

# Una propuesta de medición de daño del delito y su impacto sobre los precios de Airbnb en la Ciudad de Buenos Aires

**Alejandro Norton** (Universidad Nacional de Cuyo)

**Marcos Herrera-Gomez** (Universidad Nacional de Salta/CONICET)

## DOCUMENTO DE TRABAJO 2020-21

*Publicado como Documento de trabajo Nro.24, Instituto de Estudios Laborales  
y del Desarrollo Económico (IELDE) de la Universidad Nacional de Salta (UNSa).*

Agosto de 2020

Los documentos de trabajo de la RedNIE se difunden con el propósito de generar comentarios y debate, no habiendo estado sujetos a revisión de pares. Las opiniones expresadas en este trabajo son de los autores y no necesariamente representan las opiniones de la RedNIE o su Comisión Directiva.

The RedNIE working papers are disseminated for the purpose of generating comments and debate, and have not been subjected to peer review. The opinions expressed in this paper are exclusively those of the authors and do not necessarily represent the opinions of the RedNIE or its Board of Directors.

Citar como:

Norton, Alejandro, y Marcos Herrera-Gomez (2020). Una propuesta de medición de daño del delito y su impacto sobre los precios de Airbnb en la Ciudad de Buenos Aires. *Documento de trabajo Nro.24, Instituto de Estudios Laborales y del Desarrollo Económico (IELDE) de la Universidad Nacional de Salta (UNSa)*.

# Una propuesta de medición de daño del delito y su impacto sobre los precios de Airbnb en la Ciudad de Buenos Aires

Alejandro Norton<sup>\*</sup>, Marcos Herrera-Gomez<sup>\*\*</sup>

## Resumen

Usando información de hospedajes de Airbnb en la Ciudad de Buenos Aires, este trabajo analiza la relación entre las actividades delictivas y los precios de alquiler. El trabajo plantea un modelo hedónico que permite estimar el impacto de daño del delito controlando por otros factores (características del alojamiento y variables contextuales). Los principales aportes del trabajo pueden ser resumidos en tres partes. En primer lugar, se utiliza información geográfica para construir un indicador de daño del delito que pondera los tipos de delitos con respecto al daño producido por los mismos. En segundo lugar, el estudio reconoce el problema de la endogeneidad del delito y provee evidencia robusta sobre el impacto del delito en precios de Airbnb. En tercer lugar, el modelo principal incluye una corrección por dependencia espacial endógena en precios y por heteroscedasticidad espacial en el término de error. Nuestros resultados indican que los precios de los alojamientos son afectados negativamente por la actividad delictiva en el vecindario.

**Clasificación JEL:** Z30, C21, R32, K14.

**Palabras Clave:** Precios Airbnb, Índice de Daño del Delito, Modelo Hedónico Espacial.

---

<sup>\*</sup>Facultad de Ciencias Económicas, UNCUIYO, Doctorado en Economía; Centro Universitario (M5502JMA), Ciudad de Mendoza, Argentina; email: alejandronorton@gmail.com.

<sup>\*\*</sup>Autor Correspondiente. CONICET - IELDE, Universidad Nacional de Salta; Av. Bolivia 5150 (A4408FVY), Salta, Argentina; email: mherreragomez@conicet.gov.ar.

# 1. Introducción

En la última década, el sector de la economía colaborativa (*sharing economy*) ha generado el nacimiento de nuevas formas de turismo como es el caso del boom de los alojamientos compartidos (Zervas et al. 2017). La economía colaborativa es parte de un modelo de negocios de *e-commerce peer-to-peer* (P2P) donde las plataformas digitales permiten monetizar recursos inmovilizados acercándolos a potenciales consumidores en línea (Belk 2014, Kong et al. 2020). Una de las plataformas más importantes es Airbnb, que está disponible en 62 idiomas y recibe a más de 750 millones de huéspedes, con presencia en más de 220 países y regiones alrededor del mundo (<https://news.airbnb.com/fast-facts/>, Julio 2020). El llamativo crecimiento de esta plataforma ha llevado investigar cuales son los principales factores que explican su evolución y sus precios, incluyendo temas de seguridad del turista (Adamiak et al. 2019, Priporas et al. 2017, Yang & Mao 2020, entre otros).

Nuestro interés se centra en la relación entre el delito y los precios de Airbnb. Este tipo de relación, entre actividades delictivas y el sector turístico, es bien conocido por el sector hotelero tradicional, en particular, por las diferentes medidas de seguridad para evitar estos incidentes (Chesney-Lind & Lind 1986, Allen 1999, Jones & Groenenboom 2002, Ho et al. 2017). Los huéspedes de hoteles son un claro objetivo de hechos delictivos como hurto, robo, fraudes de tarjeta de crédito, incluso abusos de tipo sexual, por mencionar los más habituales. De igual manera, los huéspedes de Airbnb son víctimas de delitos durante su estadía, y la seguridad del lugar o vecindario del alojamiento ofrecido es un tema recurrente (véase <https://www.airnbhell.com/>). A pesar de estos hechos, la investigación sobre la relación entre delitos, seguridad y hospedajes de Airbnb es realmente escasa. Recientemente, Xu et al. (2019) analiza la relación entre la densidad de hospedajes de Airbnb y diferentes tipos de delitos. Además, Tang et al. (2019) incluye la cantidad de delitos como uno de los factores adicionales que influyen en el precio, pero sin ser de principal interés. Es por ello que nuestra investigación pretende cubrir este vacío y entender los efectos del delito sobre los precios de hospedaje de Airbnb, obteniendo una medición de su impacto tal que pueda revelar información valiosa para el sector.

Nuestra investigación propone explorar el impacto del delito en los precios de alojamientos de Airbnb después de controlar por características propias y por factores contextuales. Con la idea de considerar un único indicador de los diferentes delitos, se propone un Índice de Daño del Delito (Crime Harm Index, en inglés, CHI en adelante) derivado de la aplicación de una métrica que pondera a cada delito según el daño causado por el mismo (Sherman et al. 2016). El CHI permite ponderar los delitos y agregarlos para tener una medida general del daño a la sociedad. De esta forma, es un indicador que se contrapone a la forma tradicional de considerar únicamente el volumen del delito, donde todos los delitos tienen el mismo peso.

En el sector hotelero tradicional, estudios previos han demostrado que la actividad delictiva en destinos turísticos tiene un efecto significativo en el sector (Zhao et al. 2004). En el caso de Airbnb, el menor costo y flexibilidad operativa de esta plataforma ha dado como resultado un crecimiento de su participación de mercado; sin embargo, la actividad delictiva puede destruir rápidamente la reputación debido a la falta de

seguridad de algunas ubicaciones. Pero no es tan simple identificar el impacto de la actividad delictiva debido a la posible endogeneidad entre ella y los precios. Este problema es reconocido por estudios de economía urbana, por ejemplo, Gibbons (2004) menciona que las tasas de delito registradas serán endógenas a los precios de las viviendas; además, Ceccato & Wilhelmsson (2011) consideran una corrección por endogeneidad en modelos hedónicos. Para evitar este problema, nuestro trabajo utiliza información espacial de un año previo para identificar el impacto del índice del delito.

Otra característica de importancia es la ubicación geográfica de los alojamientos. La disponibilidad de la ubicación exacta del alojamiento ha permitido explorar que la distribución espacial de los alojamientos y el delito se distribuyen de forma no aleatoria en el espacio (Eck et al. 2005, Weisburd 2015, Adamiak et al. 2019), implicando que las actividades delictivas pueden afectar a una zona particular y expandirse a mayores regiones o afectar las cercanías a esa zona. La omisión de este tipo de características espaciales podría resultar en una conclusión sesgada.

La contribución de esta investigación tiene tres pilares. Primero, utilizamos información geográfica detallada para construir un índice de daño del delito (CHI) que pondera los delitos de acuerdo al daño causado. Segundo, este estudio introduce variables instrumentales para sortear el problema de la endogeneidad del delito y provee evidencia robusta sobre el impacto del delito en los precios de Airbnb, información no disponible en la literatura internacional. Tercero, se incorporan efectos de vecindario geográficos en el modelo principal generando una corrección por dependencia espacial de los precios y en el término de error.

Nuestro enfoque empírico considera como punto de partida al modelo hedónico tradicional, considerando como variables explicativas a las características propias de los alojamientos, variables contextuales y al delito medido por el CHI. La estrategia de estimación del modelo principal utiliza mínimos cuadrados bietápicos espaciales generalizados (GS2SLS, *Generalized Spatial Two Stage Least Squares*), incluyendo una corrección por variables instrumentales que reconoce la endogeneidad del delito y la dependencia espacial de los precios de los alojamientos. Se incluyen chequeos de robustez de los resultados mediante contrastes de validez de las variables instrumentales, sensibilidad a la zona de influencia de los delitos y ante cambios en la matriz de pesos espaciales elegida.

Nuestros resultados indican que los precios de alojamientos son afectados significativamente por el delito de zonas vecinas. El indicador de daño del delito (CHI) muestra un impacto negativo utilizando mínimos cuadrados ordinarios (MCO) y variables instrumentales (VI). La elasticidad del CHI es -0.042 bajo MCO y -0.093 bajo VI. La estimación por MCO tiene dos problemas: detectamos endogeneidad del índice del delito y los residuos están correlacionados espacialmente. En el caso de la estimación por VI, se omiten los efectos espaciales invalidando los resultados inferenciales. Para resolver estos dos problemas se introduce un modelo de rezago espacial SLM VI (Spatial Lag Model) con corrección por endogeneidad del ACHI, e incluye una corrección en el término de error por heteroscedasticidad espacial. El impacto total de CHI en los precios utilizando el modelo SLM VI es de -0.074, que puede ser descompuesto en un efecto propio (efecto directo) de -0.064 y un derrame espacial (efecto indirecto) de -0.010. El efecto directo, neto de la dependencia espacial,

presenta un valor que se ubica entre las estimaciones por MCO y VI, y el efecto total es cercano al coeficiente de VI. Este resultado muestra el sesgo del modelo de VI cuando se omite la dependencia espacial. Además, el SLM VI logra capturar el efecto indirecto, conocido como *spillover* espacial, que captura el efecto del contagio en los precios del vecindario.

El artículo se organiza de la siguiente manera. En la segunda sección se presentan los materiales y los métodos estadísticos utilizados. Esta sección explica en detalle la medición del índice de daño del delito, las fuentes de Airbnb y las variables contextuales. También se explican los modelos hedónicos y algunas herramientas de econometría espacial. La tercera sección presenta el análisis estadístico descriptivo y empírico del modelo hedónico utilizando diferentes especificaciones para los datos de Airbnb. Se presta especial atención a describir las fuentes de datos y la estrategia para identificar los *outliers* espaciales. Esta sección reporta los resultados de los principales modelos hedónicos. La cuarta sección incluye la discusión de los resultados y detalles adicionales. Finalmente, la quinta sección presenta las conclusiones y comentarios adicionales.

## 2. Materiales y métodos

Nuestra área de estudio es la Ciudad Autónoma de Buenos Aires, el distrito federal y capital de Argentina. La extensión de la ciudad es de 294 km<sup>2</sup> con una densidad media poblacional de 15038 habitantes/km<sup>2</sup> en 2018 (Dirección General de Estadísticas y Censos 2018). La división oficial de Buenos Aires es de 48 barrios; sin embargo, administrativamente se utiliza una división de 15 comunas.

### 2.1. Indicador de daño del delito para la Ciudad de Buenos Aires

Usualmente el estudio del delito tiende a estar basado en el conteo de un tipo específico de cada delito o la suma del total, sin considerar la magnitud del daño que causa el delito. Sherman (2007) menciona que una medida ponderada es la mejor forma de reflejar la naturaleza del delito. Sherman et al. (2016) sugiere un enfoque objetivo donde el delito debe ser ponderado por la propia severidad y provee un resumen y evaluación de los diferentes enfoques para medir el daño del delito a la fecha. Estos incluyen los registros de sentencias efectivas, relevamientos de víctimas o del público, estimaciones del costo del delito, sentencias mínimas y tarifas. Sherman et al. (2016) realiza el CHI de Cambridge y utiliza el número de días de prisión de acuerdo a la legislación penal vigente en el Reino Unido. Por otro lado, la pena máxima es descartada porque es raramente utilizada. La idea del CHI de Cambridge fue adaptada por otros países como Nueva Zelanda (Curtis-Ham & Walton 2018), Dinamarca (Andersen & Mueller-Johnson 2018), Estados Unidos (Fenimore 2019) y Suecia (Kärrholm et al. 2020).

Para crear el CHI para la Ciudad de Buenos Aires, usaremos información de los delitos georreferenciados relevados por la Policía de la Ciudad de Buenos Aires y presentados en el Mapa del Delito (<https://mapa.seguridadciudad.gob.ar/>), una herramienta en línea con información confiable de estadísticas de la ciudad a partir de 2016. Los datos tienen como origen los reportes de las diferentes comisarías de la

ciudad, los cuales son procesados y consolidados por el mismo organismo de seguridad. La información es actualizada en forma semestral y es de libre acceso.

Para nuestro propósito se utilizaron todos los delitos reportados entre el 1 de Enero de 2017 y el 31 de Diciembre de 2017, los datos están divididos en cinco categorías de delitos con latitud y longitud, hora del hecho y barrio (basado en la fecha de ocurrencia, no reporte). Las categorías son:

- *Hurto*: Hecho registrado como el apoderamiento de una cosa sin violencia sobre la persona o el objeto, generalmente calificado los términos del Artículo 162 y subsiguientes del Código Penal, sin considerar la imputabilidad o culpabilidad del autor.
- *Hurto de Vehículo*: Todo hecho registrado como Hurto, donde el objeto sustraído es un vehículo completo con motor de cuatro o más ruedas.
- *Robo*: Hecho registrado como el apoderamiento de una cosa por violencia sobre la persona o el objeto, generalmente calificado los términos del Artículo 164 y subsiguientes del Código Penal, sin considerar la imputabilidad o culpabilidad del autor.
- *Robo de Vehículo*: Todo hecho registrado como Robo, donde el objeto sustraído es un vehículo completo con motor de cuatro o más ruedas, sin perjuicio de la presencia de su titular o conductor.
- *Homicidio*: Muerte violenta registrada como homicidio doloso en los términos del Artículo 79 del Código Penal, sus atenuantes y agravantes, sin considerar la imputabilidad o culpabilidad del autor. Incluye los hechos registrados donde la persona fallecida fue reportada cometiendo otro ilícito.

El total de delitos en 2017 fue de 108785, distribuidos en 36098 hurtos, 5952 hurtos de vehículos, 63124 robos, 3477 robos de vehículos, y 134 homicidios.

Adicionalmente, exploramos el número de días en prisión de acuerdo a la ley penal como una alternativa costo-efectiva. Las penalidades para los delitos analizados son tomadas del Código Penal Argentino (Ley 11,179, Art. 79/162/163/164/167, <http://servicios.infoleg.gob.ar/infolegInternet/anexos/15000-19999/16546/texact.htm>) y son las guías actuales para las sentencias. El Cuadro 1 muestra los días de prisión para Argentina en comparación con Reino Unido y Dinamarca. El robo de vehículos no está especificado, por lo que se lo encuadra como un robo agravado, considerando la violencia que conlleva.

El hurto y el robo tienen la misma sentencia mínima; sin embargo, el impacto para la víctima es sustancialmente diferente. En consecuencia, descartamos a los días mínimos como ponderación del daño. De la misma forma, la literatura menciona que la pena máxima no es una buena medida ya que su aplicación prácticamente no es de cumplimiento efectivo. Proponemos entonces la utilización del tiempo promedio de prisión como la mejor ponderación del daño asociado a cada tipo de delito, sin considerar el historial delictivo ni las diferentes circunstancias de cada caso. Esto nos va a permitir la comparación año a año sin considerar quién está cometiendo el delito y cuál es su historial.

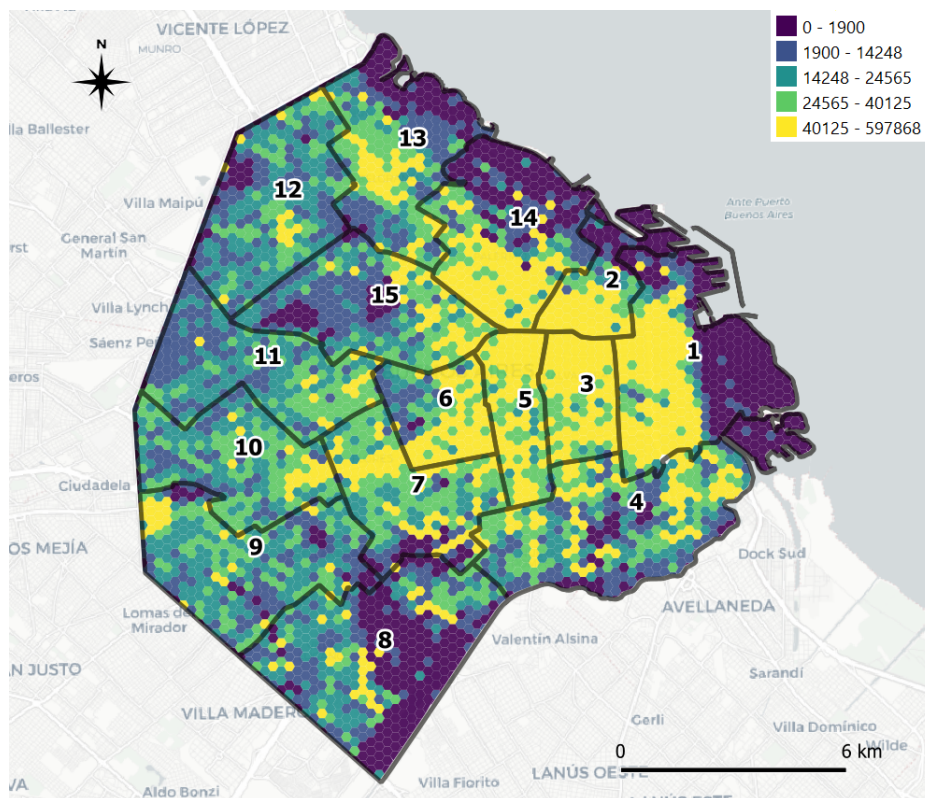
Cuadro 1: Resumen de días de prisión para Argentina, Reino Unido y Dinamarca

Tipo de Delito	Argentina				Reino Unido	Dinamarca
	Min.	Max.	Promedio	Cod. Penal	Min.	Min.
<i>Hurto</i>	30	730	380.0	Art. 162	2	10
<i>Hurto de vehículo</i>	365	2190	1277.5	Art. 163	20	14
<i>Robo</i>	30	2190	1110.0	Art. 164	365	360
<i>Robo de vehículo</i>	1095	3650	2372.5	Art. 167	—	—
<i>Homicidio</i>	2920	9125	6022.5	Art. 79	5475	4320

Nota: Art.: Artículo. Días de prisión según Ley 11179 del código penal argentino. Sherman et al. (2016, Table 2) para Reino Unido y Andersen & Mueller-Johnson (2018, Figure 2) para Dinamarca.

Aplicando el tiempo promedio de prisión como ponderador de cada tipo de delito, se construye el Índice de Daño del Delito para el año 2017. La Figura 1 presenta la distribución espacial del CHI a través de un gráfico de cuantiles. Para generar este mapa se dividieron los 204 km<sup>2</sup> del área total de la Ciudad de Buenos Aires en hexágonos de 0.054 km<sup>2</sup>, obteniendo más de 3800 hexágonos. Observamos una alta concentración de valores elevados del índice en las comunas 3 y 5, principalmente; mientras que otras comunas reflejan una mayor heterogeneidad combinando valores extremos como las comunas 1, 2, 13 y 14. Parte de la información de este índice será utilizado para medir el impacto del delito en los precios de Airbnb.

Figura 1: Distribución Espacial del CHI en 2017.





## 2.2. Descripción de base de Airbnb y variables contextuales

Las base de datos de Airbnb proviene de la información publicada por Vazquez Brust (2020),<sup>1</sup> base relevada en Febrero 2018 para la Ciudad de Buenos Aires. Vazquez Brust, utilizando un código de *web-scraping* desarrollado por Slee (2018),<sup>2</sup> provee la base de datos que incluye todos los alojamientos listados por Airbnb en la ciudad al momento del relevamiento.

La base contiene 10116 alojamientos con las siguientes variables:

- Precio (Price): tarifa en pesos argentinos por noche
- Tipo de habitación (room\_type): alojamiento entero, habitación, o habitación compartida
- Huéspedes (accommodates): número de personas de capacidad
- Habitaciones (rooms): número de habitaciones
- Baños (bathrooms): número de baños
- Evaluaciones (reviews): número de evaluaciones recibidas

Adicionalmente, se utilizó información georreferenciada para construir variables contextuales y de control. Las variables contextuales como parques, atracciones turísticas fueron obtenidas desde el sitio web de la Ciudad de Buenos Aires<sup>3</sup>. Los conjuntos de datos consultados fueron los siguientes:

- Gastronomía (<https://data.buenosaires.gob.ar/dataset/oferta-establecimientos-gastronomicos>): Listado de restaurantes y establecimientos gastronómicos en la ciudad.
- Atracciones Turísticas (<https://data.buenosaires.gob.ar/dataset/espacios-culturales>): Listado de espacios culturales públicos, privados e independientes localizados en la Ciudad de Buenos Aires. Incluye bares, bibliotecas, calesitas, centros culturales, disquerías, espacios de exhibición, espacios de formación, espacios escénicos, espacios feriales, librerías, monumentos y lugares históricos y salas de cine. Se trabajó exclusivamente con monumentos y lugares históricos.
- Parques (<https://data.buenosaires.gob.ar/dataset/espacios-verdes>): Límites y ubicación geográfica de los espacios verdes de la Ciudad (jardín, parque, patio recreativo, plaza, plazoleta, cantero y polideportivo).
- Estaciones de Trenes (<https://data.buenosaires.gob.ar/dataset/estaciones-ferrocarril>): Ubicación geográfica de líneas de trenes y estaciones en la Ciudad.
- Distrito Central Financiero-Comercial: Ubicación geográfica del Obelisco, principal monumento de la Ciudad Buenos Aires.

---

<sup>1</sup> Acceso Junio, 2020: [https://data.world/vazquez-brust/alojamientos-listados-por-airbnb-en-buenos-aires-julio-2017/workspace/file?filename=Airbnb\\_listings\\_Buenos\\_Aires\\_Jan\\_2018.csv](https://data.world/vazquez-brust/alojamientos-listados-por-airbnb-en-buenos-aires-julio-2017/workspace/file?filename=Airbnb_listings_Buenos_Aires_Jan_2018.csv)

<sup>2</sup> <https://github.com/tomslee/airbnb-data-collection>

<sup>3</sup> Acceso Junio, 2020: <https://data.buenosaires.gob.ar/dataset>

- Radios Censales/Población por radio (<https://data.buenosaires.gob.ar/dataset/informacion-censal-por-radio>): Información censal de la Ciudad desagregada por radios para el año 2010.

### 2.3. Modelo Econométrico Espacial

Nuestro principal objetivo es medir el impacto del delito sobre los precios de Airbnb. El punto de partida es un enfoque tradicional de análisis de precios de las propiedades, el modelo hedónico de precios (Rosen 1974). Este enfoque se fundamenta en el principio de que los bienes son heterogéneos y difieren en sus atributos, los que pueden ser revelados implícitamente en los precios observados. En términos formales, consideramos el siguiente modelo hedónico:

$$\ln(\mathbf{price}) = \mathbf{X}\boldsymbol{\beta} + \gamma \ln(\mathbf{CHI}) + \mathbf{u}, \quad (1)$$

donde  $\ln(\mathbf{price})$  es el vector de  $n$  observaciones en escala logarítmica de los precios de alojamientos;  $\mathbf{X}$  es una matriz  $n \times k$  de las características y atributos del vecindario,  $\boldsymbol{\beta}' = [\beta_1, \beta_2, \dots, \beta_k]$  es el vector asociado de los coeficientes de la regresión (el precio marginal implícito para cada atributo);  $\ln(\mathbf{CHI})$  es el vector que captura el logaritmo del CHI y  $\gamma$  el impacto marginal asociado; y  $\mathbf{u}$  es un vector  $n \times 1$  del término de error aleatorio.

Reconocemos dos problemas potenciales con la ecuación (1): (i) la omisión de la dependencia espacial y (ii) el problema de la endogeneidad del regresor.

Para poder aplicar los contrastes espaciales es necesario definir la matriz de pesos o ponderaciones espaciales  $\mathbf{W}$ . Esta estructura espacial es una matriz positiva de orden  $n \times n$  (el número de observaciones), usualmente pre-especificada por el investigador, que describe la disposición espacial de las observaciones (Anselin 1988). Los elementos de  $\mathbf{W}$ ,  $w_{ij}$ , son valores distintos de cero cuando suponemos que el alojamiento  $i$  y el alojamiento  $j$  son vecinos, en caso contrario será cero. Por convención, la diagonal de elementos,  $w_{ii}$ , es igual a cero, esto es, se excluye ser vecino de uno mismo. Hay múltiples alternativas para crear la matriz de pesos espaciales  $\mathbf{W}$ . En este trabajo, generamos una matriz de pesos espaciales siguiendo los siguientes pasos:

- Primero, construimos polígonos Voronoi para cada una de las ubicaciones de Airbnb, lo que permite representar a la muestra de puntos en forma de polígonos irregulares. En teoría económica, estos polígonos han sido usados como una forma fácil y sencilla para aproximar áreas de mercado (Anselin & Le Gallo 2006, Conway et al. 2010).
- En segundo lugar, usamos el criterio de contigüidad (fronteras comunes) para definir a los vecinos de cada polígono Voronoi. La matriz de pesos resultante es una matriz binaria, extremadamente dispersa (0.06 % de valores distintos de cero), que contiene como media 6 vecinos para cada ubicación (teniendo como mínimo 2 vecinos y como máximo 21).

- Finalmente, ponderamos a cada vecino utilizando una función de distancia entre los puntos, tal que el vecino más cercano tenga mayor influencia que los demás. Empleamos para esto una función exponencial de la siguiente forma:

$$\omega_{exp, ij} = \begin{cases} 0 & \text{if } i = j \\ e^{-2*dist_{ij}} & \text{caso contrario} \end{cases}, \quad (2)$$

donde  $dist_{ij}$  es la distancia en kilómetros entre los alojamientos  $i$  y  $j$ .

Como es habitual en econometría espacial, la matriz  $\mathbf{W}$  se estandariza por filas tal que la suma de cada fila es igual a la unidad. Alternativamente, se llevó a cabo un análisis de sensibilidad utilizando la inversa de distancia como función alternativa; sin embargo, este modelo produjo resultados similares al de la función exponencial. Para mayor detalle véase el Apéndice B.

Como mencionamos anteriormente, el índice del delito podría ser endógeno según la especificación de la ecuación (1). La relación causal entre los precios de Airbnb y el delito puede fluir en ambas direcciones. Áreas urbanas con altos precios pueden atraer a una mayor delincuencia; por lo tanto, el número de delitos puede ser un efecto de los precios de una zona o región. Nuestra revisión de la literatura muestra una serie de estudios que incluyen efectos espaciales para Airbnb, pero no consideran al delito como potencialmente endógeno (Tang et al. 2019, Xu et al. 2019). Formalmente, se necesita contrastar si el delito observado en los vecindarios puede ser considerado exógeno en la ecuación de precios. Para esto, aplicamos el test de Wu-Hausman para endogeneidad, en combinación con los tests de validez de los instrumentos y el test de Sargan (Sargan 1958). La metodología detrás de estos tests es de tipo estándar en econometría (véase Wooldridge 2002, pp. 129-137). Nuestros instrumentos provienen de la distribución espacial del delito observada en el periodo previo al analizado (Discutiremos los detalles de esta estrategia en la próxima sección).

Otro problema es la posible presencia de dependencia espacial. Bajo la ecuación (1), utilizamos el test I de Moran (Moran 1950) cuya hipótesis nula es no-autocorrelación espacial en el término de error  $\mathbf{u}$ . Si el test rechaza la hipótesis nula, es posible especificar tres canales de dependencia espacial mediante la creación de una matriz de contactos espaciales  $\mathbf{W}$ : (1) la inclusión de dependencia espacial global a través de una variable espacial endógena ( $\mathbf{W} \times \ln(\mathbf{price})$ ); (2) la inclusión de dependencia espacial local a través de las variables explicativas ( $\mathbf{WX}$ ); y (3) la inclusión de una dependencia espacial global en el término de error ( $\mathbf{W} \times \mathbf{u}$ ). La alternativa (2), variables espaciales locales ( $\mathbf{WX}$ ), puede ser reemplazada por variables contextuales que busquen capturar efectos similares y es la estrategia que seguiremos en la sección aplicada. Por esta razón, las dos principales alternativas para incorporar los efectos espaciales son los canales (1) y (3).

En vez que estimar los dos canales por separado, es posible utilizar un modelo espacial general que incluya tanto la dependencia espacial endógena,  $\mathbf{W} \times \ln(\mathbf{price})$ , como la correlación espacial en el término de error,  $\mathbf{W} \times \mathbf{u}$ . A este modelo se lo conoce como *SARAR* (Kelejian & Prucha 2010):

$$SARAR : \quad \ln(\mathbf{price}) = \rho \mathbf{W} \times \ln(\mathbf{price}) + \mathbf{X}\beta + \gamma \ln(\mathbf{crime}) + \mathbf{u}, \quad \mathbf{u} = \lambda \mathbf{W} \times \mathbf{u} + \boldsymbol{\varepsilon}, \quad (3)$$

donde  $\rho$  y  $\lambda$  capturan la dependencia espacial de la variable endógena y en el término de error respectivamente;  $\varepsilon$  es un término de innovación idiosincrático con cada elemento independiente e idénticamente distribuido, *i.i.d.*  $(0, \sigma_i^2)$ . Además, se considera que el término de innovación idiosincrático pueda ser heterogéneo.

En la ecuación (3), si el parámetro de  $\mathbf{W} \times \ln(\mathbf{price})$  no es significativo, obtenemos un modelo de error espacial, *SEM*:

$$SEM : \quad \ln(\mathbf{price}) = \mathbf{X}\beta + \gamma \ln(\mathbf{crime}) + \mathbf{u}, \quad \mathbf{u} = \lambda \mathbf{W} \times \mathbf{u} + \varepsilon. \quad (4)$$

De la misma forma, si el parámetro de  $\mathbf{W} \times \mathbf{u}$  no resulta significativo, podemos estimar un modelo de rezago espacial, *SLM*:

$$SLM : \quad \ln(\mathbf{price}) = \rho \mathbf{W} \times \ln(\mathbf{price}) + \mathbf{X}\beta + \gamma \ln(\mathbf{crime}) + \mathbf{u}. \quad (5)$$

Los modelos que incluyen al coeficiente  $\rho$ , con una especificación de rezago espacial como en la ecuaciones (3) y (5), son propuestos generalmente desde un modelo que proviene de la teoría. Sin embargo, Anselin & Lozano-Gracia (2009, p. 1221) mencionan que: “Desde una perspectiva puramente pragmática, uno puede discutir que la especificación por rezagos espaciales permite filtrar una fuerte tendencia espacial (similar al pre-blanqueo en series de tiempo), es decir, asegurando una inferencia apropiada de los coeficientes  $\beta$ .” Siguiendo esta recomendación, consideramos al modelo SLM como relevante ya que nuestro principal interés radica en una estimación consistente de los coeficientes, en particular, del coeficiente  $\gamma$ .

Los modelos SEM y SARAR incluyen el término  $\mathbf{W} \times \mathbf{u}$ . La inclusión del término de error espacial podría ayudar a evaluar la magnitud del agrupamiento espacial de los precios de alojamientos no explicados por las variables explicativas. En esta misma línea, el coeficiente  $\lambda$  captura el efecto espacial de las variables explicativas omitidas o efectos espaciales no observados (Anselin 2002).

Hay varias alternativas de estimación de los modelos espaciales. En nuestro caso, utilizaremos GS2SLS ya que tiene la ventaja de no requerir ningún supuesto sobre la distribución del término de error (Kelejian & Prucha 1998, 1999, 2007, 2010) y de permitir la introducción de la corrección de la endogeneidad de las variables explicativas, como en el caso del delito. Además, Kelejian & Prucha (2010) extienden la estimación GS2SLS para obtener una versión robusta ante una heteroscedasticidad desconocida.

### 3. Resultados

#### 3.1. Estadísticas descriptivas y mapas

El conjunto de datos original de Airbnb contiene 10116 observaciones. Mediante el chequeo de los datos, se pudo detectar diferentes inconsistencias entre precios y características, generando la eliminación de 220 observaciones (2.2 % de los datos originales), obteniendo un conjunto de datos limpios de 9896 observaciones.

Además, se detectaron valores atípicos (*outliers* espaciales) mediante una secuencia de herramientas espaciales. En primer lugar, se utilizó un diagrama de caja para detectar valores extremos en los precios de los alojamientos resultando en 150 observaciones de alto valor (no se encontraron *outliers* de valores extremos bajos). En segundo lugar, se exploró la distribución espacial de estas 150 observaciones en combinación con el Indicador Local de Asociación Espacial (mapa LISA), usando el software GeoDa (Anselin et al. 2006). Esta técnica nos permitió detectar que únicamente 7 alojamientos mostraban precios extremos y que no pertenecían a un agrupamiento (*cluster* espacial) de precios altos, teniendo en cuenta el vecindario geográfico. Estas observaciones fueron identificadas con una variable binaria (DUMMY).

El Cuadro 2 reporta las estadísticas descriptivas de las principales variables bajo estudio. Las variables explicativas son clasificadas por características delictivas, características propias de los alojamientos y variables contextuales, y variables de control. Para GASTRO. COUNT y ATRAC. COUNT, combinamos distintas fuentes de datos utilizando Sistemas de Información Geográfica (GIS, en inglés), donde se aplicó un *buffer* (zona de influencia) a cada alojamiento para capturar el grado de exposición a las características ambientales de los atractivos turísticos (monumentos y lugares históricos) y de los sitios gastronómicos (restaurante, bar, café, delivery y take-away, sandwicherías y vinotecas). Para construir estas variables, generamos un *buffer* de 500 metros para cada alojamiento.

Cuadro 2: Estadísticas descriptivas de las variables.

Variables		Obs.	Media	S. D.	Min.	Max.
Dependiente	LN (PRICE)	9896	6.79	0.72	4.80	10.10
Indicador	LN (CHI)	9896	13.39	0.59	7.04	14.46
Delictivo	ACCOM 3-4	9896	0.34	0.48	0.00	1.00
	ACCOM 5+	9896	0.08	0.27	0.00	1.00
	BEDROOM	9896	1.12	0.82	0.00	10.00
	BATHROOM	9896	1.26	0.62	0.00	8.00
Características propias /contextuales	REVIEWS	9896	13.94	27.26	0.00	360.00
	PRIVATE ROOM	9896	0.18	0.38	0.00	1.00
	SHARED ROOM	9896	0.02	0.14	0.00	1.00
	GASTRO. COUNT	9896	45.48	37.55	0.00	244.00
	TRAIN DISTANCE	9896	1.02	0.48	0.01	2.36
	CBD DISTANCE	9896	3.97	2.51	0.01	13.40
	ATRAC. COUNT	9896	4.69	13.51	0.00	93.00
	LN (AREA PARK)	9896	4.47	1.12	-0.15	7.91
Controles	LN (POP. AREA)	9896	9.99	0.96	0.00	11.64
	DUMMY	9896	0.00	0.03	0.00	1.00

Nota: Estimaciones propias utilizando diferentes fuentes.

La variable TRAIN DISTANCE mide los kilómetros entre alojamientos y la estación de trenes más cercana. De la misma manera, CBD DISTANCE, mide la distancia en kilómetros al centro neurálgico de la ciudad desde cada alojamiento (distancia hasta el Obelisco).

Para la incorporación de los espacios verdes, se consideró un *buffer* de 1000 metros, una distancia razonable para caminar, y se computó la suma de áreas verdes dentro del *buffer*. Posteriormente, aplicamos una escala logarítmica.

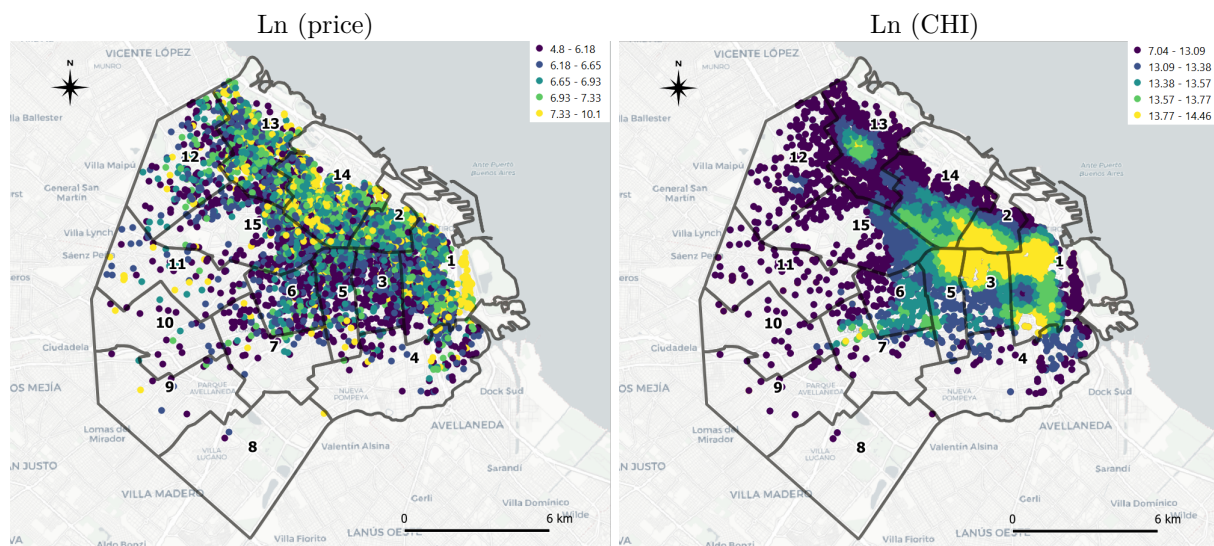
Todas estas medidas de GIS fueron realizadas utilizando QGIS software (<https://qgis.org/en/site/index.html>) y RStudio (<https://www.r-project.org/>, <https://rstudio.com/>), incluyendo el uso de paquetes como rgeos (Bivand & Rundel 2019), sp (Pebesma & Bivand 2005), rgdal (Bivand et al. 2019), GISTools (Brunsdon & Chen 2014), foreign, utils y base (R Core Team 2020).

Para el cómputo del indicador de delito, CHI, se procedió de la misma manera que la comentada para las variables contextuales. Para cada tipo de delito, se generó un *buffer* de 500 metros para cada alojamiento y se contabilizó la cantidad de delitos por cada tipo. Posteriormente, se calculó el CHI según la ponderación promedio de años. Este proceso fue realizado para el año 2017. Todas estas variables se transformaron a la escala logarítmica. Un análisis de sensibilidad del *buffer* se realiza en la siguiente sección.

Como variables de control se incluyó la densidad poblacional, además de la variable dummy para controlar los valores atípicos. La densidad poblacional utiliza información del Censo 2010, la información más reciente, y la escala considerada es a nivel de radio censal (unidad de organización de información geográfica más pequeña).

La Figura 2 presenta los mapas de cuantiles de las principales variables de interés: logaritmo de precios y el logaritmo del CHI. En comparación con la Figura 1, los valores del logaritmo del CHI son considerados solamente si pertenecen al *buffer* o zona de influencia de los alojamientos de Airbnb.

Figura 2: Distribución espacial del logaritmo de precios y de CHI para alojamientos Airbnb



### 3.2. Resultados Econométricos

En esta sección presentamos los principales resultados de nuestro modelo. Nuestra estrategia de selección de modelo fue inicialmente contrastar si el modelo no espacial, bajo MCO, mostraba evidencia de endogeneidad del indicador de delitos y de efectos espaciales omitidos. Una vez detectados ambos problemas, se procedió a estimar el SARAR VI, instrumentalizando al índice del delito. En la especificación SARAR no se encontró evidencia de efectos espaciales en el término de error pero si en el término del rezago espacial endógeno,  $\mathbf{W} \times \ln(\text{price})$ . Dados estos resultados, el modelo finalmente elegido fue el SLM VI. Para comparar los resultados con otros modelos menos complejos, el Cuadro 3 presenta cuatro modelos que van aumentando en complejidad: MCO, VI, SLM, SLM VI.

Para los modelos VI y SLM VI, utilizamos como instrumentos la distribución espacial del LN (CHI) observada en el año previo, 2016. Para construir estos instrumentos, en primer lugar se estimó el CHI para el 2016 y posteriormente se generó un rezago espacial del CHI para el mismo año, esto es, el valor promedio del índice de daño del delito para los alojamientos vecinos ( $\mathbf{W} \times \text{LN}(\text{CHI})_{2016}$ ). Además, se generó otro instrumento incorporando a los vecinos de los vecinos, un rezago espacial de segundo orden,  $\mathbf{W}^2 \times \text{LN}(\text{CHI})_{2016}$ . La Figura 3 muestra la idea principal para ambos instrumentos: los valores del LN (CHI) de 2017 para una observación central es instrumentalizada con la media de los vecinos en el año previo (Figura 3, gráfico de la izquierda), y la media de los vecinos de los vecinos del LN (CHI) (Figura 3, gráfico de la derecha) en 2016.

Figura 3: Instrumentos espaciales de 2016 para la variable LN (CHI) de 2017



El Cuadro 3 presenta los principales resultados. Todos modelos incluyen efectos fijos (FE) del nivel más bajo de agregación, es decir, por barrio. Para incluir estos efectos, se utilizó como umbral un mínimo de 2 observaciones para cada barrio.

Los resultados del modelo MCO, Cuadro 3, revelan una asociación negativa entre el valor de daño del delito y los precios de los alojamientos, después de ajustar por las características de los alojamientos y controlar por variables contextuales. El test I de Moran detecta una autocorrelación global en los residuos significativa.

Cuadro 3: Resultados de estimaciones bajo los diferentes modelos.

Modelos	MCO	VI	SLM	SLM VI
LN (CHI)	-0.042***	-0.093***	-0.032*	-0.064***
ACCOM. 3-4	0.161***	0.161***	0.161***	0.162***
ACCOM. 5+	0.462***	0.463***	0.457***	0.458***
BEDROOM	0.102***	0.102***	0.102***	0.102***
BATHROOM	0.200***	0.199***	0.199***	0.198***
PRIVATE ROOM	-0.961***	-0.961***	-0.952***	-0.952***
SHARED ROOM	-1.285***	-1.283***	-1.274***	-1.273***
REVIEWS	-0.001***	-0.001***	-0.001***	-0.001***
GASTRO. COUNT	0.001***	0.001***	0.001***	0.001***
TRAIN DISTANCE	-0.051***	-0.049***	-0.042***	-0.041***
CBD DISTANCE	-0.007	-0.012	-0.005	-0.008
ATRACC COUNT	-0.001*	-0.001**	-0.001	-0.001*
LN (AREA PARK)	0.005	-0.002	-0.003	-0.002
LN (POP. AREA)	-0.019***	-0.016***	-0.014**	-0.012**
DUMMY	1.270***	1.266***	1.343***	1.339***
CONSTANT	7.393***	8.043***	6.351***	6.795***
<b>W × ln (price)</b>			0.135***	0.132***
I de Moran (Chi2)	123.86***			
Wu-Hausman (F-test)		19.49***		
F-test instrumentos		4011.31***		
Sargan (Chi2)		1.18		
<i>R<sup>2</sup> – ajustado.</i>	62.79	62.75	63.11	63.10

Nota: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Todos los modelos incluyen barrios FE bajo errores estándar robustos.

Para detectar la posible endogeneidad, el Cuadro 3 reporta la estimación por VI. Los resultados al final del cuadro incluyen el test de Wu-Hausman (F-test) para detectar si LN (CHI) necesita ser tratado como un regresor endógeno. Se rechaza la hipótesis nula de exogeneidad del índice de daño del delito. El F-test instrumentos muestra la validez de los dos instrumentos propuestos, confirmando que los mismos son lo suficientemente fuertes. Si los instrumentos son débiles el modelo VI puede ser más sesgado que un modelo sin control por endogeneidad (para una revisión reciente ver Andrews et al. 2019). El resultado del test de Sargan confirma la validez de los instrumentos. Esto implica que los resultados reportados en la columna 2, VI, son superiores a los de la columna 1, MCO, pero no incluyen la corrección por dependencia espacial.

Las últimas dos columnas presentan los modelos espaciales. Tal como se comentó previamente, se estimó un modelo SARAR en donde el coeficiente espacial del error resultó no significativo. Esto genera que el modelo espacial se reduzca a un SLM. En el Cuadro 3, se muestra el modelo con corrección espacial, SLM, y el modelo espacial que además instrumentaliza al delito, SLM VI. En geneal los coeficientes de las variables



condicionales se mantienen estables entre los modelos y la principal corrección se observa en el coeficiente de LN (CHI), como es de esperar.

El ajuste de los modelos es bueno, indicando que más del 62 % de la variación en precios de Airbnb está explicado por las variables incluidas.

## 4. Discusión

En esta sección discutiremos el impacto del delito sobre los precios según las estimaciones alternativas propuestas. Para poder comparar cada resultado inferencial, es necesario aclarar que los modelos espaciales que incluyen una variable endógena  $\mathbf{W} \times \ln(\mathbf{price})$  modifican la interpretación marginal. Como señalan LeSage & Pace (2009), la interpretación de estos modelos requiere tener en cuenta la derivada parcial del cambio en las variables explicativas para hacer una correcta interpretación. Por ejemplo, si expresamos el valor esperado de la forma reducida del modelo (5), considerando a LN (CHI) como única variable explicativa, tenemos:

$$E[\ln(\mathbf{price})] = (\mathbf{I} - \rho\mathbf{W})^{-1} [\gamma \ln(\mathbf{CHI})] + \mathbf{R}, \quad (6)$$

donde  $E$  es el operador expectativa y  $\mathbf{R} = (\mathbf{I} - \rho\mathbf{W})^{-1} [\alpha\mathbf{1} + \mathbf{X}\beta]$ . Luego, la matriz de las derivadas parciales de  $E[\ln(\mathbf{price})]$  con respecto a la única variable explicativa  $\ln(\mathbf{CHI})$  desde la unidad 1 hasta la  $n$  puede ser representada de la siguiente forma:

$$\begin{bmatrix} \frac{\partial E[\ln(\mathbf{price})]}{\partial \ln(\mathbf{CHI})_1} & \dots & \frac{\partial E[\ln(\mathbf{price})]}{\partial \ln(\mathbf{CHI})_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial E[\ln(\mathbf{price})_1]}{\partial \ln(\mathbf{CHI})_1} & \dots & \frac{\partial E[\ln(\mathbf{price})_1]}{\partial \ln(\mathbf{CHI})_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial E[\ln(\mathbf{price})_n]}{\partial \ln(\mathbf{CHI})_1} & \dots & \frac{\partial E[\ln(\mathbf{price})_n]}{\partial \ln(\mathbf{CHI})_n} \end{bmatrix}, \quad (7)$$

$$= (\mathbf{I} - \rho\mathbf{W})^{-1} \begin{bmatrix} \gamma & 0 & \dots & 0 \\ 0 & \gamma & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \gamma \end{bmatrix}, \quad (8)$$

donde  $\gamma$  es la medida del cambio en el precio de un alojamiento en relación con el cambio en el LN (CHI), o la elasticidad del delito en nuestra especificación.

La expresión en la ecuación (8) muestra el efecto total esperado por un cambio unitario en la variable explicativa, que puede ser descompuesto en efecto directo e indirecto. El directo captura el efecto propio en cada alojamiento de un cambio unitario en LN (CHI) del vecindario, la diagonal principal de  $(\mathbf{I} - \rho\mathbf{W})^{-1} [\gamma]$ . Como este efecto es particular para cada unidad, LeSage & Pace (2009) sugieren utilizar el valor promedio de los elementos de la diagonal principal. Los efectos indirectos, o *spillover* espacial, cuantifican el impacto

de la interdependencia espacial de los alojamientos vecinos, los elementos de  $(\mathbf{I}_n - \rho \mathbf{W})^{-1} [\gamma]$  por fuera de la diagonal principal. Nuevamente, hay un valor particular para cada alojamiento que depende del número de vecinos y su distancia. LeSage & Pace sugieren utilizar el valor promedio de la suma de los elementos por fuera de la diagonal como resumen de este tipo de efecto.

Adicionalmente, la significancia de estos efectos promedio pueden ser obtenidos mediante simulaciones de Monte Carlo. Un punto importante es que los efectos deben ser interpretados como efectos de largo plazo. Esto es debido a que el término  $(\mathbf{I}_n - \rho \mathbf{W})^{-1}$  puede ser escrito como una progresión geométrica:  $(\mathbf{I} - \rho \mathbf{W})^{-1} = \mathbf{I} + \rho \mathbf{W} + \rho^2 \mathbf{W}^2 + \rho^3 \mathbf{W}^3 + \dots$  (conocido como el multiplicador de Leontieff). Bajo condiciones normales,  $-1 < \rho < 1$ , la progresión va a converger a un estado estacionario.

Cuadro 4: Resultados de las estimaciones bajo los diferentes modelos.

Modelos	MCO	VI	SLM	SLM VI
Efecto directo (LN (CHI))	-0.042**	-0.093***	-0.032*	-0.064***
Efecto indirecto (LN (CHI))			-0.005*	-0.010***
Efecto total (LN (CHI))	-0.042**	-0.093***	-0.037*	-0.074***

Nota: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Todos los modelos incluyen FE por barrios y errores std. robustos.

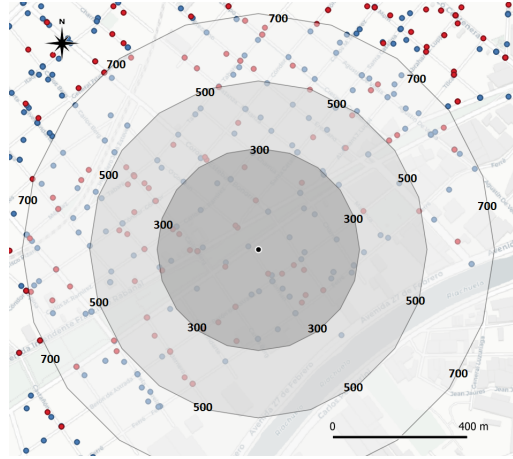
El Cuadro 4 muestra los resultados inferenciales para cada uno de los modelos focalizando la discusión en el impacto del delito. El efecto directo de los modelos espaciales es equivalente al coeficiente de los modelos no-espaciales. Puede observarse que en todos los casos, el impacto es negativo con valores inferiores (en valor absoluto) para la estimación SLM y MCO. Los modelos que incluyen la corrección por endogeneidad, VI y SLM VI muestran valores mayores (en valor absoluto).

El impacto directo de nuestro principal modelo, SLM VI, presenta un valor intermedio entre la estimación MCO y VI. Este resultado implica que un incremento del 10% en el CHI en el vecindario, genera una disminución de los precios de un 0.64%. Si se incorpora el efecto contagio, este impacto alcanza a una reducción del 0.74% en el precio. Este resultado es inferior, en valor absoluto, del impacto estimado por el modelo VI que es próximo a un 1%, 0.93% de reducción del precio ante el incremento del daño del delito de un 10% en el área.

Estos resultados no pueden compararse con otro estudio similar ya que es el primero que utiliza esta doble corrección (por endogeneidad y por dependencia espacial). El estudio más cercano en términos de metodología y especificación es el trabajo de Ceccato & Wilhelmsson (2011). Estos autores muestran, para el mercado de departamentos residenciales en Suecia, que el impacto total del crimen (no utiliza un indicador ponderado) de un incremento del 10% en el área genera una reducción esperada del 0.4% en el precio (Ceccato & Wilhelmsson 2011, p. 94). Nuestros resultados para el caso de alojamientos en Airbnb muestran una mayor sensibilidad al crimen.

Para comprobar la robustez de las estimaciones, a continuación se propone modificar la distancia de la zona de influencia utilizando *buffers* de 300 y 700 metros. Gráficamente, la modificación del *buffer* afecta a la densidad del delito para cada alojamiento (Figura 4).

Figura 4: Ejemplo de buffers (zonas de influencias)



En el Cuadro 5 se presentan los principales coeficientes de interés para el modelo SLM VI con buffers de 300, 500 y 700 metros. Puede observarse que en los tres modelos, el signo y significancia de la variable de LN (CHI) se mantiene. El coeficiente de dependencia espacial endógena es estable, entorno a 0.130. La bondad de ajuste de los modelos es también buena, en todos los casos superior al 63 %.

Cuadro 5: Resultados de estimaciones bajo diferentes tamaños de buffers.

Modelos	SLM VI 300	SLM VI 500	SLM VI 700
LN (CHI)	-0.046***	-0.064***	-0.054***
$\mathbf{W} \times \ln(\text{price})$	0.128***	0.132***	0.133***
Efecto Directo (LN (CHI))	-0.037***	-0.064***	-0.054***
Efecto Indirecto (LN (CHI))	-0.006***	-0.010***	-0.008***
Efecto Total (LN (CHI))	-0.043***	-0.074***	-0.062***
$R^2 - \text{ajustado}$	63.06	63.10	63.11

Nota: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Todos los modelos incluyen FE por barrios y errores std. robustos.

Respecto a los efectos directos e indirectos, los modelos de 300 y 700 metros muestran valores levemente inferiores en valor absoluto para estos efectos. Sin embargo, para estos modelos, se mantiene la significancia de todos los efectos y el efecto total oscila entre una disminución del 0.43 % y 0.62 % en el precio del alojamiento ante un incremento del 10 % en el daño del delito. El detalle completo para estos modelos puede observarse en el Apéndice C.

## 5. Conclusiones

El objetivo de este trabajo ha sido estudiar la relación entre el daño del delito y los precios de los alojamientos de Airbnb. Esta investigación provee una primera estimación del índice de daño del delito para alojamientos de Airbnb en Argentina. La principal motivación fue mostrar evidencia en el sector de economía colaborativa (*sharing economy*) y proveer una medición del impacto de las actividades delictivas en el sector, en particular, para Airbnb.

Nuestro trabajo contribuye en tres diferentes direcciones a la literatura existente. En primer lugar, utilizando información geográfica detallada, se construye un índice de daño del delito (CHI) que pondera los delitos de acuerdo al daño causado. Este índice es la primera vez que se propone para una ciudad en Argentina, según nuestra revisión de la literatura. En segundo lugar, nuestro estudio propone variables instrumentales de naturaleza espacio-temporal para sortear el problema de la endogeneidad del delito y provee evidencia robusta sobre el impacto del CHI en los precios de Airbnb. Por último, tanto el delito como la distribución de los precios de los alojamientos presentan estructura espacial y se consideran estas características incorporando efectos espaciales al modelo principal para corregir por dependencia espacial a los coeficientes estimados y a los errores estándares del modelo.

Los resultados inferenciales de los modelos hedónicos indican que los precios de alojamientos son afectados significativamente por el delito de zonas vecinas. El indicador de daño del crimen muestra un impacto significativo utilizando mínimos cuadrados ordinarios (MCO) y variables instrumentales (VI). Ambas estimaciones muestran problemas que sesgan la estimación de los coeficientes. El modelo de rezago espacial SLM VI (Spatial Lag Model) con Variables Instrumentales, y corrección del error por heteroscedasticidad espacial es un modelo superador a los dos previos. El impacto total de CHI en los precios utilizando el modelo SLM VI es de -0.074, descompuesto en un efecto propio (efecto directo) de -0.064 y un derrame espacial (efecto indirecto) de -0.010. El efecto directo, neto de la dependencia espacial, presenta un valor que se ubica entre las estimaciones por MCO y VI, y el efecto total es cercano al coeficiente de VI. Este resultado muestra el sesgo de los modelos de VI cuando se omite la dependencia espacial. Además, el SLM VI permite conocer el efecto indirecto, conocido como *spillover* espacial, que captura el efecto del contagio en los precios causado por el crimen en el vecindario.

Hay algunas limitaciones del presente trabajo que deben ser mencionadas. El indicador de daño del delito, CHI, solo incluye cinco categorías y no incluye delitos contra la propiedad. Además, los principales resultados provienen de un corte transversal de 2017 y 2016, teniendo todas las limitaciones que son habitualmente mencionadas para este tipo de datos. Futuras investigaciones deberían incluir información de la dinámica temporal para tener una mejor comprensión del fenómeno investigado. Además, es posible considerar añadir información sobre el nivel de inseguridad subjetiva del usuario de Airbnb mediante encuestas.

A pesar de estas limitaciones, consideramos que estos resultados son de importancia para comprender el impacto de las actividades delictivas en el sector de la economía colaborativa, y en particular para el caso de Airbnb.

## Referencias

- Adamiak, C., Szyda, B., Dubownik, A. & García-Álvarez, D. (2019), ‘Airbnb offer in Spain - spatial analysis of the pattern and determinants of its distribution’, *ISPRS international journal of geo-information* **8**(3), 155.
- Allen, J. (1999), *Crime against international tourists*, Sydney, New South Wales, Australia: Bureau of Crime Statistics and Research.
- Andersen, H. A. & Mueller-Johnson, K. (2018), ‘The Danish crime harm index: how it works and why it matters’, *Cambridge Journal of Evidence-Based Policing* **2**(1-2), 52–69.
- Andrews, I., Stock, J. H. & Sun, L. (2019), ‘Weak instruments in instrumental variables regression: Theory and practice’, *Annual Review of Economics* **11**, 727–753.
- Anselin, L. (1988), *Spatial econometrics: Methods and models*, Vol. 4, Springer Netherlands.
- Anselin, L. (2002), ‘Under the hood issues in the specification and interpretation of spatial regression models’, *Agricultural Economics* **27**(3), 247–267.
- Anselin, L. & Le Gallo, J. (2006), ‘Interpolation of air quality measures in hedonic house price models: spatial aspects’, *Spatial Economic Analysis* **1**(1), 31–52.
- Anselin, L. & Lozano-Gracia, N. (2009), Spatial hedonic models, in T. Mills & P. Kerry, eds, ‘Palgrave Handbook of Econometrics. Volume 2: Applied Econometrics’, Springer, pp. 1213–1250.
- Anselin, L., Syabri, I. & Kho, Y. (2006), ‘GeoDa: An introduction to spatial data analysis’, *Geographical analysis* **38**(1), 5–22.
- Belk, R. (2014), ‘You are what you can access: Sharing and collaborative consumption online’, *Journal of business research* **67**(8), 1595–1600.
- Bivand, R., Keitt, T. & Rowlingson, B. (2019), *rgdal: Bindings for the ‘Geospatial’ Data Abstraction Library*. R package version 1.4-8.  
**URL:** <https://CRAN.R-project.org/package=rgdal>
- Bivand, R. & Rundel, C. (2019), *rgeos: Interface to Geometry Engine - Open Source (‘GEOS’)*. R package version 0.5-2.  
**URL:** <https://CRAN.R-project.org/package=rgeos>

- Brunsdon, C. & Chen, H. (2014), *GISTools: Some further GIS capabilities for R*. R package version 0.7-4.  
**URL:** <https://CRAN.R-project.org/package=GISTools>
- Ceccato, V. & Wilhelmsson, M. (2011), ‘The impact of crime on apartment prices: Evidence from stockholm, sweden’, *Geografiska Annaler: Series B, Human Geography* **93**(1), 81–103.
- Chesney-Lind, M. & Lind, I. Y. (1986), ‘Visitors as victims crimes against tourists in hawaii’, *Annals of Tourism Research* **13**(2), 167–191.
- Conway, D., Li, C., Wolch, J., Kahle, C. & Jerrett, M. (2010), ‘A spatial autocorrelation approach for examining the effects of urban greenspace on residential property values’, *The Journal of Real Estate Finance and Economics* **41**(2), 150–169.
- Curtis-Ham, S. & Walton, D. (2018), ‘The new zealand crime harm index: Quantifying harm using sentencing data’, *Policing: A Journal of Policy and Practice* **12**(4), 455–467.
- Dirección General de Estadísticas y Censos (2018), ‘Anuario estadístico ciudad de buenos aires’.
- Eck, J., Chainey, S., Cameron, J. & Wilson, R. (2005), *Mapping crime: Understanding hotspots*, National Institute of Justice.
- Fenimore, D. M. (2019), ‘Mapping harmspots: An exploration of the spatial distribution of crime harm’, *Applied geography* **109**, 102034.
- Gibbons, S. (2004), ‘The costs of urban property crime’, *The Economic Journal* **114**(499), F441–F463.
- Ho, T., Zhao, J. & Dooley, B. (2017), ‘Hotel crimes: An unexplored victimization in the hospitality industry’, *Security Journal* **30**(4), 1097–1111.
- Jones, P. & Groenenboom, K. (2002), ‘Crime in london hotels’, *Tourism and Hospitality Research* **4**(1), 21–35.
- Kärrholm, F., Neyroud, P. & Smaaland, J. (2020), ‘Designing the swedish crime harm index: an evidence-based strategy’, *Cambridge Journal of Evidence-Based Policing* pp. 1–19.
- Kelejian, H. & Prucha, I. (1998), ‘A generalized spatial two-stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances’, *The Journal of Real Estate Finance and Economics* **17**(1), 99–121.
- Kelejian, H. & Prucha, I. (1999), ‘A generalized moments estimator for the autoregressive parameter in a spatial model’, *International economic review* **40**(2), 509–533.
- Kelejian, H. & Prucha, I. (2007), ‘Hac estimation in a spatial framework’, *Journal of Econometrics* **140**(1), 131–154.
- Kelejian, H. & Prucha, I. (2010), ‘Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances’, *Journal of Econometrics* **157**(1), 53–67.

- Kong, Y., Wang, Y., Hajli, S. & Featherman, M. (2020), ‘In sharing economy we trust: Examining the effect of social and technical enablers on millennials’ trust in sharing commerce’, *Computers in Human Behavior* **108**, 105993.
- LeSage, J. & Pace, R. (2009), *Introduction to spatial econometrics*, CRC press.
- Moran, P. A. (1950), ‘Notes on continuous stochastic phenomena’, *Biometrika* pp. 17–23.
- Pebesma, E. J. & Bivand, R. S. (2005), ‘Classes and methods for spatial data in R’, *R News* **5**(2), 9–13.  
**URL:** <https://CRAN.R-project.org/doc/Rnews/>
- Priporas, C.-V., Stylos, N., Vedanthachari, L. N. & Santiwatana, P. (2017), ‘Service quality, satisfaction, and customer loyalty in airbnb accommodation in thailand’, *International Journal of Tourism Research* **19**(6), 693–704.
- R Core Team (2020), *foreign: Read Data Stored by 'Minitab', 'S', 'SAS', 'SPSS', 'Stata', 'Systat', 'Weka', 'dBase', ...* R package version 0.8-75.  
**URL:** <https://CRAN.R-project.org/package=foreign>
- Rosen, S. (1974), ‘Hedonic prices and implicit markets: product differentiation in pure competition’, *Journal of political economy* **82**(1), 34–55.
- Sargan, J. D. (1958), ‘The estimation of economic relationships using instrumental variables’, *Econometrica: Journal of the Econometric Society* pp. 393–415.
- Sherman, L., Neyroud, P. W. & Neyroud, E. (2016), ‘The cambridge crime harm index: Measuring total harm from crime based on sentencing guidelines’, *Policing: A Journal of Policy and Practice* **10**(3), 171–183.
- Sherman, L. W. (2007), ‘The power few: experimental criminology and the reduction of harm’, *Journal of Experimental Criminology* **3**(4), 299–321.
- Slee, T. (2018), ‘airbnb-data-collection’, <https://github.com/tomslee>.
- Tang, L. R., Kim, J. & Wang, X. (2019), ‘Estimating spatial effects on peer-to-peer accommodation prices: Towards an innovative hedonic model approach’, *International Journal of Hospitality Management* **81**, 43–53.
- Vazquez Brust, A. (2020), ‘Alojamientos listados por airbnb en buenos aires. dataset 2018’. Access June. Retrieved from <https://data.world/vazquez-brust>.
- Weisburd, D. (2015), ‘The law of crime concentration and the criminology of place’, *Criminology* **53**(2), 133–157.
- Wooldridge, J. (2002), *Econometric analysis of cross section and panel data*, Cambridge, Massachusetts: MIT Press.

Xu, Y.-H., Pennington-Gray, L. & Kim, J. (2019), 'The sharing economy: a geographically weighted regression approach to examine crime and the shared lodging sector', *Journal of travel research* **58**(7), 1193–1208.

Yang, Y. & Mao, Z. (2020), 'Location advantages of lodging properties: A comparison between hotels and airbnb units in an urban environment', *Annals of Tourism Research* **81**, 102861.

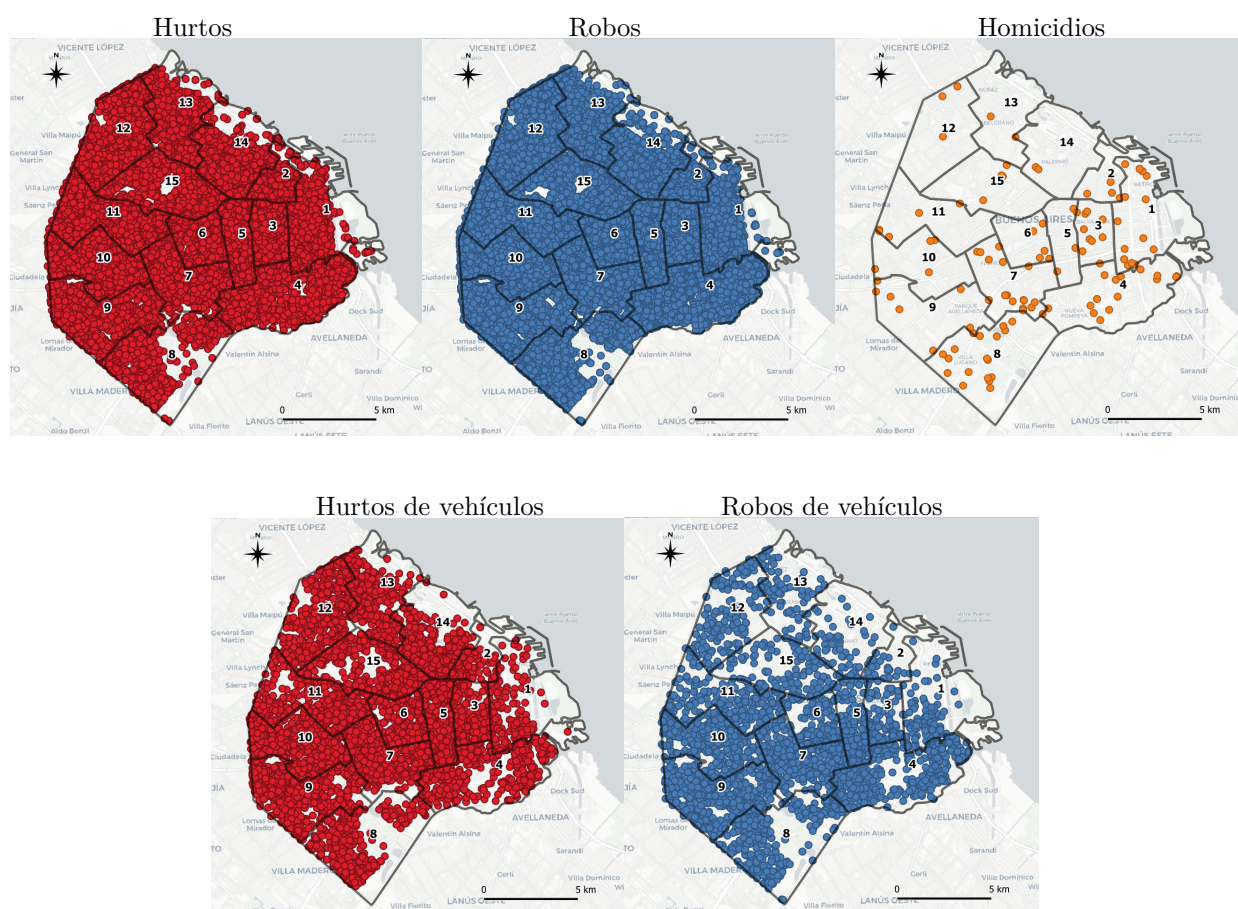
Zervas, G., Proserpio, D. & Byers, J. W. (2017), 'The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry', *Journal of marketing research* **54**(5), 687–705.

Zhao, J., Ho, T. & Chen, C.-y. (2004), 'Crimes against hotel visitors: An empirical study in miami-dade county, florida', *Journal of Hospitality Legal, Safety and Security Research* **2**, 1–12.

## Apéndices

### Apéndice A: Distribución espacial de los delitos

Figure A.1: Distribución espacial de los delitos en el año 2017





## Apéndice B: Chequeo de robustez de W

Para validar la robustez de los resultados espaciales, estimamos una especificación alternativa de la matriz de contactos espaciales. Para generar esta matriz, en vez de usar una función exponencial, se propone utilizar una función de inversa distancia:

$$w_{ij} = \begin{cases} \left(\frac{1}{d_{ij}}\right), & \text{si } i \neq j, \\ 0, & \text{otro caso.} \end{cases}$$

Usando esta especificación alternativa de **W**, los resultados de los modelos son presentados en la Tabla B.1.

Table B.1: Resultados de estimaciones de los modelos bajo W inversa

Modelos	OLS	VI	SLM	SLM VI
LN (CHI)	-0.042***	-0.090***	-0.034**	-0.068***
ACCOM. 3-4	0.161***	0.161***	0.161***	0.162***
ACCOM. 5+	0.462***	0.463***	0.458***	0.459***
BEDROOM	0.102***	0.102***	0.102***	0.102***
BATHROOM	0.200***	0.199***	0.199***	0.198***
PRIVATE ROOM	-0.961***	-0.961***	-0.953***	-0.953***
SHARED ROOM	-1.285***	-1.283***	-1.275***	-1.274***
REVIEWS	-0.001***	-0.001***	-0.001***	-0.001***
GASTRO. COUNT	0.001***	0.001***	0.001***	0.001***
TRAIN DISTANCE	-0.051***	-0.049***	-0.045***	-0.044***
CBD DISTANCE	-0.007	-0.012	-0.006	-0.009
ATRACC COUNT	-0.001*	-0.001**	-0.001	-0.001*
LN (AREA PARK)	0.005	-0.002	-0.003	-0.002
LN (POP. AREA)	-0.019***	-0.016***	-0.016***	-0.013**
DUMMY	1.270***	1.267***	1.327***	1.323***
CONSTANT	7.393***	8.007***	6.568***	7.027***
<b>W × ln (CHI)</b>			0.109***	0.107***
Direct effect (CHI)	-0.042**	-0.090***	-0.034**	-0.068***
Indirect effect (CHI)			-0.004**	-0.008***
Total effect (CHI)	-0.042**	-0.090***	-0.038**	-0.076***
I de Moran (Chi2)	113.93***			
Wu-Hausman (F-test)		21.31***		
F-test instruments		3923.08***		
Sargan (Chi2)		1.29		
<i>R</i> <sup>2</sup> – ajustado	62.79	62.76	63.10	63.08

Nota: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Todos los modelos incluyen FE por barrios y errores estándar robustos.

## Apéndice C: Estimaciones usando diferentes zonas de influencia (buffers)

Table C.1: Modelo principal ante diferentes buffers de delito

Modelos	SLM VI (300)	SLM VI (500)	SLM VI (700)
LN (CHI)	−0.037***	−0.064***	−0.054***
ACCOM. 3-4	0.161***	0.162***	0.161***
ACCOM. 5+	0.457***	0.458***	0.457***
BEDROOM	0.102***	0.102***	0.102***
BATHROOM	0.199***	0.198***	0.199***
PRIVATE ROOM	−0.953***	−0.952***	−0.952***
SHARED ROOM	−1.274***	−1.273***	−1.275***
REVIEWS	−0.001***	−0.001***	−0.001***
GASTRO. COUNT	0.001***	0.001***	0.001***
TRAIN DISTANCE	−0.043***	−0.041***	−0.040***
CBD DISTANCE	−0.005	−0.008	−0.008
ATRACC COUNT	−0.000	−0.001*	−0.001*
LN (AREA PARK)	0.002	−0.002	−0.001
LN (POP. AREA)	−0.013**	−0.012**	−0.014**
DUMMY	1.341***	1.339***	1.343***
CONSTANT	6.379***	6.795***	6.687***
<b>W × ln (price)</b>	0.136***	0.132***	0.133***
<i>R<sup>2</sup> − ajustado</i>	63.06	63.10	63.11

Nota: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Todos los modelos incluyen FE por barrios y errores estándar robustos.