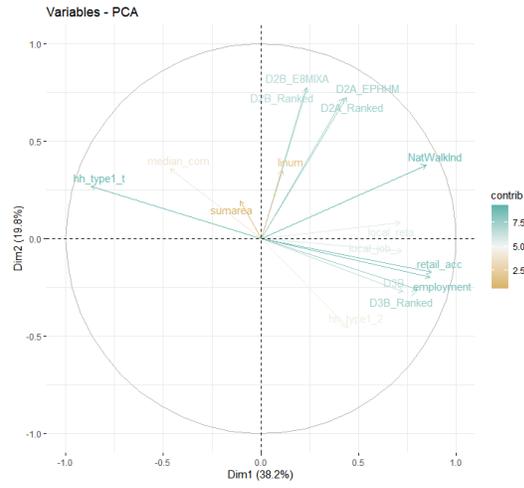
### US06 Laredo



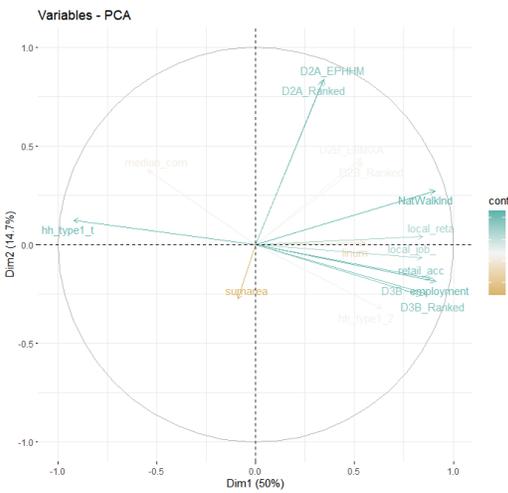
### US07 McAllen



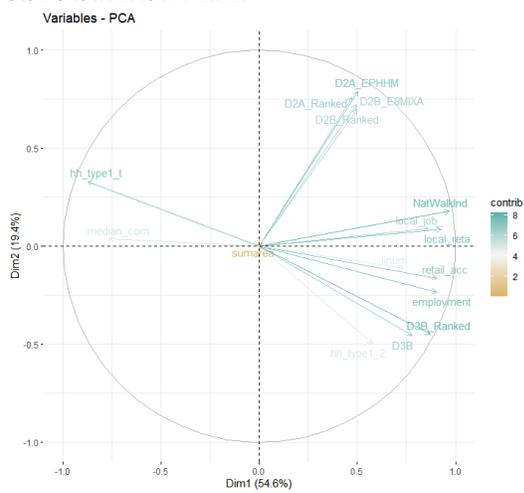
### US08 Brownsville



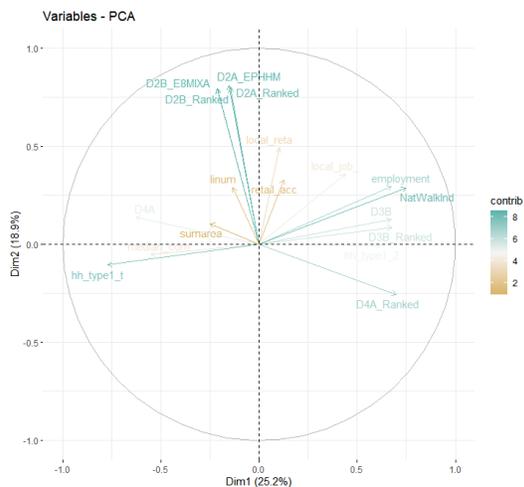
### US09 Whatcom



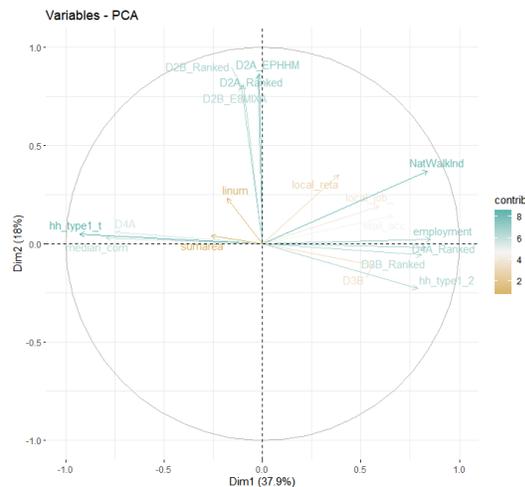
### US10 Sault St. Marie



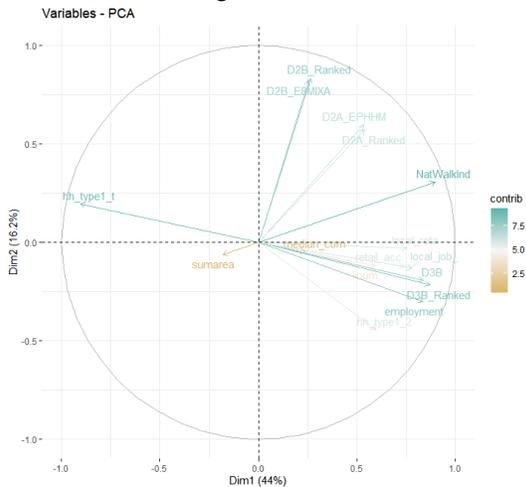
### US11 Detroit



### US12 Point Edward



### US13 Buffalo-Niagara



Fuente: Elaboración propia.

### 3. Canadá

En el caso de las ciudades canadienses, los datos empleados son de tipo numéricos continuos, por lo cual fue aplicada la metodología de PCA convencional. Los resultados de la clasificación se detallan a continuación.

**Tabla 23 Resumen PCA: número de factores y varianza acumulada**

	09MVR	10SSM	11WIN	12PTE	13GGH
Número de factores*	77.0941	80.9459	65.2199	80.8954	75.1239
% Varianza acumulada 2 dimensiones	2	1	2	1	2

\*Gráfica de sedimentación, método de agrupamiento del codo o *Elbowmethod*.

Fuente: Elaboración propia.

**Tabla 24 Contribución de las variables originales en los primeros dos componentes principales por ciudad, PCA**

<b>CA09 Metro Vancouver</b>					<b>CA10 Sault Sainte Marie</b>			
	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>
1°	prox_idx_e		prox_idx_e		prox_idx_li		amenity_de	
var.	mp	0.9286	ducpri	0.6555	b	0.9718	nse	0.9764
2°	prox_idx_g		prox_idx_e		prox_idx_e		prox_idx_e	
var.	rocery	0.9116	ducsec	0.6446	mp	0.9521	ducpri	0.2187
3°	prox_idx_p		prox_idx_c		prox_idx_e		prox_idx_g	
var.	harma	0.8693	hildcare	0.2340	ducsec	0.9328	rocery	0.1087
4°	prox_idx_c		prox_idx_li		prox_idx_p		prox_idx_c	
var.	hildcare	0.8636	b	0.2243	harma	0.9213	hildcare	0.0885
5°	prox_idx_tr		prox_idx_p		prox_idx_g		prox_idx_li	
var.	ansit	0.8463	arks	0.1304	rocery	0.9172	b	-0.0081

<b>CA11 Winsor</b>					<b>CA12 Port Huron</b>			
	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>
1°	prox_idx_e		amenity_de		prox_idx_g		prox_idx_h	
var.	mp	0.8762	nse	0.9456	rocery	0.9427	ealth	0.5914
2°	prox_idx_g		prox_idx_tr		prox_idx_li		prox_idx_e	
var.	rocery	0.8719	ansit	0.2646	b	0.9206	mp	0.3446
3°	prox_idx_p		prox_idx_p		prox_idx_p		prox_idx_p	
var.	harma	0.8327	harma	0.0890	arks	0.9084	harma	0.1516
4°	prox_idx_e		prox_idx_g		prox_idx_p		prox_idx_e	
var.	ducsec	0.8287	rocery	0.0800	harma	0.8757	ducsec	0.0582
5°	prox_idx_tr		prox_idx_e		prox_idx_e		prox_idx_g	
var.	ansit	0.7895	mp	-0.0315	mp	0.8753	rocery	0.0339

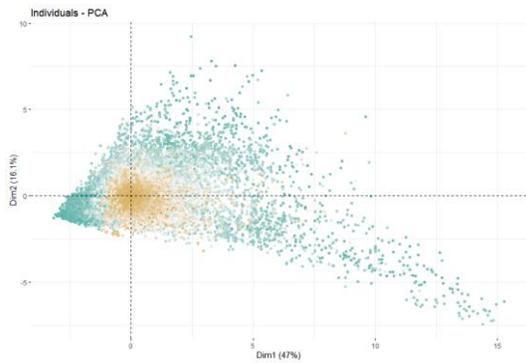
<b>CA13 Great Golden Horseshoe</b>				
	<i>Dim. 1</i>	<i>Coord.<sup>a</sup></i>	<i>Dim. 2</i>	<i>Coord.<sup>a</sup></i>
1°	prox_idx_e		amenity_de	
var.	mp	0.9197	nse	0.6481
2°	prox_idx_p		prox_idx_h	
var.	harma	0.9060	ealth	0.3840
3°	prox_idx_p		prox_idx_tr	
var.	arks	0.8918	ansit	0.2845
4°	prox_idx_e		prox_idx_e	
var.	ducsec	0.8794	mp	0.1036
5°	prox_idx_g		prox_idx_p	
var.	rocery	0.8674	harma	0.0707

<sup>a)</sup> Varianza de las coordenadas principales correspondientes al valor propio de la dimensión.

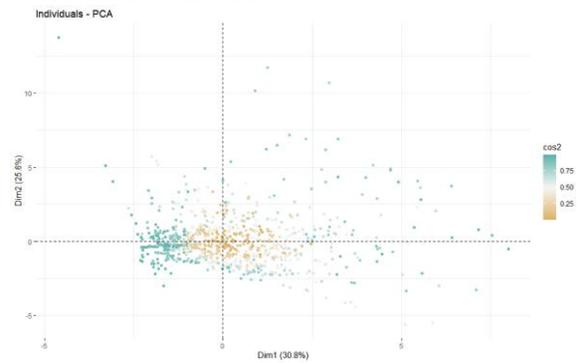
Fuente: Elaboración propia.

## Gráfica 12. Cuadrado del eigenvalor de cada observación del entorno urbano ( $\cos^2$ ), PCA Canadá

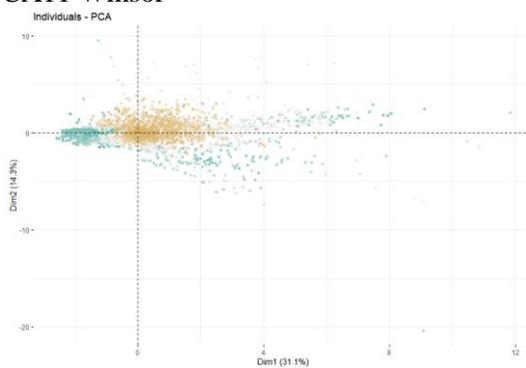
CA09 Metro Vancouver



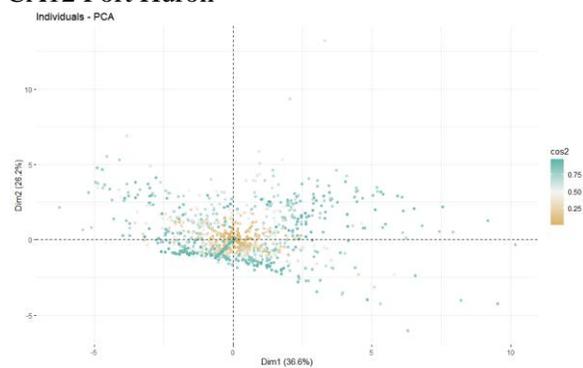
CA10 Sault Sainte Marie



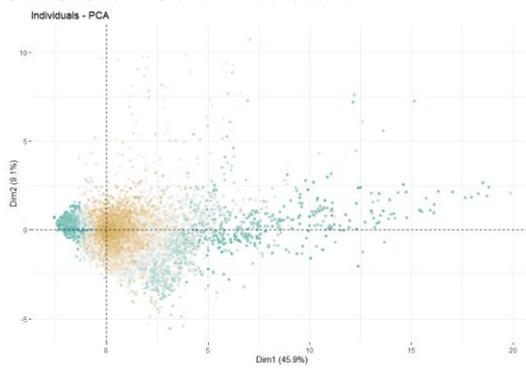
CA11 Winsor



CA12 Port Huron



CA13 Great Golden Horseshoe



Fuente: Elaboración propia.

## 5.2. Submercados por características de la vivienda

### 5.2.1. Análisis de Conglomerados (K-Means)

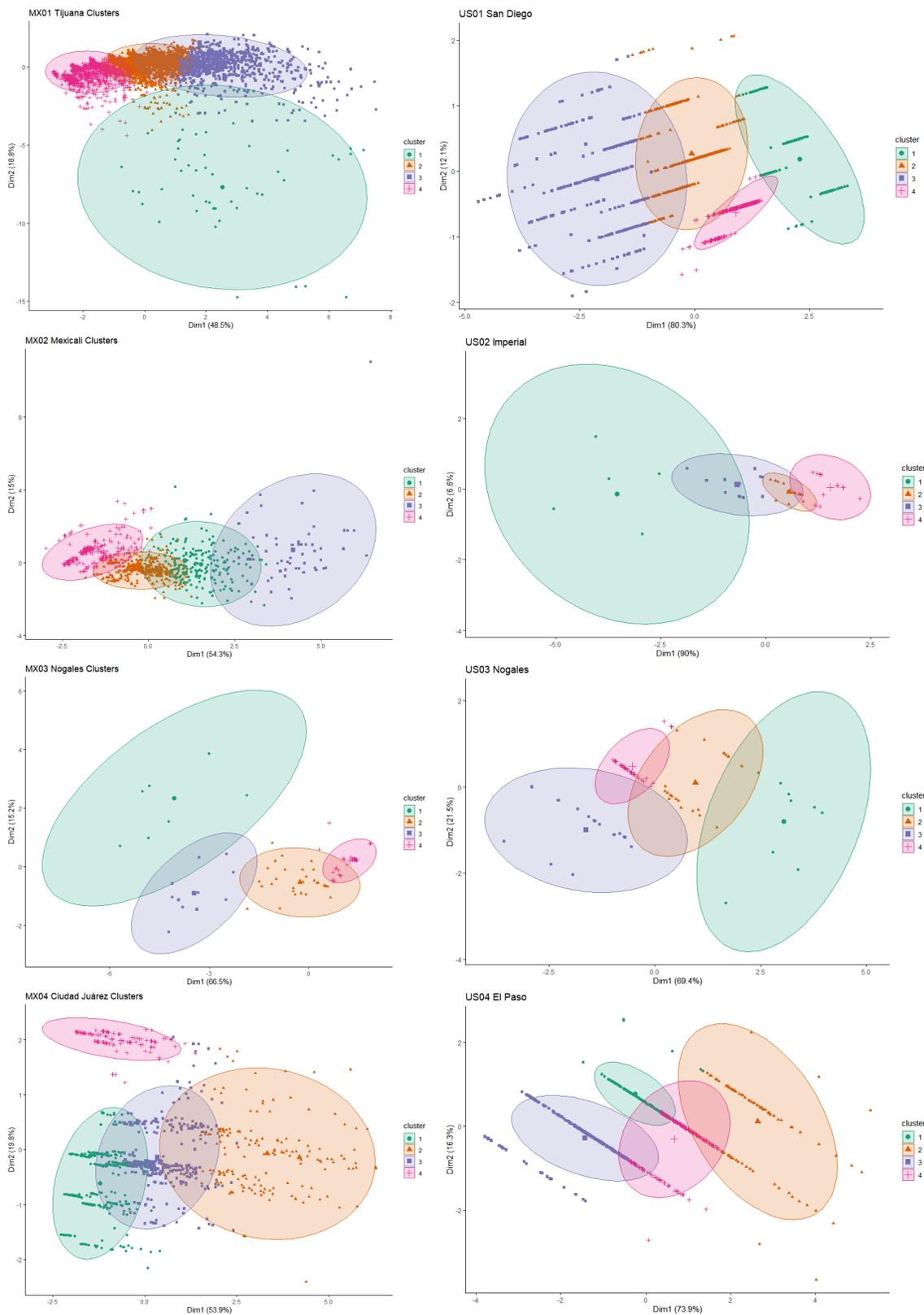
Los submercados por características de la vivienda permiten la identificación de atributos que distinguen a dichas propiedades en grupos, los cuales expresan una combinación de factores intrínsecos que permiten diferenciar a las viviendas entre sí y con ello, contribuir en la explicación de las características de la oferta y comportamiento de los precios en dicho mercado.

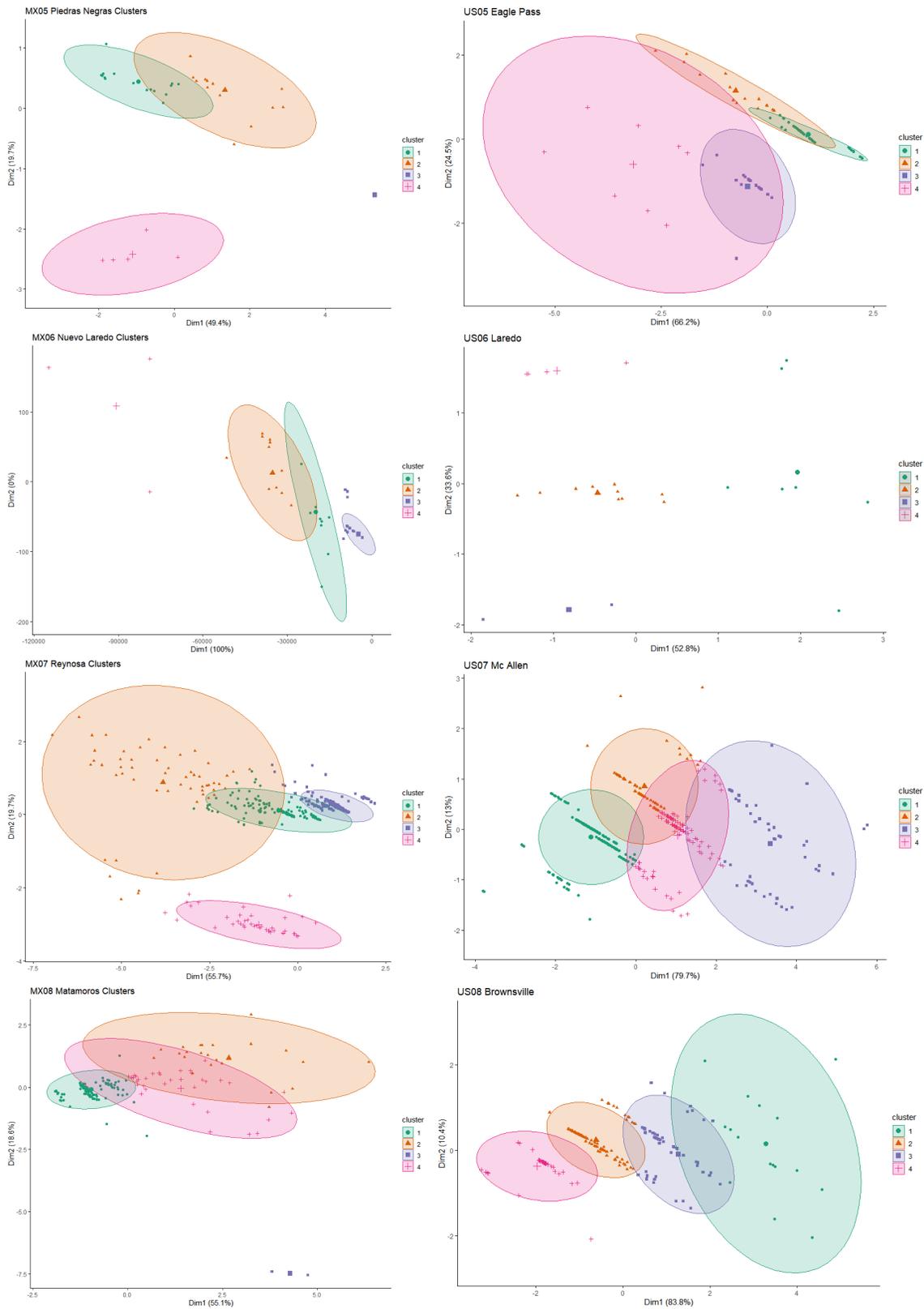
Para agrupar a las viviendas según sean más parecidas entre sí y diferentes del resto de observaciones, se emplearon técnicas de clasificación supervisada como K-Medias (Lloyd, 1982), K-NN y HCA, las cuales agrupan a las observaciones a partir de la minimización de las distancias intra-clúster, es decir de semejanza, a la vez que mantienen la mayor diferencia inter-cluster o entre conglomerados.

En primer lugar, para el análisis de K-Medias el número óptimo de clúster se determinó a partir de tres estadísticos: Suma Total de Cuadrados Intra-Cluster –TWSS–, Promedio del Ancho de Silhouette y Estadístico Gap. El método TWSS, también llamado *Elbow* por la forma que toma su representación gráfica, busca el número de clúster que minimiza la variación al interior de cada grupo (Milligan & Cooper, 1985). Por su parte, la Silhouette mide la calidad de los conglomerados al determinar que tan bien clasificada está cada observación dentro de cada grupo (Batoool & Hennig, 2021; Kaufman & Rousseuw, 1990). Por último, el Estadístico Gap compara el cambio en la dispersión intra-clúster a partir del cálculo de resultados de clasificaciones con diferente número de k, este método se diferencia de los dos anteriores los cuales miden únicamente características globales (Tibshirani, Walther, & Hastie, 2001).

Para la mayoría de las ciudades estudiadas, los resultados coincidieron en un número de clusters  $k=4$  bajo el método TWSS o Elbow,  $k=2$  de Silhouette y  $k=1$  según el estadístico de Gap. Se estableció un mismo k para todas las ciudades, con el fin de que permitir la comparación que satisfaga los criterios óptimos y sea coherente con las características propias del tipo de observaciones a clasificar, en este caso, viviendas. A manera de ejemplo, se presentan a continuación los resultados para México y Estados Unidos.

**Gráfica 13. Clústeres resultantes, clasificación por K-Medias, frontera México-Estados Unidos**





Fuente: Elaboración propia.

Bajo el criterio homologado de 4 clústeres en todas las ciudades, se observa que el tamaño de los subgrupos resultantes presenta un marcado desbalance para algunas áreas de estudio. Esto ocurre en particular en aquellas ciudades con un menor número de observaciones, como Piedras Negras y Nuevo Laredo, o con un comportamiento más disperso en sus características, como Matamoros. Para las ciudades de Estados Unidos, el caso que muestra mejor la dificultad de clasificar observaciones muy dispersas es Eagle Pass y Laredo, en el primer caso los clústeres se superponen uno arriba del otro, lo que indica que no es una buena clasificación, mientras que en el caso de Laredo la agrupación es casi imperceptible a simple vista.

Todos estos resultados apuntan a una de las principales condicionantes de la clasificación por K-Medias, la esfericidad de los datos. Debido a que la varianza de las características intrínsecas de las viviendas no es constante, los datos no están agrupados de forma esférica que permita aplicar de la mejor manera la identificación de los centroides como parámetros iniciales, ni el posterior cálculo de la distancia euclidiana entre los mismos y cada uno de las observaciones a clasificar dentro de los conglomerados.

Asimismo, es notorio que debido a que las variables número de cuartos y número de baños son numéricas discretas, existen líneas paralelas de puntos dentro del espacio bidimensional mostrado en las gráficas, lo cual representa un patrón difícil de clasificar de manera óptima por medio de esta técnica. Es por ello que se procedió a la aplicación de la clasificación jerárquica de clústeres, con el fin de subsanar las falencias del método aquí presentado.

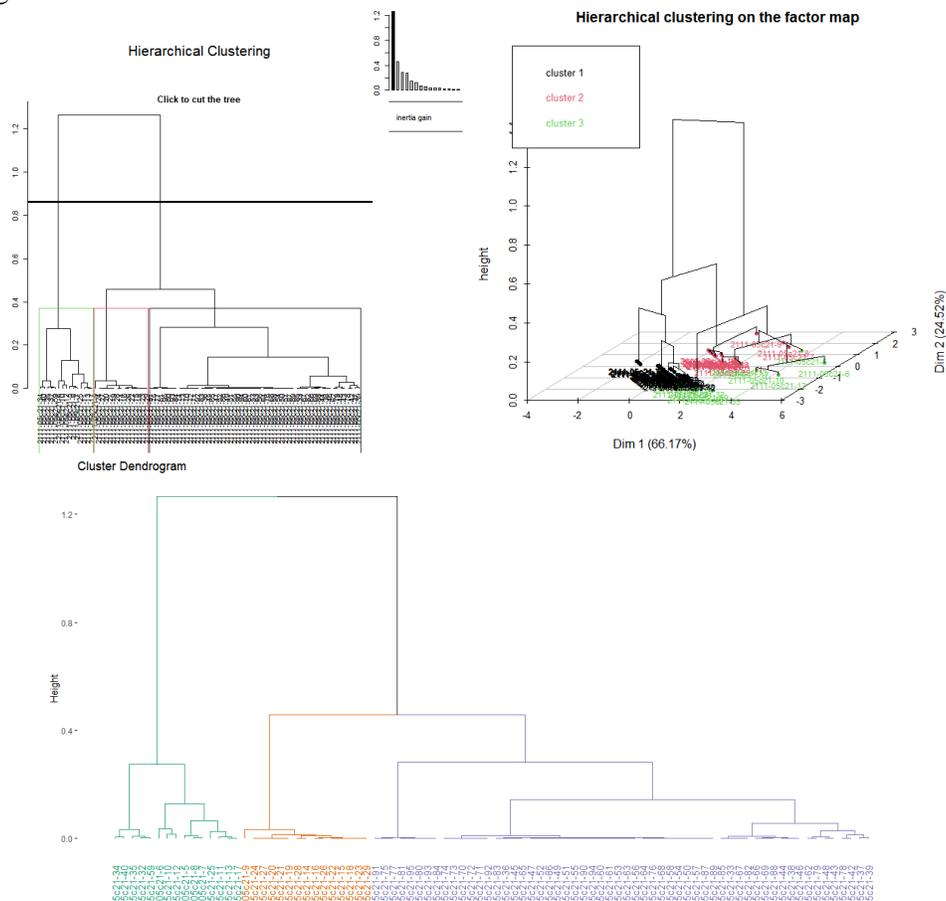
### *5.2.2. Análisis de Conglomerados Jerárquico (HC)*

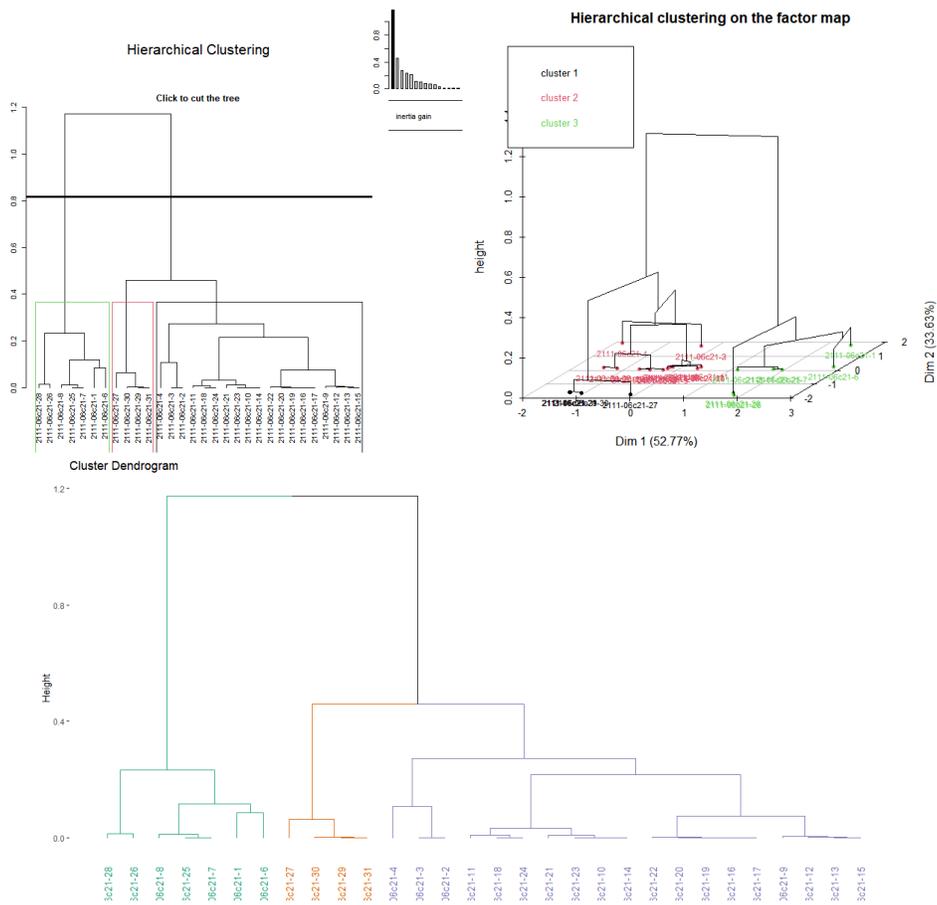
La clasificación de clústeres por medio de la técnica de Conglomerados Jerárquico – *Hierarchical Clustering*, HC– permite identificar clases a partir de la agregación o subdivisión de los individuos en diferentes grupos. Dicha división se elabora a partir de una estructura jerárquica que se muestra a manera de árbol en el dendograma. Para el objetivo planteado en esta investigación, la ventaja de la clasificación HC frente a la de K-Medias, es en primer que no requiere cumplir con la condición de esfericidad de los datos. En segundo lugar, permite distinguir el número de grupos o clases que retenga la mayor inercia con el menor número de clases posible, lo que facilita la identificación de los submercados por características intrínsecas de la vivienda a partir de los datos.

El análisis de conglomerados jerárquico inicia con el cálculo de la raíz o punto de inicio para las clasificaciones por división, o la agrupación de las observaciones más parecidas entre sí, una a una. El resultado de estos cálculos se muestra en el dendograma, en el cual la base está compuesta por todas las observaciones ordenadas por proximidad o similitud, mismas que a su vez se van agrupando hacia arriba. El siguiente paso del análisis consiste en elegir el punto a partir del cual se retendrán las clases, dicho de otro modo, se establece el punto a partir del cual se corta el árbol. El criterio de decisión se basa en los valores de similitud y disimilitud y la inercia, la cual se muestra en la altura de las ramas. Con el fin de ilustrar esta técnica, a continuación se muestran los resultados más destacados para Estados Unidos.

### Gráfica 14. Clústeres resultantes, clasificación por HC, Estados Unidos

US05 Eagle Pass





Como se aprecia en las gráficas anteriores, la clasificación por HC brinda mejores resultados que el análisis de K-Medias, incluso para conjuntos de datos como el de Eagle Pass y Laredo, los cuales mostraban problemas de agrupación al presentar una diferencia muy pequeña entre conglomerados y una similitud no tan alta entre las observaciones agrupadas, es decir, se carecía de una suficiente disimilitud externa y similitud interna. La única desventaja de la técnica de clasificación por HC, es que requiere de mayor poder de cómputo y tiende a ser más compleja en muestras grandes, no obstante, sus ventajas resultan determinantes para su selección. Por lo anterior, se concretó la clasificación de los submercados por características de las viviendas a partir del análisis HC para todos los mercados del área de estudio.

## VI. DISCUSIÓN Y CONCLUSIONES

---

El estudio de los mercados de vivienda en los espacios limítrofes entre México, Estados Unidos y Canadá supone un caso particular de mercados bajo múltiples jurisdicciones, los cuales exhiben similitudes y diferencias tanto intraregionales como interregionales. A su vez, los mercados de bienes raíces presentan condiciones de asimetría de la información entre los agentes, distintas en cada país y ciudad, lo cual complejiza el fenómeno. La condición de proximidad geográfica e interacción entre las ciudades fronterizas de América del Norte, es el marco a partir del cual se realizó esta investigación, en la cual se buscó identificar las particularidades de cada mercado de vivienda a partir de la identificación de sus submercados.

En primer lugar, para cada una de las veintiséis ciudades estudiadas, se estimó un modelo de precios hedónicos, el cual permiten captar la valoración de no mercado de las características de las viviendas. Al cálculo del modelo restringido por mínimos cuadrados ordinarios, le siguió la estimación del modelo semilogarítmico, con el fin de especificar la forma funcional y el modelo semilog con ruido blanco HC para obtener errores estándar robustos. A partir de los modelos antes señalados, se identificó la estructura de los datos que muestra las relaciones entre las variables. En general, las variables metros cuadrados de vivienda, número de cuartos y número de baños fueron significativas tanto en lo individual como en conjunto.

En segundo lugar, mediante la aplicación del análisis de componentes principales, se encontraron los submercados por características del vecindario a partir de un conjunto de variables que dan cuenta del entorno urbano, el cual está constituido por el espacio exterior de la vivienda en el cual se puede acceder a bienes y servicios para la satisfacción de las necesidades de los hogares. A partir de la identificación de los dos factores subyacentes que concentran la mayor información, se obtuvieron indicadores compuestos a partir de los cuales se elaboró una clasificación del espacio urbano en submercados de vivienda como expresión de las condiciones materiales y de accesibilidad de los vecindarios.

En tercer lugar, se realizó la clasificación por conglomerados, en primera instancia con la técnica de K-Medias. Al obtener resultados poco satisfactorios, se optó por realizar una nueva clasificación, esta vez mediante el análisis de Conglomerados Jerárquico, con el cual se logró identificar a los submercados por características intrínsecas de las viviendas. Es así el modelo de precios hedónicos o base, fue enriquecido con la incorporación de los dos tipos de submercados de vivienda. A su vez, las clasificaciones de submercados por características del vecindario y por atributos de las viviendas, permiten distinguir los rasgos más característicos de cada mercado.

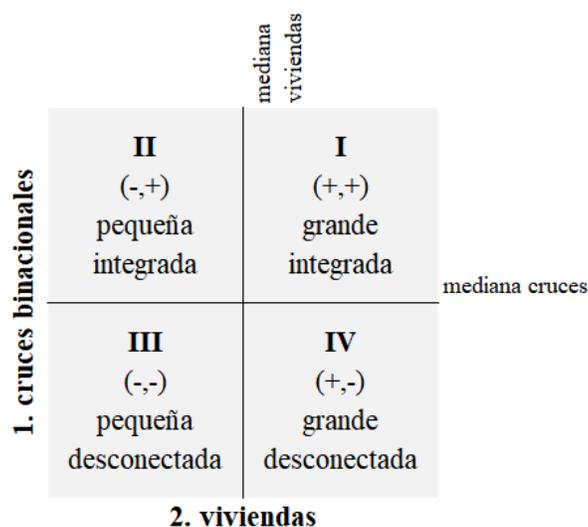
Debido a lo extenso de la investigación, en esta sección sólo se comentan algunos de los resultados generales a partir de los cuales se destacan diferencias y similitudes entre las ciudades del área de estudio. En particular, la revisión de los principales hallazgos relacionados con los objetivos de la investigación. En cuanto a la evaluación de los modelos de precios hedónicos, éste explica de mejor manera las relaciones entre las variables a partir de los datos para las ciudades mexicanas de Tijuana, Mexicali y Ciudad Juárez, las ciudades estadounidenses de San Diego, Brownsville, Detroit y Niagara Falls; y la ciudad canadienses de Metro Vancouver. Por el contrario, el modelo no presenta el mejor ajuste para explicar la valoración del precio de las viviendas en los mercados de Piedras Negras, Imperial, Whatcom y Windsor.

Respecto a los submercados de vivienda por características del entorno urbano, destacan como los factores más importantes para las ciudades mexicanas la cercanía a fuentes de empleo y consultorios médicos; para las ciudades estadounidenses el acceso por vía peatonal a bienes y servicios entendido como caminabilidad y el acceso a comercios; mientras que para Canadá destaca la proximidad a los centros de empleo y abastecimiento de alimentos.

Adicionalmente, la diversidad que presentan las ciudades fronterizas entre México, Estados Unidos y Canadá, permite estudiar el fenómeno de la vivienda bajo diferentes escenarios. Con el fin de identificar de manera más clara a estas conurbaciones binacionales y sus características, a continuación se propone una clasificación de las ciudades fronterizas seleccionadas dentro del área de estudio. Con ello se aporta un marco de referencia para el análisis y la comparación de pares de ciudades a partir de categorías que explicitan la

magnitud de la interacción trasfronteriza y el peso relativo de las comunidades en ambos lados de la frontera.

**Esquema 2. Plano bidimensional de clasificación de ciudades fronterizas**

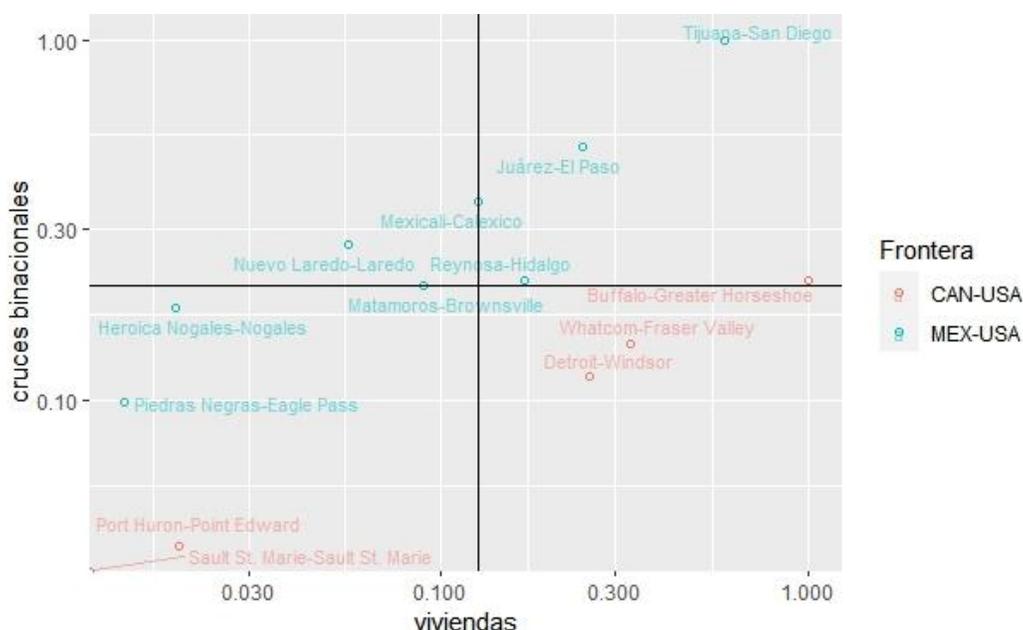


Fuente: Elaboración propia.

La clasificación se realizó a partir de dos criterios que obedecen a los principios de la economía espacial. En primer lugar, el nivel de interacción representado por el volumen de cruces de personas a través de las garitas terrestres. En segundo lugar, el peso a partir de la ponderación del tamaño del inventario de vivienda como variable aproximativa de la dimensión del mercado residencial. La combinación de ambos criterios se muestra en el esquema a manera de plano cartesiano que se muestra a continuación.

En una primera etapa, se definieron cada una de las cuatro tipologías tomando como puntos de referencia a las medianas de los dos criterios antes mencionados (ver esquema 2). Posteriormente, se graficaron en un plano bidimensional los valores del número de vivienda de ambos lados de la frontera en el eje de las abscisas y el volumen de cruces terrestres de personas en el eje de las ordenadas, como se muestra a continuación en la gráfica.

**Gráfica 15. Clasificación de ciudades fronterizas entre México, Estados Unidos y Canadá**

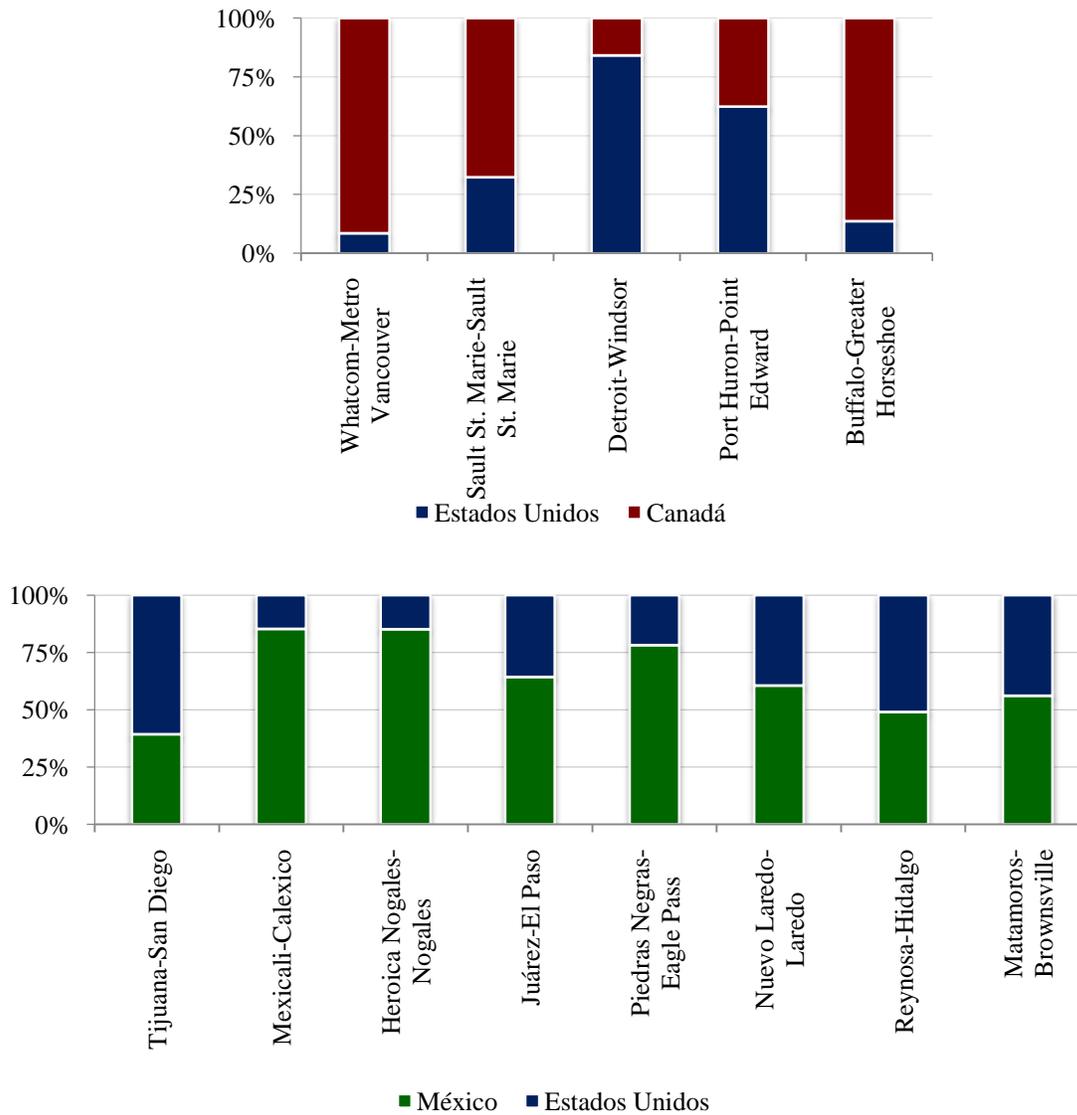


I. Grande e integrada	II. Pequeña e integrada	III. Pequeña y desconectada	III. Grande y desconectada
<ul style="list-style-type: none"> <li>•01Tijuana-San Diego</li> <li>•04Juárez-El Paso</li> <li>•13Buffalo-Greater Horseshoe</li> <li>•07Reynosa-Hidalgo</li> </ul>	<ul style="list-style-type: none"> <li>•02Mexicali-Calexico</li> <li>•06Nuevo Laredo-Laredo</li> <li>•08Matamoros-Brownsville</li> </ul>	<ul style="list-style-type: none"> <li>•03Heróica Nogales-Nogales</li> <li>•05Piedras Negras-Eagle Pass</li> <li>•12Port Huron-Point Edward</li> <li>•10Sault St. Marie-Sault St. Marie</li> </ul>	<ul style="list-style-type: none"> <li>•09Watcom-Metro Vancouver</li> <li>•11Detroit-Windsor</li> </ul>

Fuente: Elaboración propia con datos de BTS, INEGI, US Census y StatCan, 2010-2020.

En una segunda etapa, se ordenaron los pares de ciudades dentro de la categoría correspondiente al cuadrante en el que se ubica. Por otro lado, debido a que existen condiciones y características distintas para cada ciudad emparejada con su contraparte binacional, a continuación se distinguen dentro de cada conurbación fronteriza, la contribución de las ciudades de uno y otro lado de la frontera en el peso específico de las variables de interacción y tamaño, tomando en cuenta la proporción de población, hogares y viviendas de uno y otro lado de la frontera.

**Gráfica 16. Población de las conurbaciones fronterizas por país (%)**



En suma, la clasificación de ciudades fronterizas facilitará la identificación de las características distintivas de las ciudades fronterizas y sus mercados de vivienda de una forma que permita establecer de forma esquemática sus similitudes y diferencias. De este modo, permitirá dilucidar los resultados de los estudios de submercados de vivienda y asimetría de la información bajo un marco analítico común. Finalmente, la clasificación de ciudades propuesta a partir de las variables con escala de razón seleccionadas es sistemática, comparable, medible y replicable.

## VI. BIBLIOGRAFÍA

---

### **Libros:**

- Anselin, L. (2010). *Spatial Econometrics: Methods and Models* (Vol. Studies in Operational Regional Science). Amsterdam: Kluwer Academic Publishers.
- Anselin, L., Florax, R. J., & Rey, S. J. (2004). *Advances in Spatial Econometrics: Methodology, Tools and Applications*. Berlin: Springer-Verlag Berlin Heidelberg.
- Ball, M. (2006). *Markets and Institutions in Real Estate and Construction*. India: Blackwell Publishing Ltd.
- Capello, R. (2007). *Regional Economics*. New York: Routledge.
- Comisión para la Cooperación Ambiental. (1997). *Regiones Ecológicas de América del Norte Nivel I. Atlas de América del Norte*. Montreal.
- Demographia. (2021). *Demographia International Housing Affordability 2021 Edition*. Winnipeg: Urban Reform Institute and the Frontier Centre for Public Policy.
- Demographia. (2020). *Demographia International Housing Affordability Survey: 2020*. Belleville, IL: Demographia.
- Durlauf, S. N. (2004). Chapter 50. Neighborhood Effects. In V. J. Henderson, & J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (1 ed., Vol. 4 Cities and Geography, pp. 2173-2242). Amsterdam: North-Holland Elsevier.
- Fallis, G. (1985). *Housing Economics*. Toronto, Canada: Butterworth & Co. (Canada) Ltd.
- Fujita, M. (1989). *Urban economic theory: land use and city size*. New York: Cambridge University Press.
- Fujita, M., & Thisse, J.-F. (2002). *Economics of agglomeration: cities, industrial location, and regional growth*. UK: Cambridge University Press.
- Gibb, K. (2003). Chapter 2. Urban Housing Models. In T. O'Sullivan, & K. Gibbs (Eds.), *Housing Economics and Public Policy: Essays in honour of Duncan Maclennan* (pp. 22-37). Oxford: Blackwell Science.
- Gilbert, A. (2012). Rental Market and Rental Policies in Less Developed Countries. In S. J. Smith (Ed.), *International Encyclopedia of Housing and Home* (Vol. 6, pp. 35-45). Amsterdam, Netherlands: Elsevier.
- Haffner, M., Elsinga, M., & Hoekstra, J. (2012). Access and Affordability: Rent Regulation. In S. J. Smith (Ed.), *International Encyclopedia of Housing and Home* (Vol. 1, pp. 40-45). Amsterdam, Netherlands: Elsevier.
- International Joint Commission. (2020). *Transboundary Waters*. Washington & Ottawa.
- Isard, W. (1967). *Methods of Regional Analysis: an Introduction to Regional Science*. Cambridge, Massachusetts: The M.I.T. Press.
- Kandlbinder, K. (2018). The Role of Information in Real Estate Markets. *Dissertation, Doktors der Wirtschaftswissenschaft*, 109. Regensburg, Deutschland: Universität Regensburg.
- Malpezzi, S. (2003). Chapter 5. Hedonic pricing models: A selective and applied review. In T. O'Sullivan, & K. Gibb (Eds.), *Housing Economics and Public Policy: Essays in honour of Duncan Maclennan* (pp. 67-89). Oxford: Blackwell Science.
- North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

- OECD. (2017). *How's Life? 2017: Measuring Well-being*. Paris: OECD Publishing.
- ONU, A. G. (1948). Declaración Universal de los Derechos Humanos. (217 [III] A) . Paris.
- Quigley, Q. (2003). Chapter 4. Transactions Costs and Housing Markets. In T. O'Sullivan, & K. Gibb (Eds.), *Housing Economics and Public Policy: Essays in honour of Duncan MacLennan* (pp. 56-66). Oxford: Blackwell Science.
- Santos, M. (2009). Espacio y Método. Algunas reflexiones sobre el concepto de espacio. *Gestión y Ambiente* , 12 (1), 147-148.
- Takayama, T., & Labys, W. (1987). Chapter 5. Spatial Equilibrium Analysis. In P. Nijkamp (Ed.), *Handbook of Regional and Urban Economics* (Vol. 1, pp. 171-199). North Holland: Elsevier.
- Tokunaga, S. (1996). *Landownership and Residential Land Use in Urban Economies: Existence and Uniqueness of the Equilibrium*. Japón: Springer Japan.
- Tu, Y. (2003). Chapter 3. Segmentation, Adjustment and Disequilibrium. In T. O'Sullivan, & K. Gibbs (Eds.), *Housing Economics and Public Policy: Essays in honour of Duncan MacLennan* (pp. 38-55). Oxford: Blackwell Science.
- U.S. Census Bureau. (2011). *Statistical Abstract of the United States*.
- UBS. (2019). *UBS Global Real Estate Bubble Index 2019*. Chief Investment Office GWM. Zurich: UBS Switzerland AG.
- Weber, A. (1929). *Theory of the Location of Industries*. (C. J. Friedrich, Ed.) Chicago: University of Chicago Press.

### **Artículos**

- Akerlof, G. A. (1970). The Market for “Lemons”: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3), 488–500. <https://doi.org/10.2307/1879431>
- Alti, A., Engelberg, J., & Parsons, C. A. (2012). Information Spillovers and House Prices. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1786824>
- Arrow, K. J. (1963). Uncertainty and the Welfare Economics of Medical Care. *The American Economic Review*, 53(5), 941–973.
- Badarinza, C. (2018). The Informational Role of Housing Market Liquidity. *SSRN Electronic Journal*, 1–53. <https://doi.org/10.2139/ssrn.2880023>
- BARNETT, C. J. (1985). An Application of the Hedonic Price Model to the Perth Residential Land Market. *Economic Record*, 61(1), 476–481. <https://doi.org/10.1111/j.1475-4932.1985.tb01999.x>
- Bates, L. K. (2006). Does Neighborhood Really Matter? Comparing Historically Defined Neighborhood Boundaries with Housing Submarkets. *Journal of Planning Education and Research*, (26), 5–17. <https://doi.org/10.1177/0739456X05283254>
- Batool, F., & Hennig, C. (2021). Clustering with the Average Silhouette Width. *Computational Statistics and Data Analysis*, 158. <https://doi.org/10.1016/j.csda.2021.107190>
- Belasco, E., Farmer, M. C., & Lipscomb, C. A. (2012). Using a Finite Mixture Model of Heterogeneous Households to Delineate Housing Submarkets. *The Journal of Real Estate Research*, 34(4), 577–594. Retrieved from <https://www.jstor.org/stable/24888423?seq=1>
- Bhagwati, J. N. (1984). Why are Services Cheaper in the Poor Countries? *The Economic*

- Journal*, 94(374), 279–286. <https://doi.org/10.2307/2232350>
- Black, S. E. (1999). Do Better Schools Matter? Parental Valuation of Elementary Education. *The Quarterly Journal of Economics*, 114(2), 577–599. <https://doi.org/10.1162/003355399556070>
- Bourassa, S. C., Hoesli, M., & Peng, V. S. (2003). Do housing submarkets really matter? *Journal of Housing Economics*, 12, 12–28. [https://doi.org/10.1016/S1051-1377\(03\)00003-2](https://doi.org/10.1016/S1051-1377(03)00003-2)
- Brito, M., Macias, J., Ramirez, L., Jacquin, C., & Zubicaray, G. (2021). *Índice De Desigualdad Urbana*. WRI Mexico. Ciudad de México. Retrieved from <https://wrimexico.org/publication/indice-de-desigualdad-urbana>
- Butsic, V., Hanak, E., & Valletta, R. (2011). Climate Change and Housing Prices: Hedonic Estimates for Ski Resorts in Western North America. *Land Economics*, 87(1), 75–91. Retrieved from <https://www.jstor.org/stable/27920304?seq=1>
- Calanog, V., Lynford, L., & Yang, H. (2017). In the Age of Global Cities, Submarkets Still Matter. *Real Estate Issues*, 141–147.
- Cattell, R. B. (1996). The Scree Test For The Number Of Factors. [https://doi.org/10.1207/S15327906mbr0102\\_10](https://doi.org/10.1207/S15327906mbr0102_10), 1(2), 245–276. [https://doi.org/10.1207/S15327906MBR0102\\_10](https://doi.org/10.1207/S15327906MBR0102_10)
- Chau, K. W., & Choy, L. H. T. (2011). Let the Buyer or Seller Beware: Measuring Lemons in the Housing Market under Different Doctrines of Law Governing Transactions and Information. *The Journal of Law and Economics*, 54(S4), S347–S365. <https://doi.org/10.1086/661941>
- Chinco, A., & Mayer, C. (2016). Misinformed speculators and mispricing in the housing market. *Review of Financial Studies*, 29(2), 486–522. <https://doi.org/10.1093/rfs/hhv061>
- Chiu, H. C., Hsieh, Y. I. C., & Kao, C. Y. I. (2005). Website quality and customer's behavioural intention: An exploratory study of the role of information asymmetry. *Total Quality Management and Business Excellence*, 16(2), 185–197. <https://doi.org/10.1080/14783360500054277>
- Claycamp, H. J., & Massy, W. F. (1968). A Theory of Market Segmentation. *Journal of Marketing Research*, 5(4), 388. <https://doi.org/10.2307/3150263>
- Court, A. T. (1939). Hedonic price indexes with automotive examples. In *The dynamics of automobile demand* (pp. 99–119). New York. Retrieved from [http://www.econ.wayne.edu/agoodman/research/pubs/court\\_hedonic.pdf](http://www.econ.wayne.edu/agoodman/research/pubs/court_hedonic.pdf)
- Culbertson, J. M. (1957). The Term Structure of Interest Rates. *The Quarterly Journal of Economics*, 71(4), 485–517.
- Dale-Johnson, D. (1982). An alternative approach to housing market segmentation using hedonic price data. *Journal of Urban Economics*, 11(3), 311–332. [https://doi.org/10.1016/0094-1190\(82\)90078-X](https://doi.org/10.1016/0094-1190(82)90078-X)
- Eerola, E., & Lyytikäinen, T. (2015). On the role of public price information in housing markets. *Regional Science and Urban Economics*, 53, 74–84. <https://doi.org/10.1016/j.regsciurbeco.2015.05.006>
- Ellen, I. (2012). Economics of housing market segmentation. In S. J. Smith (Ed.), *International Encyclopedia of Housing and Home* (pp. 51–54). Elsevier. <https://doi.org/10.1016/B978-0-08-047163-1.00103-X>
- Ellickson, B. (1981). An alternative test of the hedonic theory of housing markets. *Journal of Urban Economics*, 9(1), 56–79. [https://doi.org/10.1016/0094-1190\(81\)90048-6](https://doi.org/10.1016/0094-1190(81)90048-6)

- Fereidouni, H. G., Al-MMulali, U., Lee, J. Y. M., & Mohammed, A. H. (2016). Dynamic Relationship between House Prices in Malaysia's Major Economic Regions and Singapore House Prices. *Regional Studies*, 50(4), 657–670. <https://doi.org/10.1080/00343404.2014.928408>
- Fischer, M. M., Hewings, G. J. D., Snickars, F., Editor, C., Fischer, M. M., & Eds, G. P. (1998). *Advances in Spatial Science*.
- Freemark, Y. (2019). Doubling housing production in the Paris region: a multi-policy, multi-jurisdictional response. *International Journal of Housing Policy*, 1–15. <https://doi.org/10.1080/19491247.2019.1682233>
- Fullerton, T. M., & Bujanda, A. (2017). Commercial property values in a border metropolitan economy. *Asia-Pacific Journal of Regional Science*. <https://doi.org/10.1007/s41685-017-0065-x>
- Galster, G. (1996). William Grigsby and the analysis of housing sub-markets and filtering. *Urban Studies*, 33(10), 1797–1805. <https://doi.org/10.1080/0042098966376>
- Galster, G., & Rothenberg, J. (1991). Filtering in Urban Housing: A Graphical Analysis of a Quality-Segmented Market. *Journal of Planning Education and Research*, 11(1), 37–50. <https://doi.org/10.1177/0739456X9101100106>
- Garmaise, M. J., & Moskowitz, T. J. (2004). Confronting Information Asymmetries: Evidence from Real Estate Markets. *The Review of Financial Studies*, 17(2), 405–437. <https://doi.org/10.1093/rfs/hhg037>
- Goodman, A. C. (1978). Hedonic prices, price indices and housing markets. *Journal of Urban Economics*, 5(4), 471–484. [https://doi.org/10.1016/0094-1190\(78\)90004-9](https://doi.org/10.1016/0094-1190(78)90004-9)
- Goodman, A. C. (1998). Andrew Court and the Invention of Hedonic Price Analysis. *Journal of Urban Economics*, 44, 291–298. Retrieved from <http://www.econ.wayne.edu/agoodman/research/pubs/Deep/Court-JUE.pdf>
- Goodman, A. C., & Thibodeau, T. G. (2003). Housing market segmentation and hedonic prediction accuracy. *Journal of Housing Economics*, 12(3), 181–201. [https://doi.org/10.1016/S1051-1377\(03\)00031-7](https://doi.org/10.1016/S1051-1377(03)00031-7)
- Griliches, Z. (1961). Hedonic Price Indexes for Automobiles: An Econometric of Quality Change. In Report of the Price Statistics Review Committee (Ed.), *The Price Statistics of the Federal Government* (pp. 173–196). Massachusetts: National Bureau of Economic Research. Retrieved from <http://www.nber.org/books/repo61-1>
- Helgers, R., & Buyst, E. (2016). Spatial and Temporal Diffusion of Housing Prices in the Presence of a Linguistic Border: Evidence from Belgium. *Spatial Economic Analysis*, 11(1), 92–122. <https://doi.org/10.1080/17421772.2016.1102961>
- Herzog, L. A., & Sohn, C. (2017). The co-mingling of bordering dynamics in the San Diego–Tijuana cross-border metropolis. *Territory, Politics, Governance*, (May 2017), 1–24. <https://doi.org/10.1080/21622671.2017.1323003>
- Holmes, C., & Xie, J. (2018). Distortions in Real Estate Transactions with Out-of-State Participants. *Journal of Real Estate Finance and Economics*, 57(4), 592–617. <https://doi.org/10.1007/s11146-017-9624-1>
- Honkanen, A., Pitkänen, K., & Hall, M. C. (2016). A Local Perspective on Cross-Border Tourism. Russian Second Home Ownership in Eastern Finland. *International Journal of Tourism Research*, 18(2), 149–158. <https://doi.org/10.1002/jtr.2041>
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24(6), 417–441. <https://doi.org/10.1037/h0071325>

- Josse, J., & Husson, F. (2016). missMDA: A package for handling missing values in multivariate data analysis. *Journal of Statistical Software*, 70(1), 1–31. <https://doi.org/10.18637/jss.v070.i01>
- Kaiser, H. F. (1991). Coefficient Alpha for a Principal Component and the Kaiser-Guttman Rule. *Psychological Reports*, 68(3). <https://doi.org/10.2466/pr0.1991.68.3.855>
- Kandlbinder, K. (2017). Leveling the playing field: Information Efficiency in US Housing Markets over time. In *24th Annual European Real Estate Society Conference. ERES: Conference*. Delft, Netherlands: Stichting European Real Estate Society. [https://doi.org/10.15396/eres2017\\_134](https://doi.org/10.15396/eres2017_134)
- Kandlbinder, K., Miller, N. G., & Sklarz, M. (2019). Leveling the playing field: out-of-town buyer premiums in US housing markets over time. *International Journal of Housing Markets and Analysis*, 12(3), 377–404. <https://doi.org/10.1108/IJHMA-02-2018-0017>
- Kaufman, L., & Rousseeuw, P. J. (1990). *Finding Groups in Data: An Introduction to Cluster Analysis* (1st ed.). NJ, United States: John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470316801.CH1>
- Keskin, B., & Watkins, C. (2017). Defining spatial housing submarkets: Exploring the case for expert delineated boundaries. *Urban Studies*, 54(6), 1446–1462. <https://doi.org/10.1177/0042098015620351>
- Kincal, G., Fullerton, T. M., Holcomb, J. H., & Barraza de Anda, M. P. (2010). Cross Border Business Cycle Impacts on the El Paso Housing Market. *Estudios Fronterizos*, 11(22), 49–77. Retrieved from [http://www.scielo.org.mx/scielo.php?script=sci\\_arttext&pid=S0187-69612010000200002&lng=es&nrm=iso&tlng=en](http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0187-69612010000200002&lng=es&nrm=iso&tlng=en)
- Kravis, I. B., & Lipsey, R. E. (1983). *Toward an Explanation of National Price Levels* (No. 52). *Princeton Studies in International Finance*. Princeton.
- Lancaster, K. J. (1966). A New Approach to Consumer Theory. *The Journal of Political Economy*, 74(2), 132–157. Retrieved from <http://www.jstor.org/stable/1828835>
- Lee, S., & Lin, J. (2018). Natural amenities, neighbourhood dynamics, and persistence in the spatial distribution of income. *Review of Economic Studies*, 85(1), 663–694. <https://doi.org/10.1093/restud/rdx018>
- Lee, Y. S., & Sasaki, Y. (2018). Information technology in the property market. *Information Economics and Policy*, 44, 1–7. <https://doi.org/10.1016/j.infoecopol.2018.07.002>
- Levitt, S. D., & Syverson, C. (2008). Market distortions when agents are better informed: The value of information in real estate transactions. *Review of Economics and Statistics*, 90(4), 599–611. Retrieved from <http://www.mlsmi.com>.
- Lloyd, S. P. (1982). Least Squares Quantization in PCM. *IEEE Transactions on Information Theory*, 28(2). <https://doi.org/10.1109/TIT.1982.1056489>
- Lucas, R. E. B. (1975). Hedonic Price Functions. *Economic Inquiry*, 13(2), 157–178. <https://doi.org/10.1111/j.1465-7295.1975.tb00985.x>
- Lucas, S., & Purkayastha, B. (2007). “Where is home?” Here and there: Transnational experiences of home among Canadian migrants in the United States. *GeoJournal*, 68(2–3), 243–251. <https://doi.org/10.1007/s10708-007-9073-0>
- Micheli, M., Rouwendal, J., & Dekkers, J. (2019). Border Effects in House Prices. *Real Estate Economics*, 47(3), 757–783. <https://doi.org/10.1111/1540-6229.12255>
- Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining

- the number of clusters in a data set. *Psychometrika*, 50(2).  
<https://doi.org/10.1007/BF02294245>
- Mills, E. S., & MacKinnon, J. (1973). Notes on the New Urban Economics. *The Bell Journal of Economics and Management Science*, 4(2), 601.  
<https://doi.org/10.2307/3003054>
- Norman, L. M., Feller, M., & Guertin, D. P. (2009). Forecasting urban growth across the United States–Mexico border. *Computers, Environment and Urban Systems*, 33, 150–159. <https://doi.org/10.1016/j.compenvurbsys.2008.10.003>
- Páez, A., & Scott, D. M. (2004). Spatial statistics for urban analysis: A review of techniques with examples. *GeoJournal*, 61, 53–67.
- Pagès, J. (2004). Analyse Factorielle de Données Mixtes. *Revue de Statistique Appliquée*, 52(4), 93–111.
- Ramírez Franco, L. D. (2008). Revisión teórica de la estructura espacial y la identificación de subcentros de empleo. *Ensayos de Economía*, (33), 141–178.
- Rérat, P., Moine, A., Gertsch, K., & Signoret, P. (2011). La mobilité résidentielle transfrontalière dans l'Arc jurassien franco-suisse. *Géo-Regards: Revue Neuchâteloise de Géographie*, 4, 153–167. Retrieved from <https://halshs.archives-ouvertes.fr/halshs-00763253>
- Richardson, H. W. (1976). The new urban economics: An evaluation. *Socio-Economic Planning Sciences*, 10(4), 137–147. [https://doi.org/10.1016/0038-0121\(76\)90014-8](https://doi.org/10.1016/0038-0121(76)90014-8)
- Roberts, S. (1991). A critical evaluation of the city life cycle idea. *Urban Geography*, 12(5), 431–449. <https://doi.org/10.2747/0272-3638.12.5.431>
- Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *The Journal of Political Economy*, 82(1), 34–55. Retrieved from <http://www.jstor.org/stable/1830899>
- Rothschild, M., & Stiglitz, J. (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *Quarterly Journal of Economics*, 90(4), 629–649.
- Rubin-Kurtzman, J. R., Ham-Chande, R., & Arsdol, M. D. Van. (1996). Population in Trans-Border Regions: The Southern California-Baja California Urban System. *The International Migration Review*, 30(4), 1020–1045. Retrieved from <http://www.jstor.org/stable/2547602>
- Shonkwiler, J., & Reynolds, J. (1986). A note on the use of hedonic price models in the analysis of land prices at the urban fringe. *Land Economics*, 62(1), 58–63. Retrieved from <http://www.jstor.org/stable/3146563>
- Soaita, A. M., Maclaren, D., & Gibb, K. D. B. (2020). The social and economic consequences of housing wealth inequalities. *The Quarterly Journal of the International Union for Housing Finance*, 35(2), 27–35.
- Sobek, M. (2020). The IPUMS approach to harmonizing the world's population census data. In *Studies in Classification, Data Analysis, and Knowledge Organization* (pp. 495–511). Springer Science and Business Media Deutschland GmbH. [https://doi.org/10.1007/978-981-15-3311-2\\_39](https://doi.org/10.1007/978-981-15-3311-2_39)
- Sobek, M., & Cleveland, L. L. (2017). *IPUMS Approach to Harmonizing International Census and Survey Data* (No. 31). Geneva.
- Spence, M. (1973). Job Market Signaling. *Quarterly Journal of Economics*, 87(3), 355–374. <https://doi.org/10.2307/1882010>
- Spence, M. (1981). Signaling, Screening, and Information. In S. Rosen (Ed.), *Studies in*

- Labor Markets* (pp. 319–358). Chicago: University of Chicago Press. Retrieved from <http://www.nber.org/chapters/c8915>
- Stevenson, S. (2004). House price diffusion and inter-regional and cross-border house price dynamics. *Journal of Property Research*, 21(4), 301–320. <https://doi.org/10.1080/09599910500151228>
- Stigler, G. J. (1961). The Economics of Information. *The Journal of Political Economy*, 69(3), 213–225. Retrieved from <https://www.jstor.org/stable/1829263>
- Stiglitz, J. E. (1975). The Theory of “Screening”, Education, and the Distribution of Income. *American Economic Review*, 65(3), 283–300. Retrieved from <https://www.jstor.org/stable/1804834>
- Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 63(2). <https://doi.org/10.1111/1467-9868.00293>
- Tu, Y. (1997). The local housing sub-market structure and its properties. *Urban Studies*, 34(2), 337–353. <https://doi.org/10.1080/0042098976203>
- Tu, Y., Sun, H., & Yu, S. M. (2007). Spatial autocorrelations and urban housing market segmentation. *Journal of Real Estate Finance and Economics*, 34(3), 385–406. <https://doi.org/10.1007/s11146-007-9015-0>
- Watkins, C. A. (2001). The definition and identification of housing submarkets. *Environment and Planning A: Economy and Space*, 33, 2235–2253. <https://doi.org/10.1068/a34162>
- Wexler, M. E. (1996). A Comparison of Canadian and American Housing Policies. *Urban Studies*, 33(10), 1909–1921. <https://doi.org/10.1080/0042098966457>
- Williamson, O. E. (1979). Transaction-Cost Economics: The Governance of Contractual Relations. *The Journal of Law and Economics*, 22(2), 233–261. Retrieved from <https://www.jstor.org/stable/725118?seq=1>
- Williamson, S. D. (1986). Costly monitoring, financial intermediation, and equilibrium credit rationing. *Journal of Monetary Economics*, 18(2), 159–179. [https://doi.org/10.1016/0304-3932\(86\)90074-7](https://doi.org/10.1016/0304-3932(86)90074-7)
- Wilson, P., White, M., Dunse, N., Cheong, C., & Zurbrugg, R. (2011). Modelling Price Movements in Housing Micro Markets: Identifying Long-term Components in Local Housing Market Dynamics. *Urban Studies*, 48(9), 1853–1874. <https://doi.org/10.1177/0042098010380960>
- Wu, C., & Sharma, R. (2012). Housing submarket classification: The role of spatial contiguity. *Applied Geography*, 32, 746–756. Retrieved from [https://ac.els-cdn.com/S0143622811001640/1-s2.0-S0143622811001640-main.pdf?\\_tid=ef110ad0-085f-11e8-9c52-00000aacb360&acdnat=1517607065\\_22fc7296d8bff0575075091afd747329](https://ac.els-cdn.com/S0143622811001640/1-s2.0-S0143622811001640-main.pdf?_tid=ef110ad0-085f-11e8-9c52-00000aacb360&acdnat=1517607065_22fc7296d8bff0575075091afd747329)
- Xie, J. (2018). Who Is “Misleading” Whom in Real Estate Transactions? *Real Estate Economics*, 46(3), 527–558. <https://doi.org/10.1111/1540-6229.12196>
- Young, S. (2013). Transaction Cost Economics. In S. O. Idowu, N. Capaldi, L. Zu, & A. Das Gupta (Eds.), *Encyclopedia of Corporate Social Responsibility* (pp. 2547–2552). Berlin, Heidelberg: Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-28036-8\\_221](https://doi.org/10.1007/978-3-642-28036-8_221)
- Zhao, C., Jensen, J., & Zhan, B. (2017). A comparison of urban growth and their influencing factors of two border cities: Laredo in the US and Nuevo Laredo in Mexico. *Applied Geography*, 79, 223–234.

<https://doi.org/10.1016/j.apgeog.2016.12.017>

Zhou, X., Gibler, K., & Zahirovic-Herbert, V. (2015). Asymmetric buyer information influence on price in a homogeneous housing market. *Urban Studies*, 52(5), 891–905. <https://doi.org/10.1177/0042098014529464>

Zhou, Y., & Guo, H. (2015). Is Shenzhen Housing Price Bubble that High? A Perspective of Shenzhen Hong Kong Cross-Border Integration. *International Real Estate Review*, 18(3), 365–382. Retrieved from <https://ideas.repec.org/a/ire/issued/v18n032015p365-382.html>

### **Repositorios digitales:**

Cox, Murray (Noviembre 2020). About Inside Airbnb [repositorio digital]. Inside Airbnb. Recuperado de: <http://insideairbnb.com/get-the-data.html> [18 de noviembre de 2020].

Slee, Tom (Noviembre 2020). Airbnb Data Collection Methodology and Accuracy [repositorio digital]. Airbnb Data Collection. Recuperado de: <http://tomslee.net/airbnb-data-collection-methodology-and-accuracy> [18 de noviembre de 2020].

Minnesota Population Center. “IPUMS, Integrated Public Use Microdata Series, International: Version 7.3” [dataset]. Minneapolis, MN: IPUMS, 2020. <https://doi.org/10.18128/D020.V7.3>

IPUMS (2020) “Harmonization”, Minneapolis: IPUMS International. Recuperado de: <https://international.ipums.org/international/harmonization.shtml>

OECD (2019). “Indicators on the Housing Market”, OECD Affordable Housing Database [dataset]. Recuperado de: <https://www.oecd.org/housing/data/affordable-housing-database/>

OECD (2019). “Indicators of Housing Conditions and Affordability”, OECD Affordable Housing Database [dataset]. Recuperado de: <https://www.oecd.org/housing/data/affordable-housing-database/>

## VII. ANEXOS

### MX01 Tijuana

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 7157

Included observations: 7157

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-367.1062	20.49400	-17.91287	0.0000
SQM	0.000676	5.86E-05	11.53199	0.0000
ROOMS	-0.088596	0.012918	-6.858199	0.0000
BATHS	0.438558	0.013223	33.16743	0.0000
CARS	0.201590	0.013988	14.41132	0.0000
LAT	0.043357	0.117464	0.369108	0.7121
LONG	-3.216725	0.159535	-20.16314	0.0000
R-squared	0.431176	Mean dependent var		11.94771
Adjusted R-squared	0.430698	S.D. dependent var		0.894776
S.E. of regression	0.675127	Akaike info criterion		2.053146
Sum squared resid	3258.946	Schwarz criterion		2.059871
Log likelihood	-7340.184	Hannan-Quinn criter.		2.055461
F-statistic	903.2972	Durbin-Watson stat		1.101436
Prob(F-statistic)	0.000000	Wald F-statistic		763.4800
Prob(Wald F-statistic)	0.000000			

### US01 San Diego

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 1602

Included observations: 1602

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-152.0360	13.24606	-11.47783	0.0000
SQM	0.005165	0.000309	16.73917	0.0000
ROOMS	0.074133	0.016093	4.606565	0.0000
BATHS	0.016309	0.021770	0.749168	0.4539
LAT	-0.476493	0.071232	-6.689273	0.0000
LONG	-1.538682	0.123962	-12.41257	0.0000
R-squared	0.489159	Mean dependent var		13.45374
Adjusted R-squared	0.487558	S.D. dependent var		0.572264
S.E. of regression	0.409655	Akaike info criterion		1.056736
Sum squared resid	267.8365	Schwarz criterion		1.076882
Log likelihood	-840.4457	Hannan-Quinn criter.		1.064217
F-statistic	305.6515	Durbin-Watson stat		1.715404
Prob(F-statistic)	0.000000	Wald F-statistic		227.5339
Prob(Wald F-statistic)	0.000000			

MX02 Mexicali

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 1348

Included observations: 1348

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	191.1504	45.27414	4.222066	0.0000
SQM	0.000683	0.000134	5.077508	0.0000
ROOMS	0.204926	0.034551	5.931083	0.0000
BATHS	0.441287	0.026713	16.51969	0.0000
CARS	0.021683	0.026725	0.811330	0.4173
LAT	0.683865	0.174439	3.920377	0.0001
LONG	1.765225	0.436303	4.045868	0.0001
R-squared	0.508606	Mean dependent var		11.45552
Adjusted R-squared	0.506407	S.D. dependent var		0.870663
S.E. of regression	0.611694	Akaike info criterion		1.860011
Sum squared resid	501.7619	Schwarz criterion		1.887047
Log likelihood	-1246.648	Hannan-Quinn criter.		1.870137
F-statistic	231.3285	Durbin-Watson stat		1.376756
Prob(F-statistic)	0.000000	Wald F-statistic		220.2286
Prob(Wald F-statistic)	0.000000			

US02 Imperial

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 48

Included observations: 48

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.734438	29.27874	0.093393	0.9260
SQM	0.005798	0.001578	3.673276	0.0007
ROOMS	-0.009982	0.068358	-0.146033	0.8846
BATHS	-0.052021	0.050952	-1.020986	0.3131
LAT	-0.078193	0.445730	-0.175427	0.8616
LONG	-0.100190	0.306887	-0.326471	0.7457
R-squared	0.586566	Mean dependent var		12.53137
Adjusted R-squared	0.537347	S.D. dependent var		0.556047
S.E. of regression	0.378215	Akaike info criterion		1.009761
Sum squared resid	6.007959	Schwarz criterion		1.243661
Log likelihood	-18.23427	Hannan-Quinn criter.		1.098152
F-statistic	11.91761	Durbin-Watson stat		2.287545
Prob(F-statistic)	0.000000	Wald F-statistic		37.90074
Prob(Wald F-statistic)	0.000000			

### MX03 Nogales

Dependent Variable: LOG(USD\_PRICE)

Method: Least Squares

Sample: 1 134

Included observations: 134

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-477.5604	319.4586	-1.494906	0.1374
SQM	0.002275	0.000677	3.359466	0.0010
ROOMS	0.177988	0.086388	2.060329	0.0414
BATHS	0.238199	0.067282	3.540327	0.0006
CARS	0.304750	0.104422	2.918445	0.0042
LAT	-5.509319	3.337841	-1.650563	0.1013
LONG	-5.940673	2.837648	-2.093520	0.0383
R-squared	0.684738	Mean dependent var		10.79431
Adjusted R-squared	0.669843	S.D. dependent var		0.957837
S.E. of regression	0.550366	Akaike info criterion		1.694357
Sum squared resid	38.46860	Schwarz criterion		1.845737
Log likelihood	-106.5219	Hannan-Quinn criter.		1.755873
F-statistic	45.97320	Durbin-Watson stat		0.879701
Prob(F-statistic)	0.000000	Wald F-statistic		37.67484
Prob(Wald F-statistic)	0.000000			

### US03 Nogales

Dependent Variable: LOG(USD\_PRICE)

Method: Least Squares

Sample: 1 88

Included observations: 88

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-53.72205	22.25350	-2.414094	0.0180
SQM	0.002643	0.000560	4.721811	0.0000
ROOMS	0.010468	0.039864	0.262601	0.7935
BATHS	0.075317	0.076456	0.985103	0.3275
LAT	2.245275	0.350396	6.407814	0.0000
LONG	0.045646	0.193832	0.235494	0.8144
R-squared	0.766942	Mean dependent var		12.70752
Adjusted R-squared	0.752732	S.D. dependent var		0.439458
S.E. of regression	0.218525	Akaike info criterion		-0.138083
Sum squared resid	3.915773	Schwarz criterion		0.030826
Log likelihood	12.07566	Hannan-Quinn criter.		-0.070034
F-statistic	53.96889	Durbin-Watson stat		2.188800
Prob(F-statistic)	0.000000	Wald F-statistic		47.36822
Prob(Wald F-statistic)	0.000000			

MX04 Ciudad Juárez

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 1396

Included observations: 1396

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	370.3398	49.51459	7.479409	0.0000
SQM	0.001526	0.000119	12.77581	0.0000
ROOMS	-0.067720	0.013773	-4.916945	0.0000
BATHS	0.412038	0.021926	18.79246	0.0000
CARS	0.109121	0.020177	5.408300	0.0000
LAT	7.390916	0.430945	17.15048	0.0000
LONG	5.586111	0.496623	11.24820	0.0000
R-squared	0.733369	Mean dependent var		11.36745
Adjusted R-squared	0.732217	S.D. dependent var		1.057980
S.E. of regression	0.547481	Akaike info criterion		1.638023
Sum squared resid	416.3321	Schwarz criterion		1.664305
Log likelihood	-1136.340	Hannan-Quinn criter.		1.647849
F-statistic	636.7416	Durbin-Watson stat		1.159193
Prob(F-statistic)	0.000000	Wald F-statistic		788.7185
Prob(Wald F-statistic)	0.000000			

US04 El Paso

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 818

Included observations: 818

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	14.23835	6.689671	2.128408	0.0336
SQM	0.004032	0.000273	14.74603	0.0000
ROOMS	0.049911	0.017193	2.902966	0.0038
BATHS	0.114552	0.022443	5.104094	0.0000
LAT	0.094525	0.125965	0.750405	0.4532
LONG	0.056986	0.079707	0.714949	0.4748
R-squared	0.678803	Mean dependent var		12.38007
Adjusted R-squared	0.676825	S.D. dependent var		0.374632
S.E. of regression	0.212973	Akaike info criterion		-0.247996
Sum squared resid	36.83025	Schwarz criterion		-0.213471
Log likelihood	107.4303	Hannan-Quinn criter.		-0.234747
F-statistic	343.2083	Durbin-Watson stat		1.298264
Prob(F-statistic)	0.000000	Wald F-statistic		253.3218
Prob(Wald F-statistic)	0.000000			

MX05 Piedras Negras

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 49

Included observations: 49

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-300.8363	560.3626	-0.536860	0.5942
SQM	0.000260	0.000314	0.827583	0.4126
ROOMS	-0.107318	0.136017	-0.789006	0.4345
BATHS	0.469703	0.096089	4.888228	0.0000
CARS	0.058740	0.271104	0.216670	0.8295
LAT	4.481762	4.169429	1.074910	0.2886
LONG	-1.816644	5.228287	-0.347464	0.7300
R-squared	0.365949	Mean dependent var		11.29336
Adjusted R-squared	0.275371	S.D. dependent var		0.809549
S.E. of regression	0.689130	Akaike info criterion		2.224789
Sum squared resid	19.94578	Schwarz criterion		2.495049
Log likelihood	-47.50733	Hannan-Quinn criter.		2.327325
F-statistic	4.040129	Durbin-Watson stat		0.965120
Prob(F-statistic)	0.002770	Wald F-statistic		9.142167
Prob(Wald F-statistic)	0.000002			

US05 Eagle Pass

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 88

Included observations: 88

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	194.0915	140.5884	1.380566	0.1712
SQM	0.002703	0.000864	3.127297	0.0024
ROOMS	-0.220082	0.065436	-3.363318	0.0012
BATHS	0.291608	0.099477	2.931405	0.0044
LAT	1.966391	1.395274	1.409323	0.1625
LONG	2.376118	1.753484	1.355085	0.1791
R-squared	0.428321	Mean dependent var		12.15967
Adjusted R-squared	0.393462	S.D. dependent var		0.454584
S.E. of regression	0.354032	Akaike info criterion		0.826890
Sum squared resid	10.27780	Schwarz criterion		0.995799
Log likelihood	-30.38314	Hannan-Quinn criter.		0.894939
F-statistic	12.28742	Durbin-Watson stat		1.066951
Prob(F-statistic)	0.000000	Wald F-statistic		14.91409
Prob(Wald F-statistic)	0.000000			

MX06 Nuevo Laredo

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 142

Included observations: 142

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-551.6434	257.8582	-2.139328	0.0342
SQM	-0.003023	0.001429	-2.115814	0.0362
ROOMS	1.396349	0.090701	15.39515	0.0000
BATHS	0.192119	0.109112	1.760748	0.0805
CARS	0.193970	0.157417	1.232207	0.2200
LAT	-5.907522	2.593840	-2.277520	0.0243
LONG	-7.240564	2.815303	-2.571859	0.0112
R-squared	0.831756	Mean dependent var		8.902570
Adjusted R-squared	0.824279	S.D. dependent var		0.968449
S.E. of regression	0.405965	Akaike info criterion		1.082940
Sum squared resid	22.24903	Schwarz criterion		1.228650
Log likelihood	-69.88873	Hannan-Quinn criter.		1.142150
F-statistic	111.2346	Durbin-Watson stat		1.894736
Prob(F-statistic)	0.000000	Wald F-statistic		1347.747
Prob(Wald F-statistic)	0.000000			

US06 Laredo

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 30

Included observations: 30

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	174.7906	169.4122	1.031747	0.3125
SQM	0.002299	0.000873	2.633059	0.0146
ROOMS	-0.103208	0.056098	-1.839777	0.0782
BATHS	-0.010583	0.039026	-0.271168	0.7886
LAT	1.330159	0.475027	2.800174	0.0099
LONG	1.999074	1.759740	1.136005	0.2672
R-squared	0.675666	Mean dependent var		12.65321
Adjusted R-squared	0.608097	S.D. dependent var		0.264407
S.E. of regression	0.165525	Akaike info criterion		-0.582536
Sum squared resid	0.657563	Schwarz criterion		-0.302296
Log likelihood	14.73804	Hannan-Quinn criter.		-0.492885
F-statistic	9.99579	Durbin-Watson stat		1.720510
Prob(F-statistic)	0.000029	Wald F-statistic		28.97212
Prob(Wald F-statistic)	0.000000			

MX07 Reynosa

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 766

Included observations: 766

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	54.23794	24.05913	2.254360	0.0245
SQM	0.002368	0.000206	11.51388	0.0000
ROOMS	0.197381	0.032448	6.082923	0.0000
BATHS	0.358929	0.032098	11.18245	0.0000
CARS	0.048469	0.066476	0.729124	0.4662
LAT	0.988478	0.546916	1.807369	0.0711
LONG	0.721115	0.249918	2.885406	0.0040
R-squared	0.643645	Mean dependent var		10.50875
Adjusted R-squared	0.640828	S.D. dependent var		0.732148
S.E. of regression	0.438783	Akaike info criterion		1.199475
Sum squared resid	146.1310	Schwarz criterion		1.241888
Log likelihood	-452.3990	Hannan-Quinn criter.		1.215802
F-statistic	228.4832	Durbin-Watson stat		1.018139
Prob(F-statistic)	0.000000	Wald F-statistic		248.3854
Prob(Wald F-statistic)	0.000000			

US07 McAllen

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 468

Included observations: 468

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-23.16445	11.95284	-1.937987	0.0532
SQM	0.003098	0.000257	12.07446	0.0000
ROOMS	-0.021196	0.029424	-0.720368	0.4717
BATHS	0.104879	0.023532	4.456946	0.0000
LAT	0.923030	0.205139	4.499522	0.0000
LONG	-0.107384	0.123751	-0.867743	0.3860
R-squared	0.679536	Mean dependent var		12.45793
Adjusted R-squared	0.676068	S.D. dependent var		0.464353
S.E. of regression	0.264287	Akaike info criterion		0.189173
Sum squared resid	32.26954	Schwarz criterion		0.242359
Log likelihood	-38.26657	Hannan-Quinn criter.		0.210102
F-statistic	195.9318	Durbin-Watson stat		1.182597
Prob(F-statistic)	0.000000	Wald F-statistic		152.2038
Prob(Wald F-statistic)	0.000000			

MX08 Matamoros

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 221

Included observations: 221

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-69.25125	71.64537	-0.966584	0.3348
SQM	0.000705	0.000221	3.184345	0.0017
ROOMS	0.284839	0.061618	4.622657	0.0000
BATHS	0.267620	0.053326	5.018593	0.0000
CARS	0.334696	0.090309	3.706105	0.0003
LAT	2.644968	1.081314	2.446068	0.0152
LONG	-0.101724	0.512126	-0.198630	0.8427
R-squared	0.674377	Mean dependent var		10.68354
Adjusted R-squared	0.665247	S.D. dependent var		0.729978
S.E. of regression	0.422350	Akaike info criterion		1.145195
Sum squared resid	38.17315	Schwarz criterion		1.252829
Log likelihood	-119.5441	Hannan-Quinn criter.		1.188656
F-statistic	73.86688	Durbin-Watson stat		1.365382
Prob(F-statistic)	0.000000	Wald F-statistic		55.40532
Prob(Wald F-statistic)	0.000000			

US08 Brownsville

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 253

Included observations: 253

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	167.6801	17.59820	9.528255	0.0000
SQM	0.003164	0.000772	4.098636	0.0001
ROOMS	-0.029330	0.069674	-0.420963	0.6741
BATHS	0.310683	0.062520	4.969304	0.0000
LAT	1.426630	0.396035	3.602279	0.0004
LONG	1.985692	0.166204	11.94731	0.0000
R-squared	0.725795	Mean dependent var		13.04421
Adjusted R-squared	0.720244	S.D. dependent var		0.893667
S.E. of regression	0.472678	Akaike info criterion		1.362624
Sum squared resid	55.18579	Schwarz criterion		1.446420
Log likelihood	-166.3719	Hannan-Quinn criter.		1.396338
F-statistic	130.7571	Durbin-Watson stat		1.259608
Prob(F-statistic)	0.000000	Wald F-statistic		95.43894
Prob(Wald F-statistic)	0.000000			

### US09 Whatcom

Dependent Variable: LOG(USD\_PRICE)

Method: Least Squares

Sample: 1 42

Included observations: 42

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	41.05996	39.14740	1.048856	0.3012
SQM	0.004067	0.001269	3.204601	0.0028
ROOMS	0.040030	0.123607	0.323847	0.7479
BATHS	-0.034730	0.175888	-0.197453	0.8446
LAT	-1.080033	0.733844	-1.471746	0.1498
LONG	-0.197597	0.317328	-0.622690	0.5374
R-squared	0.417228	Mean dependent var		13.29658
Adjusted R-squared	0.336288	S.D. dependent var		0.465941
S.E. of regression	0.379596	Akaike info criterion		1.032143
Sum squared resid	5.187340	Schwarz criterion		1.280381
Log likelihood	-15.67500	Hannan-Quinn criter.		1.123132
F-statistic	5.154755	Durbin-Watson stat		2.090970
Prob(F-statistic)	0.001146	Wald F-statistic		5.861386
Prob(Wald F-statistic)	0.000464			

### CA09 Metro Vancouver

Dependent Variable: LOG(USD\_PRICE)

Method: Least Squares

Sample: 1 509

Included observations: 509

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-120.5410	17.99297	-6.699341	0.0000
SQM	0.003567	0.000303	11.77986	0.0000
ROOMS	0.096082	0.019105	5.029223	0.0000
BATHS	-0.049906	0.024789	-2.013210	0.0446
LAT	0.447309	0.222342	2.011801	0.0448
LONG	-0.906577	0.126153	-7.186323	0.0000
R-squared	0.685274	Mean dependent var		13.74458
Adjusted R-squared	0.682145	S.D. dependent var		0.647423
S.E. of regression	0.365008	Akaike info criterion		0.833923
Sum squared resid	67.01514	Schwarz criterion		0.883815
Log likelihood	-206.2335	Hannan-Quinn criter.		0.853486
F-statistic	219.0426	Durbin-Watson stat		1.230093
Prob(F-statistic)	0.000000	Wald F-statistic		204.1302
Prob(Wald F-statistic)	0.000000			

US10 Sault St. Marie

Dependent Variable: LOG(USD\_PRICE)

Method: Least Squares

Sample: 1 105

Included observations: 105

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-8.338477	17.96390	-0.464180	0.6435
SQM	0.004007	0.001013	3.957180	0.0001
ROOMS	-0.092833	0.050089	-1.853380	0.0668
BATHS	0.159175	0.070025	2.273116	0.0252
LAT	0.242251	0.280278	0.864321	0.3895
LONG	-0.098195	0.059196	-1.658827	0.1003
R-squared	0.368573	Mean dependent var		11.98359
Adjusted R-squared	0.336683	S.D. dependent var		0.410129
S.E. of regression	0.334027	Akaike info criterion		0.700255
Sum squared resid	11.04582	Schwarz criterion		0.851910
Log likelihood	-30.76337	Hannan-Quinn criter.		0.761708
F-statistic	11.55756	Durbin-Watson stat		1.519455
Prob(F-statistic)	0.000000	Wald F-statistic		20.39555
Prob(Wald F-statistic)	0.000000			

CA10 Sault Sainte Marie

Dependent Variable: LOG(USD\_PRICE)

Method: Least Squares

Sample: 1 27

Included observations: 27

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-281.8310	249.4213	-1.129939	0.2712
SQM	0.002865	0.001703	1.681959	0.1074
ROOMS	-0.161583	0.107424	-1.504154	0.1474
BATHS	0.324899	0.147546	2.202017	0.0390
LAT	14.77042	7.041209	2.097711	0.0482
LONG	4.667116	2.826862	1.650988	0.1136
R-squared	0.615928	Mean dependent var		12.18356
Adjusted R-squared	0.524482	S.D. dependent var		0.502152
S.E. of regression	0.346273	Akaike info criterion		0.909953
Sum squared resid	2.518010	Schwarz criterion		1.197917
Log likelihood	-6.284371	Hannan-Quinn criter.		0.995580
F-statistic	6.735442	Durbin-Watson stat		1.226896
Prob(F-statistic)	0.000686	Wald F-statistic		9.702775
Prob(Wald F-statistic)	0.000065			

## US11 Detroit

Dependent Variable: LOG(USD\_PRICE)

Method: Least Squares

Sample: 1 1659

Included observations: 1659

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-26.52994	11.20467	-2.367757	0.0180
SQM	0.003180	0.000699	4.546722	0.0000
ROOMS	-0.053701	0.030657	-1.751661	0.0800
BATHS	0.539434	0.045819	11.77304	0.0000
LAT	-0.996873	0.143706	-6.936896	0.0000
LONG	-0.953964	0.105984	-9.001003	0.0000
R-squared	0.465646	Mean dependent var		11.91102
Adjusted R-squared	0.464030	S.D. dependent var		0.970794
S.E. of regression	0.710718	Akaike info criterion		2.158529
Sum squared resid	834.9639	Schwarz criterion		2.178109
Log likelihood	-1784.499	Hannan-Quinn criter.		2.165786
F-statistic	288.0915	Durbin-Watson stat		0.685246
Prob(F-statistic)	0.000000	Wald F-statistic		337.8287
Prob(Wald F-statistic)	0.000000			

## CA11 Windsor

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 50

Included observations: 50

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	33.06587	319.1151	0.103617	0.9179
SQM	0.003547	0.004762	0.744742	0.4604
ROOMS	-0.004360	0.090142	-0.048367	0.9616
BATHS	0.398574	0.475814	0.837668	0.4067
LAT	-0.152190	1.481187	-0.102748	0.9186
LONG	0.184917	3.175006	0.058242	0.9538
R-squared	0.332512	Mean dependent var		13.13292
Adjusted R-squared	0.256661	S.D. dependent var		1.293562
S.E. of regression	1.115272	Akaike info criterion		3.168241
Sum squared resid	54.72861	Schwarz criterion		3.397683
Log likelihood	-73.20602	Hannan-Quinn criter.		3.255614
F-statistic	4.383757	Durbin-Watson stat		0.897024
Prob(F-statistic)	0.002523	Wald F-statistic		5.666460
Prob(Wald F-statistic)	0.000407			

US12 Port Huron

Dependent Variable: LOG(USD\_PRICE)

Method: Least Squares

Sample: 1 43

Included observations: 43

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.474042	147.7720	0.003208	0.9975
SQM	0.002919	0.000827	3.531559	0.0011
ROOMS	0.070491	0.047462	1.485223	0.1460
BATHS	0.165043	0.086592	1.905987	0.0644
LAT	8.286829	4.045561	2.048376	0.0477
LONG	4.171207	2.519276	1.655717	0.1062
R-squared	0.767053	Mean dependent var		12.54204
Adjusted R-squared	0.735573	S.D. dependent var		0.561495
S.E. of regression	0.288735	Akaike info criterion		0.482171
Sum squared resid	3.084607	Schwarz criterion		0.727920
Log likelihood	-4.366673	Hannan-Quinn criter.		0.572795
F-statistic	24.36682	Durbin-Watson stat		1.521053
Prob(F-statistic)	0.000000	Wald F-statistic		17.15793
Prob(Wald F-statistic)	0.000000			

CA12 Point Edward

Dependent Variable: LOG(USD\_PRICE)

Method: Least Squares

Sample: 1 20

Included observations: 20

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-11.96908	27.89580	-0.429064	0.6744
SQM	0.000745	0.000734	1.015047	0.3273
ROOMS	-0.097368	0.051446	-1.892637	0.0793
BATHS	0.390957	0.095934	4.075263	0.0011
LAT	0.624978	0.296259	2.109566	0.0534
LONG	0.030868	0.241198	0.127976	0.9000
R-squared	0.719129	Mean dependent var		12.96586
Adjusted R-squared	0.618818	S.D. dependent var		0.397572
S.E. of regression	0.245461	Akaike info criterion		0.271967
Sum squared resid	0.843515	Schwarz criterion		0.570686
Log likelihood	3.280333	Hannan-Quinn criter.		0.330280
F-statistic	7.168997	Durbin-Watson stat		2.230053
Prob(F-statistic)	0.001609	Wald F-statistic		15.80260
Prob(Wald F-statistic)	0.000025			

### US13 Buffalo-Niagara Falls

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 1006

Included observations: 1006

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	48.80297	12.23210	3.989747	0.0001
SQM	0.004913	0.000732	6.707799	0.0000
ROOMS	-0.088822	0.022372	-3.970280	0.0001
BATHS	0.385988	0.033597	11.48875	0.0000
LAT	-0.135896	0.105464	-1.288556	0.1979
LONG	0.404217	0.139994	2.887381	0.0040
R-squared	0.435941	Mean dependent var		12.17826
Adjusted R-squared	0.433121	S.D. dependent var		0.618983
S.E. of regression	0.466041	Akaike info criterion		1.316859
Sum squared resid	217.1939	Schwarz criterion		1.346165
Log likelihood	-656.3799	Hannan-Quinn criter.		1.327994
F-statistic	154.5728	Durbin-Watson stat		0.888698
Prob(F-statistic)	0.000000	Wald F-statistic		157.1624
Prob(Wald F-statistic)	0.000000			

### CA13 Great Golden Horseshoe

**Dependent Variable: LOG(USD\_PRICE)**

Method: Least Squares

Sample: 1 56

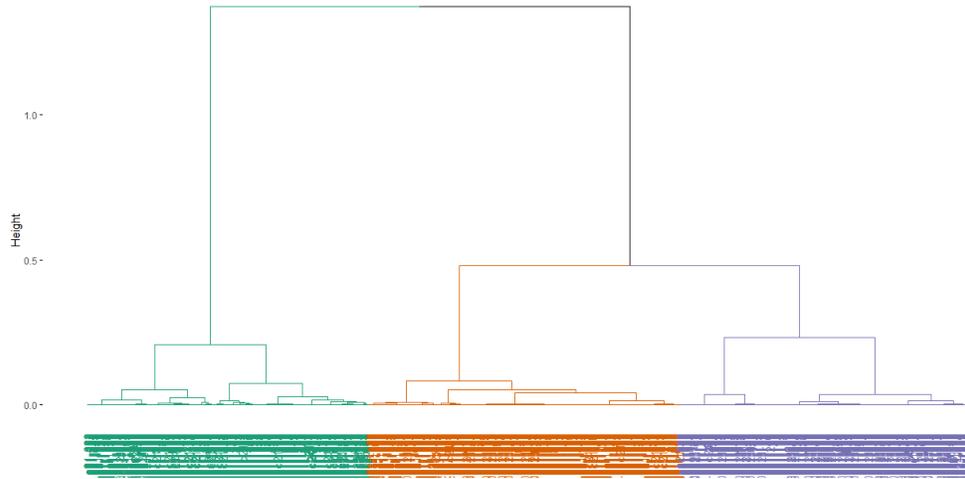
Included observations: 56

White heteroskedasticity-consistent standard errors & covariance

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-29.39952	12.37018	-2.376643	0.0213
SQM	0.002464	0.000538	4.580621	0.0000
ROOMS	-0.040878	0.069066	-0.591871	0.5566
BATHS	0.178243	0.078223	2.278665	0.0270
LAT	0.469426	0.607496	0.772724	0.4433
LONG	-0.274537	0.293174	-0.936428	0.3536
R-squared	0.768044	Mean dependent var		13.77191
Adjusted R-squared	0.744848	S.D. dependent var		0.821822
S.E. of regression	0.415123	Akaike info criterion		1.180473
Sum squared resid	8.616355	Schwarz criterion		1.397475
Log likelihood	-27.05325	Hannan-Quinn criter.		1.264604
F-statistic	33.11161	Durbin-Watson stat		1.533647
Prob(F-statistic)	0.000000	Wald F-statistic		42.55648
Prob(Wald F-statistic)	0.000000			

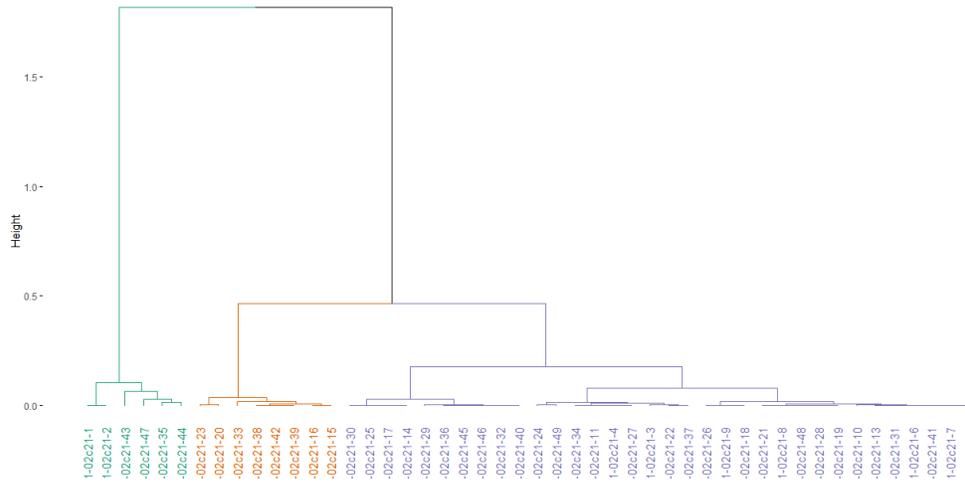
### US01 San Diego

Cluster Dendrogram



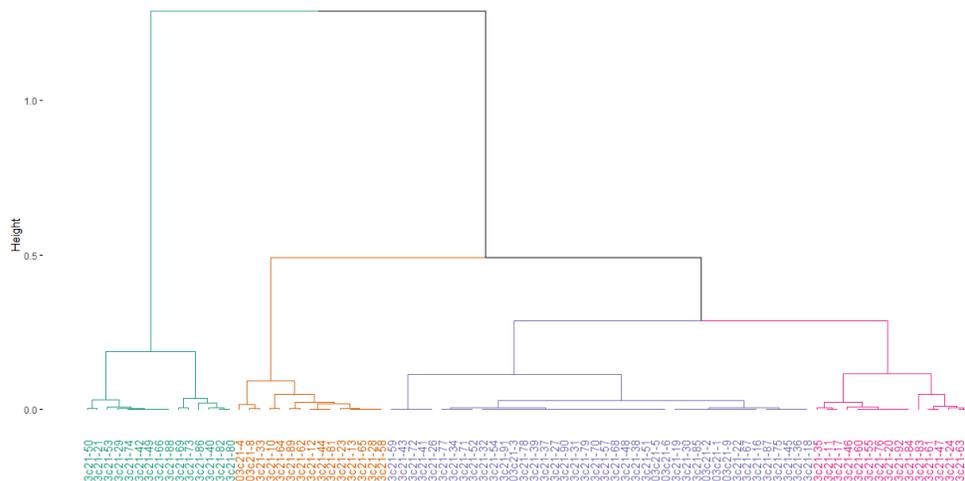
### US02 Imperial

Cluster Dendrogram

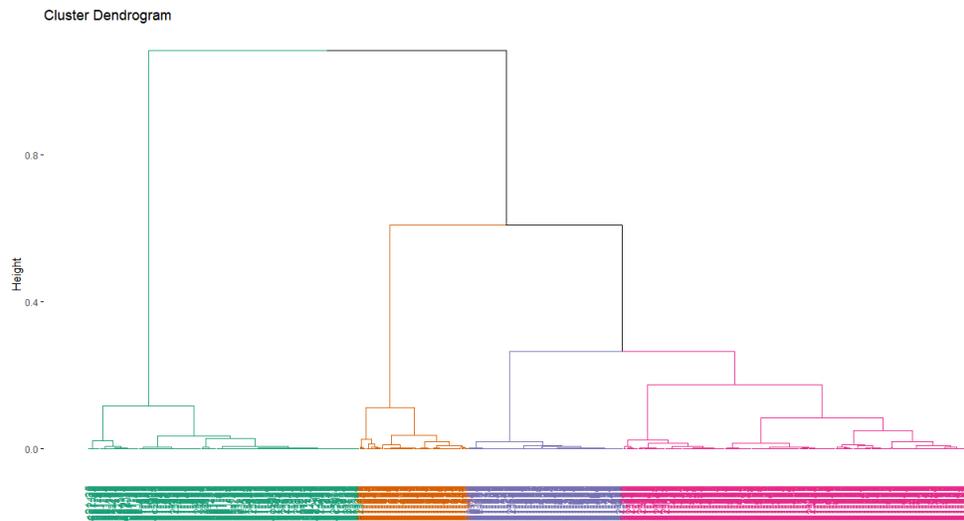


### US03 Nogales

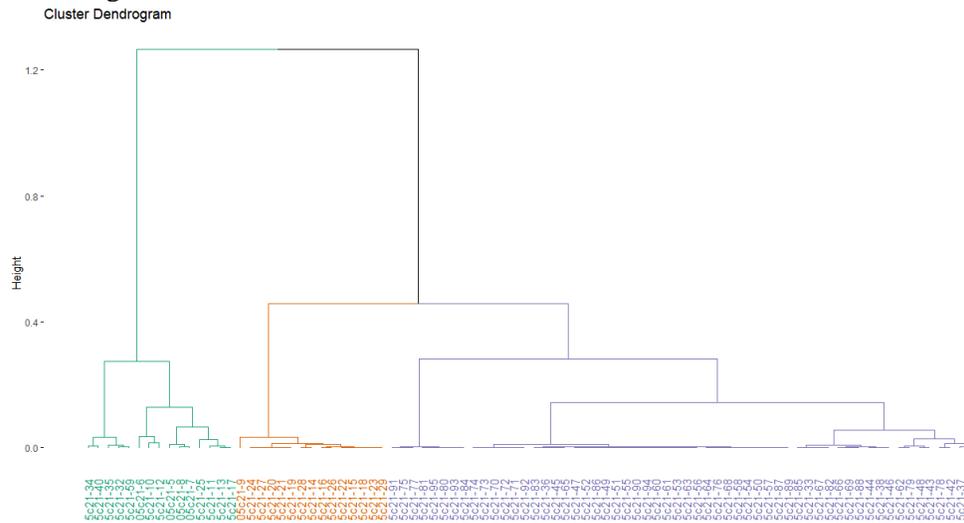
Cluster Dendrogram



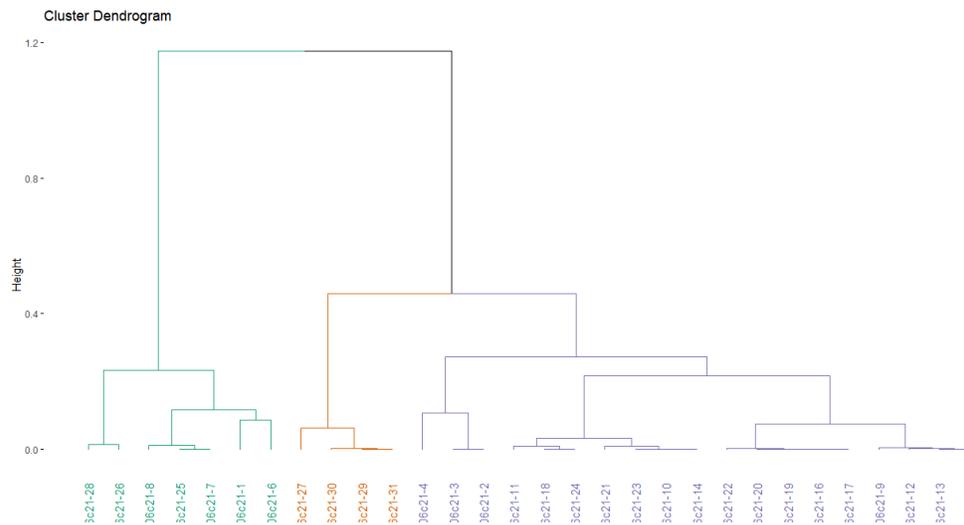
## US04 El Paso



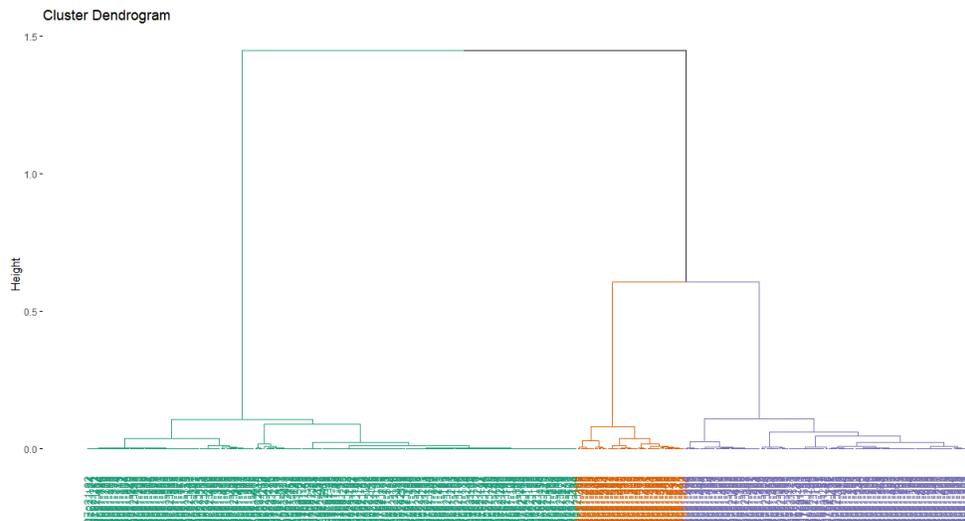
## US05 Eagle Pass



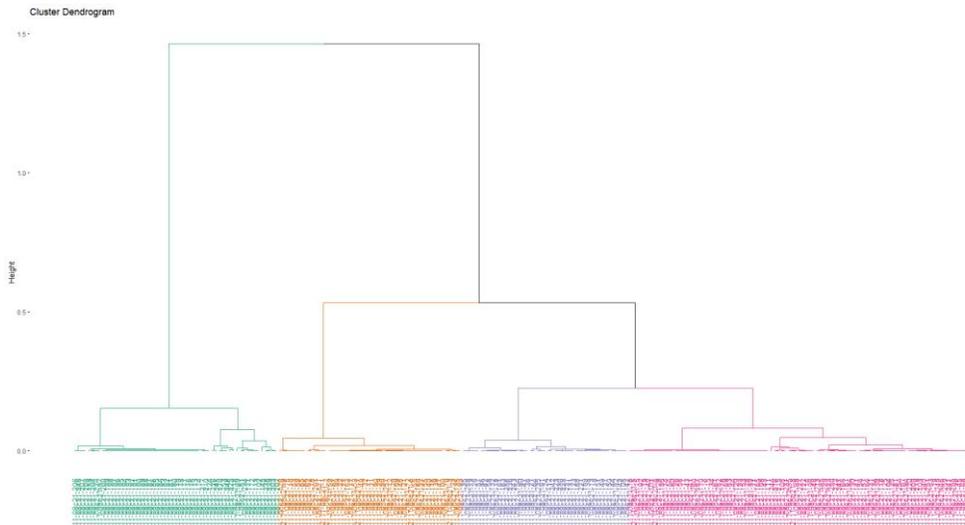
## US06 Laredo



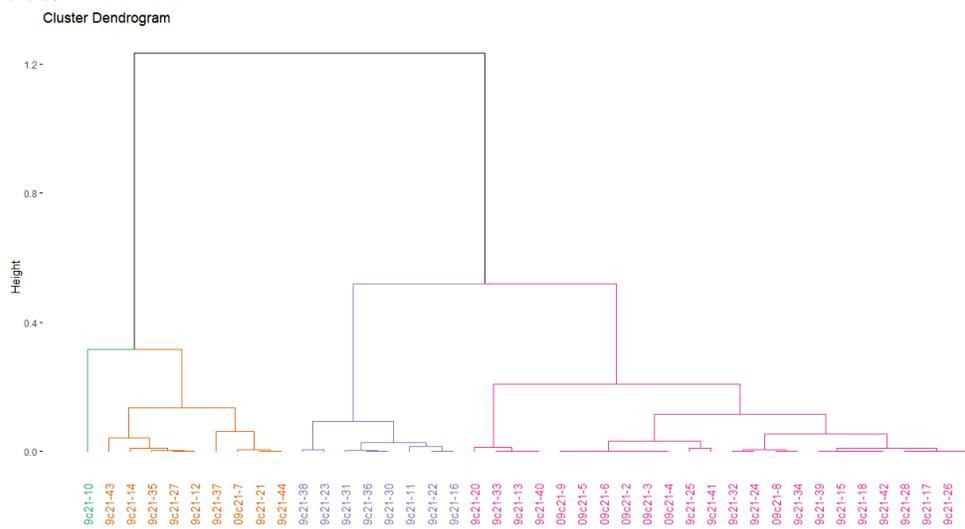
### US07 McAllen



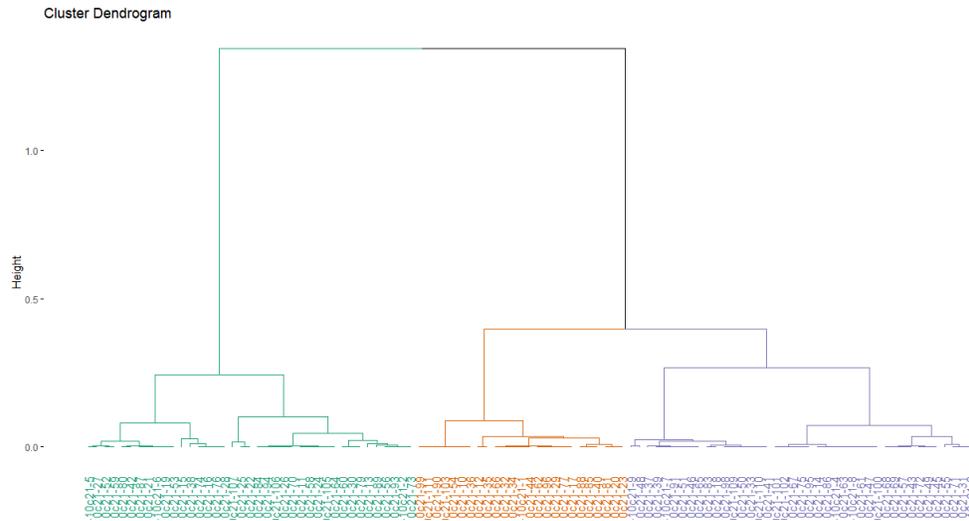
### US08 Brownsville



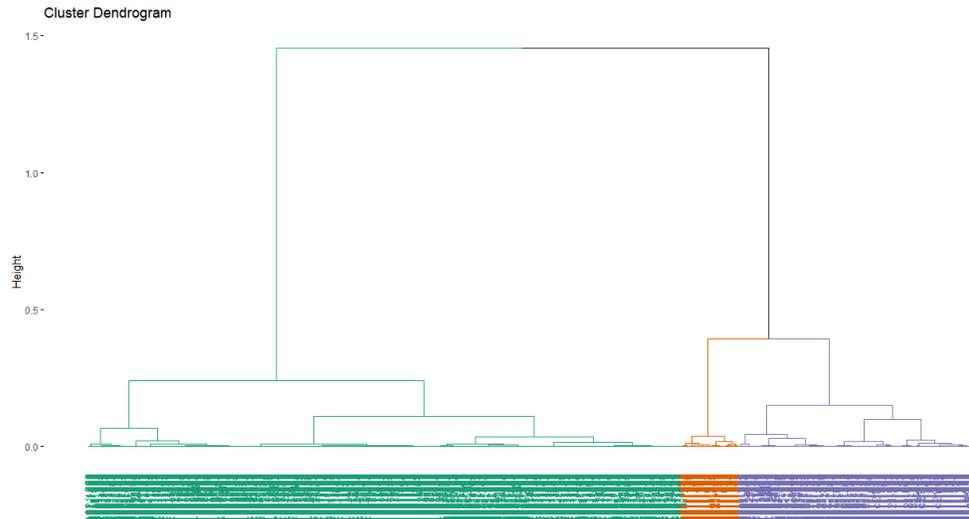
### US09 Whatcom



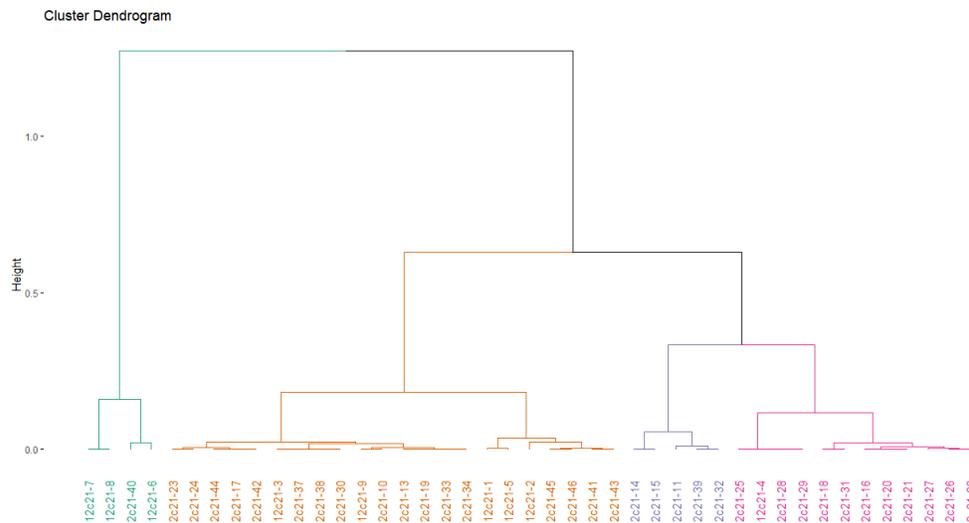
## US10 Sault St. Marie



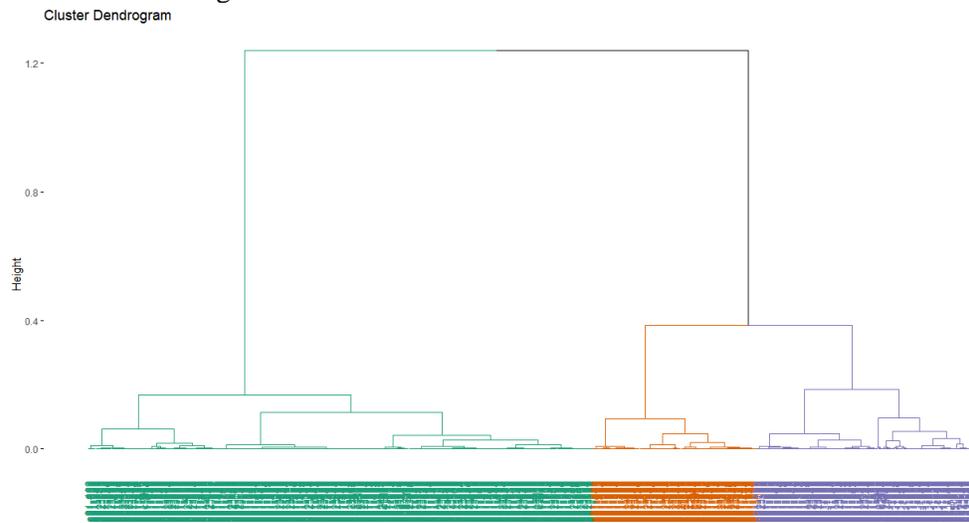
## US11 Detroit



## US12 Point Edward



# US13 Buffalo-Niagara



La autora es Maestra Economía Aplicada por El Colegio de la Frontera Norte, Licenciada en Relaciones Internacionales por la Universidad Autónoma de Baja California, UABC, Campus Tijuana. Egresada del Doctorado en Ciencias Económicas de la Facultad de Economía y Relaciones Internacionales, Universidad Autónoma de Baja California, UABC, Campus Tijuana. Ha realizado estancias de investigación en la Luskin School of Public Affairs de la UCLA y ha cursado estudios en la Université Paris Ouest Nanterre La Défense en Francia, la Universidad de Granada en España, la Universidad Nacional Tres de Febrero en Buenos Aires y la Universidad Nacional de Córdoba en Argentina.

Correo electrónico: [atenea.cb@gmail.com](mailto:atenea.cb@gmail.com)

© *Todos los derechos reservados.*

Forma de citar:

De La Cruz Brito, Atenea (2022). “Mercados de vivienda transfronterizos entre México, Estados Unidos y Canadá: submercados e información asimétrica”. Tesis de Doctorado en Ciencias Económicas. Universidad Autónoma de Baja California, UABC, México. 128 pp.