

Tres ensayos empíricos de economía aplicada

por

Juan Sebastián Ordóñez Herrera

Enviado en cumplimiento  
de los  
requisitos para el grado  
Doctor en Economía

Supervisado por  
Profesor Juan Daniel Oviedo Arango

Facultad de Economía  
Doctorado en Economía  
Universidad del Rosario

Bogotá - Colombia

2022

*Este trabajo y estos años de esfuerzo son de Ana María y Santiago, a quienes quiero agradecer su presencia, su apoyo incondicional y la paciencia de entender que culminar este proceso nos dará oportunidades como familia. Espero que la vida del conocimiento los apasione como a mí, y que vean en ella un mundo de oportunidades para ser felices*

*“El da esfuerzo al cansado, y multiplica las fuerzas al que no tiene ningunas”. (Isaías 40:29)*

## Curriculum Vitae

Juan Sebastián Ordoñez es Asesor de la Dirección General y Coordinador de los Grupos Internos de Trabajo de Registros Administrativos y Pobreza y Desigualdad del Departamento Administrativo Nacional de Estadística – DANE. Es candidato a doctor en economía de la Universidad del Rosario (Bogotá, Colombia), Magíster en Economía, y Economista de esta universidad. Antes de trabajar en el DANE, Juan Sebastián Ordoñez se desempeñó como docente e investigador en la Facultad de Economía de la Universidad del Rosario y ha trabajado para varias firmas consultoras en temas de política de la competencia, economía comercial, entre otros.

## Agradecimientos

Después de largas noches de estudio, visitas académicas y presentaciones, culminé quizá lo que más me ha costado en la vida. Este proceso tuvo frustraciones, tristezas, cambios de ánimo y afanes que aportaron aprendizajes diferentes a los académicos, como poner a prueba la estabilidad de mi familia durante meses eternos en los que me ausente. Sin embargo, la felicidad y tranquilidad de terminar cambiaron mis sensaciones donde espero reír y aportar al desarrollo de mi país y a la ciencia económica. Además, me permitió conocer muchas personas y ver el apoyo incondicional de muchas de ellas. A todos aquellos que participaron directa o indirectamente en el desarrollo de este proyecto les deseo lo mejor en la vida, en este mundo de conocimiento. Y en especial:

### *A mi familia*

A mi amado papá, gracias por guiarme, protegerme y enseñarme todo lo que sé en mi vida. A ti te debo lo que soy, gracias por creer siempre en tu ratón. Ha sido un placer conocerte en la vida.

A Hato María, quien, con su ejemplo, consejos, apoyo incondicional y sobre todo como mi mejor amiga me impulsó constantemente, además por darme a la sobrinita más tierna que cuando más lo necesité con sus juegos y abrazos me dieron energía para soportar.

A mi mamá, a mis tías, a mi hermano y a mí tío por anhelar esto como yo, por darme todo su apoyo durante estos años, por alentarme, alcahuetearme y quererme sin medida.

Al compañero de mi travesía quien durante años ha vivido esto como propio, y a quien le agradezco infinitamente su apoyo incondicional, el cariño en sus consejos, su cantaleta y su dureza en momentos que no podía desfallecer.

### *A mi equipo*

Siempre he tenido un gran equipo en los que debo agradecer los consejos, afecto y su ayuda. Además, por formar un grupo donde la relación no ha sido solo laboral, sino donde la amistad y solidaridad se mantuvo hasta el último día donde no pararon de apoyarme y alentarme. De este equipo debo agradecer a mi mejor amigo Sergio Montoya, por su solidaridad con su conocimiento. Y como no a José Lobo, Andrés García, Cristhyan Naranjo, John Quinchua, Gabriel Lombo y Juan Fernando Duque, quienes con sus comentarios y trabajo enriquecieron mis ideas.

A Juan Daniel Oviedo, quien me ha guiado y enseñado a ser profesional en estos arduos años, quien desde hace tiempo ha sido la persona que ha visto por mí, un futuro exitoso. Y que además de ser mi jefe, asesor de tesis, ha sido un gran apoyo y un amigo muy especial.

A las K, Moni, Mateo y a Camilo por acompañarme en los momentos finales dándome todo su apoyo por mi estabilidad, para poder culminar mi trabajo. A Diana J y Erika Jo quienes por años han estado para aconsejarme y escucharme en todo este proceso.

*A la Universidad del Rosario, Toulouse school of economics*

A Toulouse School of Economics, quienes abrieron sus puertas con sus comentarios a mi trabajo. A Marc Ivaldi por su paciencia y dedicación en sus comentarios.

A todos los profesores de la Universidad del Rosario, quienes, desde mi pregrado en Economía en su hermosa Facultad, me han enseñado como con un arduo trabajo colaborativo se consiguen las metas en la vida. Así mismo al grupo de Organización Industrial quienes con sus comentarios enriquecieron estos artículos. A mis estudiantes quienes día a día escucharon mis historias de hoteles y de medición económica. Al Departamento Administrativo Nacional de Estadística – DANE, y todos aquellos funcionarios que han hecho sentir en un mar de información y conocimiento.

A Carlos Sepúlveda y Carlos Ivan Cañón quienes me abrieron la puerta a la universidad del rosario y al final con sus comentarios enriquecieron mi trabajo. Y a Jorge Tovar quien desde mi primera propuesta de tesis aportó con sus valiosos comentarios

## Resumen

Tesis entregada para optar al título de Doctor de Economía de la Universidad del Rosario del candidato Juan Sebastián Ordoñez Herrera.

Esta tesis consta de tres estudios empíricos de economía aplicada, en los cuales se utilizan diferentes metodologías, técnicas y métodos estadísticos, así como diferentes fuentes de información entre las cuales se encuentran fuentes tradicionales y no tradicionales de datos. El primer capítulo tiene como objetivo comprender la relación entre diferenciación y resiliencia para la formulación e implementación de medidas de iniciativa pública y privada que permitan identificar anticipadamente riesgos para minimizar la fragilidad de ciertos mercados de alojamiento. En el sector hotelero existe un alto grado de diferenciación horizontal, que se puede presentar en la localización o la oferta de servicios que se adapta para los diversos tipos de clientes. Este evalúa los efectos heterogéneos que se presentan ante caídas exógenas de la demanda, de acuerdo con el nivel de diferenciación horizontal. Para esto se utilizan diferentes fuentes de datos que se integran, como lo son: i) Datos Administrativos, ii) encuestas, y iii) fuentes no tradicionales de información, obtenidas a través de búsquedas en portales web facilitadores para la venta de habitaciones de hotel y la API de Google. A partir de la estimación de especificaciones de diferencias en diferencias, los resultados sugieren una mayor resiliencia en aquellos establecimientos que realizan un mayor esfuerzo por diferenciarse geográficamente o en su oferta de servicios.

El segundo capítulo se propone estudiar el impacto del cambio temporal en la recolección de datos de la Gran Encuesta Integrada de Hogares (GEIH), de donde se obtiene información relacionada con los ingresos y cálculos de pobreza, como consecuencia de las medidas sanitarias implementadas por el Gobierno Nacional de Colombia entre marzo y julio del 2020. El confinamiento estricto implementado por el Gobierno Nacional como medida de contención a la expansión de la pandemia causada por el COVID-19 generó retos en los operativos de recolección de datos a través de encuestas de hogares. Como resultado, las encuestas con métodos de recolección presenciales migraron a modalidad remota, a través de encuestas telefónicas, lo cual podría haber modificado los posibles sesgos de reporte de variables como los ingresos. Este artículo estudia el efecto de cambio en el modelo de recolección de información en la Gran Encuesta Integrada de Hogares (GEIH) de Colombia sobre el reporte de ingresos laborales. Para ello, se explota la variación geográfica en la implementación de métodos de recolección

y la integración de la encuesta con un registro administrativo de seguridad social para cuantificar la variación en el reporte. Los resultados son relevantes para estudiar el impacto de los métodos de recolección en los sesgos de reporte de las encuestas y el uso métodos de integración de datos provenientes de diferentes fuentes. Este capítulo fue coautorado con Juan Daniel Oviedo, Andrés García Suaza y José Lobo.

Finalmente, el tercer capítulo tiene como objetivo analizar el modo en el que las plataformas de compras en línea influyen en la toma de decisiones de los viajeros al momento de elegir hoteles, así como con los efectos sobre su rentabilidad y la de su competencia. El creciente uso de plataformas para compras online de bienes experiencia cuya compra no es recurrente y en las que el oferente tiene privilegio de información, como los alojamientos, genera externalidades positivas, ya que las experiencias pasadas de los consumidores proveen información útil para la toma de decisiones por los viajeros y los hoteles competidores. De este modo, este artículo evidenciará, a través de un modelo espacial auto-regresivo Durbin, cómo la divulgación de información acerca de experiencias pasadas en hoteles competidores con atributos similares genera un efecto en la rentabilidad propia. Esta investigación utiliza una base de datos innovadora que vincula datos de encuestas de alojamiento y registros administrativos de turismo con datos recopilados por un web-crawler de las plataformas web facilitadoras y la API de Google Cloud.

# Índice general

Dedicatoria . . . . .	II
Curriculum Vitae . . . . .	III
Agradecimientos . . . . .	IV
Resumen . . . . .	VI
Lista de figuras . . . . .	IX
Lista de tablas . . . . .	X
<b>1. Diferenciación horizontal y resiliencia: el caso del sector de alojamiento en Bogotá-Colombia en tiempos de pandemia</b>	<b>XI</b>
1.1. Introducción . . . . .	XI
1.2. Literatura relacionada . . . . .	XIV
1.3. Datos . . . . .	XVIII
1.4. Estrategia de identificación . . . . .	XXVI
1.5. Resultados . . . . .	XXX
1.5.1. Cadenas de hoteles . . . . .	XXXIV
1.6. Robustez . . . . .	XXXVI
1.6.1. Especificación con variables geográficas y controles por diferenciación vertical . . . . .	XXXVII
1.6.2. Medida alternativa de intensidad de tratamiento . . . . .	XXXVIII
1.6.3. Medida de diferenciación por atributos . . . . .	XLI
1.7. Conclusiones . . . . .	XLIII
1.8. Referencias . . . . .	XLV
<b>2. Impacto del modelo de recolección sobre la captura de ingresos. Un estudio en los tiempos del COVID-19</b>	<b>LVII</b>
2.1. Introducción . . . . .	LVII
2.2. Cambios en los métodos de recolección en la encuesta de hogares . . . . .	LXI

2.3. Datos y estratégica empírica . . . . .	LXIII
2.3.1. Datos . . . . .	LXIII
2.3.2. Metodología . . . . .	LXV
2.4. Resultados . . . . .	LXVIII
2.5. Chequeos de robustez . . . . .	LXXI
2.6. Conclusiones . . . . .	LXXIV
2.7. Referencias . . . . .	LXXV
<b>3. Competencia a través de plataformas facilitadoras en el mercado de alojamiento: la relación de la diferenciación y la reputación con los beneficios del mercado hotelero en Bogotá-Colombia</b>	<b>LXXIX</b>
3.1. Introducción . . . . .	LXXIX
3.2. Plataformas y Diferenciación . . . . .	LXXXII
3.2.1. Aglomeración y competencia hotelera . . . . .	LXXXIV
3.2.2. Diferenciación por atributos en el mercado hotelero . . . . .	LXXXVI
3.3. Metodología y datos . . . . .	LXXXVI
3.3.1. Metodología . . . . .	LXXXVI
3.3.2. Datos . . . . .	LXXXIX
3.4. Resultados . . . . .	XCIII
3.5. Conclusiones . . . . .	XCIV
3.6. Referencias . . . . .	XCVII
<b>Appendices</b>	<b>CVIII</b>
<b>A. Chapter 1</b>	<b>CIX</b>
<b>B. Chapter 2</b>	<b>CXI</b>

# Índice de figuras

1.2. Distribución geográfica de establecimientos de alojamiento en Bogotá D.C.	XXIII
1.3. Variables dicotómicas para la construcción de las distancias por atributos	XXV
1.4. Número total de huéspedes por mes: residentes y no residentes . . . . .	XXVII
1.5. Rentabilidad, precio, tasa de ocupación promedio hoteles según residen- tes no residentes . . . . .	XXIX
1.6. Modelo sin efectos heterogéneos por diferenciación. . . . .	XXXI
1.7. Resultados para rentabilidad por trabajador. . . . .	XXXII
1.8. Resultados para precio por habitación. . . . .	XXXIII
1.9. Resultados para tasa de ocupación. . . . .	XXXIV
1.10. Resultados para rentabilidad por trabajador, usando pertenencia a ca- denas. . . . .	XXXV
1.11. Resultados para precio por habitación, usando pertenencia a cadenas.. .	XXXVI
1.12. Resultados para tasa de ocupación, usando pertenencia a cadenas. . . .	XXXVI
1.13. Resultados para rentabilidad por trabajador, precio por habitación y tasa de ocupación, usando especificación alternativa con efectos fijos geográficos y controles de diferenciación vertical (estrellas). . . . .	XXXVIII
1.14. Resultados para rentabilidad por trabajador, usando el porcentaje de huéspedes que viajan por negocios como medida de exposición, diferen- ciación medida por proximidad en atributos. . . . .	XL
1.15. Resultados para precio por habitación, usando el porcentaje de huéspe- des que viajan por negocios como medida de exposición, diferenciación medida por proximidad en atributos. . . . .	XL
1.16. Resultados para tasa de ocupación, usando el porcentaje de huéspedes que viajan por negocios como medida de exposición, diferenciación me- dida por proximidad en atributos.. . . . .	XLI

1.17. Resultados para rentabilidad por trabajador, diferenciación medida por proximidad en atributos por quintiles (1 y 2 vs. 3 y 4). . . . .	XLII
1.18. Resultados para precio por habitación, diferenciación medida por proximidad en atributos por quintiles (1 y 2 vs. 3 y 4). . . . .	XLII
1.19. Resultados para tasa de ocupación, diferenciación medida por proximidad en atributos por quintiles (1 y 2 vs. 3 y 4). . . . .	XLIII
2.1. Promedio de la diferencia en logaritmos de salarios entre GEIH y RELAB, para áreas metropolitanas, otras ciudades y zonas rurales por mes 2019-2021 . . . . .	LXVIII
2.2. . Impacto estimado del método de recopilación de datos por método CATI en los salarios informados . . . . .	LXIX
2.3. Impacto estimado por sexo y nivel educativo . . . . .	LXX
3.1. . . . . .	XCII
3.2. . . . . .	XCIII
A.1. Modelo sin efectos heterogéneos por diferenciación, usando porcentaje de huéspedes que viajan por negocios como variable de exposición. . . .	CX
B.1. Impacto estimado por rango de edad . . . . .	CXII
B.2. Impacto estimado por tipo de hogar . . . . .	CXIII

# Índice de cuadros

1.1. Atributos de los servicios del hotel y de las habitaciones . . . . .	XXII
1.2. Estadísticas descriptivas variables de resultados, de exposición a la pandemia y de diferenciación. . . . .	XXIV
2.1. Tabla 3. Impacto estimado del método de recopilación de datos por método CATI en los salarios informados . . . . .	LXXI
2.2. Tabla 4. Resultado ejercicios de robustez . . . . .	LXXIII
3.1. Atributos de los servicios del hotel y de las habitaciones . . . . .	XCI
3.2. Resultados . . . . .	XCV

# Capítulo 1

## Diferenciación horizontal y resiliencia: el caso del sector de alojamiento en Bogotá-Colombia en tiempos de pandemia

### 1.1. Introducción

Después de varios años de alto crecimiento del sector de alojamiento (OECD, 2020), el 2020 marcó un hito para el sector como consecuencia de la pandemia ocasionada por el coronavirus SARS-CoV-2. El 17 de abril, la Organización Mundial de Turismo (OMT) reportó que el “96 % de todos los destinos de viaje en el mundo habían impuesto restricciones de viaje como respuesta a la pandemia” (OMT, 2020). La mayoría de estas restricciones implicaron el cierre total o parcial de las fronteras estatales a los turistas. Si bien estas restricciones fueron aplicadas según el criterio de cada país, en términos generales siguieron las mismas líneas: prohibiciones de viaje, controles de visados y cuarentenas.

El sector de alojamiento estuvo afectado por estas restricciones sanitarias y la reacción de los viajeros a la evolución de la pandemia, lo que repercutió en aquellas economías dependientes de este sector. En Colombia, por ejemplo, la contribución del sector turístico al valor agregado bruto cayó del 2.8 % en 2019 a 1.0 % en 2020, de acuerdo con datos de la Organización para la Cooperación y el Desarrollo Económicos

(OCDE, 2022). Desde entonces, la recuperación de los establecimientos de alojamiento se ha dado de forma gradual y desigual. Por lo tanto, el estudio de las características asociadas con la rápida recuperación de algunos mercados y con la resiliencia que presentaron algunas firmas frente al choque de demanda que representó el cierre en el flujo de turistas, aporta elementos para el diseño de políticas públicas que tengan como objetivo disminuir el impacto negativo de cambios exógenos en la demanda.

En este sentido, uno de los elementos que puede afectar la vulnerabilidad de un establecimiento ante un choque de demanda es la diferenciación del producto que ofrece. En el sector hotelero se han encontrado evidencias de que las firmas reaccionan a incrementos en la incertidumbre de la demanda mediante el aumento de la variedad de productos ofrecidos (Chen et al., 2011), y que aumentos en los niveles de variedad en la oferta reducen el costo asociado a tener un exceso de capacidad instalada, condición particularmente inflexible en el sector (Carlton & Dana, 2008). En contravía de estos hallazgos, en el contexto de la crisis asociada al COVID-19, Lin & Chen (2022) encontraron una correlación negativa entre la evolución de los ingresos durante la pandemia y los niveles de variedad.

Este artículo desarrolla una nueva aproximación a la relación entre diferenciación horizontal y resiliencia a choques de demanda con un diseño empírico que aprovecha la exposición diferenciada al choque de demanda, con el propósito de allegar elementos de juicio que permitan organizar una estrategia de identificación del efecto causal. El estudio analiza el desempeño, a través de la tasa de ocupación y la rentabilidad, de los establecimientos de alojamiento hotelero en Bogotá durante el periodo 2019-2021. Adicionalmente, se presentan un conjunto novedoso de indicadores para describir la diferenciación horizontal de un mercado en términos geográficos y de las características de los servicios ofrecidos.

Abordamos esta pregunta usando información para el sector hotelero en la ciudad de Bogotá, Colombia. Esta ciudad concentra una alta proporción del mercado hotelero del país, con una participación del 26.6% en 2019. Adicionalmente, presentó una contracción significativa a consecuencia de la pandemia, pasando a tener una participación de tan solo 15.7% en 2021. Sin embargo, este debilitamiento fue heterogéneo, de forma que la variabilidad en desempeño a nivel de firma fue elevada, como se mostrará más adelante.

Uno de los desafíos metodológicos al comparar el desempeño de los establecimientos durante la pandemia según nivel de diferenciación es la heterogeneidad de las expo-

siones al choque de demanda que tuvieron dichos establecimientos. En ausencia de información adicional, no es posible identificar si un desempeño superior para un hotel dado es ocasionado por una perturbación de demanda específicamente más baja para el establecimiento, o por un mayor nivel de resiliencia. Dado que el nivel de exposición está correlacionado con el nivel de diferenciación, como se mostrará más adelante, es necesario buscar una metodología para diferenciar ambos efectos.

Para enfrentar este desafío, este trabajo explota la variabilidad entre firmas en términos de dos variables asociadas a la exposición al choque de demanda: la proporción prepandemia de viajeros provenientes del exterior y la proporción prepandemia de viajeros de negocios. Dado que las restricciones a la movilidad para viajeros internacionales fueron más estrictas, los hoteles que antes de la exposición atendían en mayor medida a este tipo de usuarios estaban más expuestos a la contracción en la demanda de servicios de alojamiento (esta idea ya ha sido verificada empíricamente por Duro et al. (2021) para España y por Pramana et al. (2021) para Bali). Por otra parte, la contracción en la demanda por parte de viajeros de negocios fue más fuerte que en el caso de quienes viajaban por otros motivos (principalmente ocio), por lo cual es de esperarse que los hoteles que se concentraban en este segmento tuvieran una mayor contracción de la demanda. La existencia de heterogeneidades en la intensidad de exposición al choque estudiado aunada con la variación temporal de esta permite el uso de una estrategia de diferencia en diferencias.

Para la aplicación de la metodología mencionada fue necesaria la preparación de un conjunto de datos que mezclan fuentes alternativas (*web scrapping* de plataformas web de reserva de alojamientos), registros administrativos y fuentes tradicionales (encuestas a establecimientos). La información sobre desempeño de los establecimientos, así como algunas variables de caracterización básica de los mismos, provienen de la Encuesta Mensual de Alojamiento, a cargo del Departamento Administrativo Nacional de Estadística de Colombia (DANE). Estos datos fueron vinculados y armonizados con un registro administrativo diligenciado anualmente por los mismos hoteles, el Registro Nacional Hotelero, a partir del cual se obtienen variables adicionales de caracterización, así como la ubicación geográfica, la cual permite describir a nivel censal la diferenciación horizontal en términos espaciales.

Para poder describir la diferenciación horizontal en términos de las características de los servicios ofrecidos fue necesario recurrir a fuentes no tradicionales, en particular la implementación de (*web scrapping* en dos de los sitios web facilitadores para la

venta de habitaciones de hotel más usados en Colombia. De esta manera, se capturó un conjunto de variables que dan cuenta de las características de los servicios a partir de las descripciones de los oferentes. Así podemos construir tres aproximaciones a la diferenciación horizontal: un índice de distancia promedio entre todos los competidores, otro a partir de la similaridad en la oferta de servicios, y finalmente identificar la pertenencia a cadenas hoteleras.

Los resultados de este estudio indican que la diferenciación horizontal, medida a través de la distancia geográfica, la distancia por atributos de los servicios ofrecidos o la pertenencia a una cadena hotelera, se traduce en mayores niveles de resiliencia—medida a partir de la evolución de las variables de ocupación y rentabilidad laboral—ante reducciones exógenas de la demanda. Los ajustes de precios jugaron un papel importante en la manera en que los establecimientos de alojamiento se adaptaron a las nuevas condiciones de demanda, pero la flexibilidad de las firmas para ejecutar estos cambios estuvo determinada por el grado de diferenciación.

Desde el punto de vista de política pública, estos resultados son relevantes ya que sugieren la posibilidad de crear instrumentos de política focalizados en función de las características de los hoteles y la composición del mercado, como programas locales que tengan en cuenta la vulnerabilidad causada por los niveles de diferenciación en la ciudad. Así mismo, será posible identificar anticipadamente riesgos, de modo que sea posible minimizar la fragilidad de ciertos mercados de alojamiento oportunamente. Adicionalmente, los índices de diferenciación que se proponen en este estudio podrán servir como variables para informar la focalización de programas, o como indicadores de alerta de vulnerabilidad.

El documento se divide en cinco secciones: en la primera, se presenta el marco teórico con los estudios recientes que se han realizado sobre esta temática; en la segunda, se expone el proceso de construcción del conjunto de datos; en la tercera, se plantea la estrategia de identificación del modelo; en la cuarta, se muestran los resultados, y, en la quinta, se presentan las conclusiones.

## **1.2. Literatura relacionada**

El sector hotelero presenta altos incentivos para la diferenciación (Audretsch & Feldman, 1996; Feldman & Audretsch, 1999; Jaffe, 2002). Según Marco-Lajara et. al (2016), estos incentivos ocurren porque se trata de un bien experiencia (un bien para

el cual solo se conoce la calidad a través del consumo del mismo), y porque existe tendencia a la aglomeración geográfica (debido a que los hoteles tienden a formar clústeres en zonas próximas a los sitios de interés).

En relación con el carácter de bien experiencia, la diferenciación juega un papel importante en la medida en que la demanda de habitaciones está influenciada por la reputación generada por visitantes anteriores (Liu 2017). Esta se ha medido a partir de atributos que reflejan superioridad en la calidad del bien (Gabszewicz & Thisse, 1986; Shaked & Sutton, 1987), expresado principalmente mediante el número de estrellas, el número de certificaciones y el número de valoraciones que han realizado sus visitantes. Nicolau et al. (2010) y Alonso-Almeida et al. (2012) han establecido que mejoras en estas variables aumentan la disposición a pagar de los clientes, lo que se puede traducir en mayor capacidad en la fijación de precios por parte de los hoteles.

Por su parte, la aglomeración se da porque los hoteles tienden a ubicarse alrededor de atractivos turísticos, y porque la demanda de alojamientos es complementaria a la demanda de otros servicios suplementarios que exhiben economías de escala locales. Sin embargo, existe un trade-off entre la intensidad de la competencia y los efectos de aglomeración (Tsang & Yip, 2009). A medida que los establecimientos de alojamiento puedan ubicarse lejos entre sí, podrán explotar los monopolios locales y evitar una mayor competencia, lo cual se vería reflejado en mayores ganancias. Baum & Mezas (1992) analizaron la competencia localizada en la industria hotelera de Manhattan, encontrando que los hoteles con más similitudes entre sí en términos de tamaño, precio de la habitación y ubicación física competían más intensamente. Por otro lado, Silva (2015) documenta que la proximidad entre las firmas genera presiones competitivas que pueden generar guerras de precio.

La teoría de la aglomeración sostiene que la presencia de competidores en una misma zona suele ser beneficiosa, debido a que la proximidad física genera economías de escala por aglomeración que pueden ser lo suficientemente grandes como para compensar los efectos negativos de la competencia (Marshall, 1920; Marco-Lajara et al. 2016; Durantón et al., 2020). Los beneficios en el sector hotelero pueden ser de dos tipos: i) mejoras en la producción, como consecuencia de encadenamientos productivos y economías de escala; y ii) aumentos en la demanda por la reducción en los costos de búsqueda de clientes y mayor frecuencia de visitas de potenciales compradores. Adicionalmente, se tienen ventajas por infraestructuras valiosas alrededor de los hoteles, incluyendo restaurantes, bares, atracciones, cines, entre otros, que atraen más clientes.

En este sentido, la proximidad física entre hoteles beneficia a los consumidores a través de costos de búsqueda más bajos, y aumentando así la cantidad de visitas y compras. Las ventajas de la aglomeración también han sido documentadas por Tsang & Yip (2009) para diferentes industrias que incluyen hoteles, teatros, restaurantes y tiendas especializadas

En esta misma línea, Marco-Lajara et al. (2016) encontraron que las firmas pequeñas en el sector hotelero pueden percibir en la competencia un entorno de aprendizaje. Esto se debe a que, si bien no tienen mucha capacidad para desarrollar actividades innovadoras, la experiencia de competidores exitosos les permite optimizar sus procesos. Este mismo mecanismo de aprendizaje se puede aprovechar afiliándose a franquicias o contratando mano de obra que tenga experiencias laborales previas con competidores con un nivel de eficiencia superior (Eisenhardt & Martin, 2000; Teece, Pisano, & Shuen, 1997; Wang & Ahmed, 2007). O'Neill, et al. (2006) identifican que una estrategia alternativa de diferenciación de los hoteles se da mediante las alianzas internacionales o la pertenencia a cadenas hoteleras que generalmente tienen un mayor goodwill.

La decisión sobre la ubicación geográfica está también asociada a las decisiones sobre las características de los servicios ofrecidos y la diferenciación horizontal. Baum & Haveman (1997) analizaron las decisiones multidimensionales en las elecciones de ubicación de los alojamientos en Manhattan desde dos perspectivas distintas: la elección del producto y la ubicación geográfica respecto a sus pares. Los autores encontraron evidencia que señala que los empresarios ubican nuevos hoteles lo suficientemente cerca de otros alojamientos ya establecidos, y que a pesar de ser similares en alguna dimensión, como el precio, buscan diferenciarse en otra dimensión del producto para evitar la competencia localizada y crear diferencias complementarias. De igual manera, Datta & Sudhir (2013) encuentran que las limitaciones a la diferenciación espacial conducen a una mayor diferenciación de productos. Kuksov (2004) argumenta que las empresas responden a la incapacidad de diferenciarse espacialmente en Internet con una mayor diferenciación de productos. Bar-Isaac, H et al, (2009) encuentran resultados similares.

En el caso de Airbnb, Voltes-Dorta & Inchausti-Sintes (2021) encontraron que características diferenciadoras como la cantidad de dormitorios y baños, así como su proximidad a las atracciones turísticas, son factores clave en la fijación de precios y afectan de forma considerable la rentabilidad. Sin embargo, para niveles más altos de diferenciación por calidad, la proximidad de los competidores tiene un impacto positivo en los precios, lo cual puede interpretarse como un efecto de la interacción de la

aglomeración y la diferenciación.

Los atributos que influyen en la decisión de alojamiento de los consumidores se han estudiado en trabajos como el de Wu (1999), Chu & Choi (2000) y Thrane (2005). Entre ellos se destacan: denominación de la marca, reputación, seguridad, características de las habitaciones, existencia de instalaciones deportivas, limpieza, condiciones de ubicación, entre otros. Según Dubé & Renaghan (2000) esta multidimensionalidad dificulta el análisis de la diferenciación horizontal, ya que es difícil aislar el impacto en los precios de cada atributo en particular.

La literatura que ha buscado profundizar en este asunto ha centrado su análisis en la identificación de atributos que permitan caracterizar segmentos agregados. Este es el caso de la diferenciación entre hoteles vacacionales o de negocios (Espinete et al. 2003; Haroutunian et al. 2005; así como Rigall-I-Torrent et al. 2011). Por ejemplo, cuando se planea una convención, los demandantes prefieren lugares que tengan salones de conferencias de gran tamaño, herramientas tecnológicas, con facilidad de acceso; mientras que las personas que viajan en familia prefieren hoteles con juegos para niños y lugares no muy concurridos.

La evidencia presentada hasta ahora muestra el papel central que juegan los procesos de aglomeración en la determinación de precios y beneficios de las firmas hoteleras, así como en el comportamiento de los consumidores. El contexto de la pandemia de COVID-19 ofrece una oportunidad de probar cómo estas características del mercado de alojamiento afectan su respuesta a perturbaciones negativas<sup>1</sup>.

Este trabajo contribuye también a la literatura reciente en el área de resiliencia de establecimientos. En particular, este trabajo encuentra que los hoteles que presentan una mayor diferenciación horizontal y vertical presentan mayor resiliencia a partir del análisis entre la tasa de ocupación, el precio y la rentabilidad. Estudios previos han demostrado que la competitividad, reputación y/o rentabilidad de las empresas hoteleras mejoran con la calidad del servicio (Graf, N. S. 2011; Chiang-Ming Chen, Yu-Chen Lin, Yan-Ping Chi, Sou-Chein Wu, Do, 2016), la diferenciación horizontal en la ubicación (Illescas-Manzano, M. D., Martínez-Puertas, S., & Sánchez-Pérez, M. 2022) o aumentos en la calidad de alimentos y bebidas (Chen y Lin, 2012). De igual forma, hoteles diferenciados a través de la pertenencia a una cadena hotelera cobran precios más altos (Silva, 2015). Los resultados de este trabajo complementan dicha

---

<sup>1</sup>Algunos artículos recientes han estudiado el impacto de la pandemia de COVID-19 sobre el sector hotelero. Ver Park, Kim & Kim (2020), Güliz & Akbıyık (2020) y Neuburger & Eggber (2021).

literatura.

### 1.3. Datos

El área de turismo es protagonista en innovaciones en las fuentes y métodos en las estadísticas oficiales, registros y otras fuentes alternativas, como los datos obtenidos con herramientas de web-scraping, que han permitido reducir la carga a los encuestados en los alojamientos turísticos (Adhinugroho et al. 2020) y que han servido como fuentes de información para estadísticas experimentales (Li et al., 2018). Las estadísticas experimentales de turismo se han servido de registros administrativos de empresas de alojamientos turísticos y otras fuentes alternativas como operadores de telefonía móvil o de sensores de ubicación; número de visitas a páginas web; transacciones con tarjetas de crédito; datos de rotación de productos en establecimientos de comercio; y publicaciones en redes sociales en Internet (EUROSTAT, 2017).

El web-scraping es una técnica de recolección cada vez más utilizada en la producción de estadísticas y en investigación. Esta técnica se sirve de un programa informático para recuperar información de páginas web y estructurarla en una base de datos. Con el fin de responder a las necesidades de información en una variedad de temas, el web-scraping se ha consolidado como un método complementario a las fuentes de recolección tradicionales como encuestas o censos que permite elaborar indicadores que previamente se consideraban por fuera del alcance de los institutos nacionales de estadísticas (INE).

Estas herramientas han permitido hacer análisis de caracterización bien sea a partir del análisis geoespacial o por la oferta de servicios en el sector de alojamiento en diversos países. Por ejemplo, en E.U. fue utilizada para la caracterización de la oferta de viviendas turísticas en plataformas digitales o en España, donde ha sido utilizado para medir la oferta de alojamientos en la economía colaborativa y su capacidad disponible. Recientemente Pramana et al. (2021) se basan en distintas fuentes de big data, algunas recopiladas con web-scraping, para estimar el impacto de la pandemia por COVID-19 en distintas regiones de Indonesia.

En general, estas aplicaciones integran y armonizan los datos obtenidos a través de web-scraping con otras operaciones estadísticas como encuestas y registros estadísticos. La perspectiva de EUROSTAT (2017) para las estadísticas de turismo en el largo plazo es que las nuevas fuentes de datos sean integradas y armonizadas en conjunto con las

encuestas en un ecosistema estadístico. EUROSTAT (2017) señala que las encuestas continuarán aportando información que no es posible conseguir de otra manera como es el motivo de viaje, los medios de transporte y las características sociodemográficas.

Las siguientes subsecciones proporcionan una descripción de los datos utilizados. Para la construcción de la base de datos se utilizaron tres fuentes de información. La primera fuente es un registro administrativo de turismo del Gobierno colombiano, denominado “Registro Nacional de Turismo” (RNT), el cual contiene la descripción de todos los hoteles y empresas turísticas autorizados en el país. Esta información se utilizó para estudiar la cobertura de las otras fuentes de datos, así como para determinar las distancias geográficas entre los establecimientos (utilizando la Application Programming Interface [API] de geolocalización de Google).

La segunda fuente de información es la Encuesta Mensual de Alojamiento del DANE, que monitorea las principales variables de resultados de los establecimientos hoteleros, como la ocupación y los ingresos para cada mes y región. De igual forma es posible observar atributos específicos de la demanda, como la proporción huéspedes no residentes, el motivo del viaje (ocio, negocios u otros), entre otros.

Finalmente, la tercera fuente se obtiene a través de un proceso de web scraping y minería de texto utilizando un rastreador web diseñado específicamente para este estudio<sup>2</sup>, el cual captura información de dos de los sitios web más utilizados en Colombia para la reservar alojamientos hoteleros. Estas fuentes de datos permiten la caracterización detallada del lado de la oferta hotelera<sup>3</sup>. El proceso de rastreo de esta información se realiza mediante la construcción de un algoritmo en lenguaje R que selecciona el nombre comercial de los establecimientos incluidos en la encuesta o el Registro de Turismo, con el propósito de obtener una descripción detallada de los servicios ofrecidos por los hoteles (por ejemplo, tener o no parques infantiles, centros de convenciones, zonas húmedas, etc.). También se recopiló información relacionada con los puntajes otorgados por los diferentes clientes que han visitado el hotel. Este trabajo contribuye a la literatura reciente en torno a las innovaciones metodológicas referentes a la recopilación

---

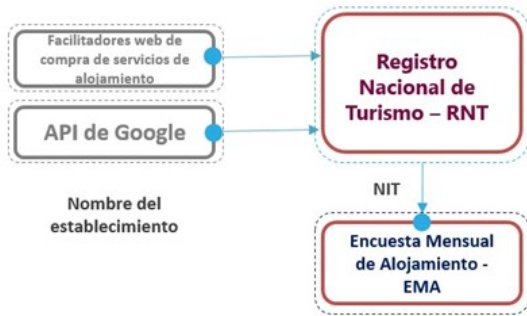
<sup>2</sup>El rastreador web está a disposición de cualquier investigador que quiera ponerse en contacto con el autor.

<sup>3</sup>Nelson (1970) sostiene que la publicidad es clave para la diferenciación, ya que indica cualidades no observables de un producto o servicio, lo que es especialmente importante para vender bienes de experiencia. Es así como los consumidores utilizan con frecuencia sistemas de clasificación de hoteles para elegir entre las alternativas de alojamiento. Es por esto que en este estudio se utilizaron las características del servicio publicitadas por los hoteles como insumo para medir la diferenciación horizontal.

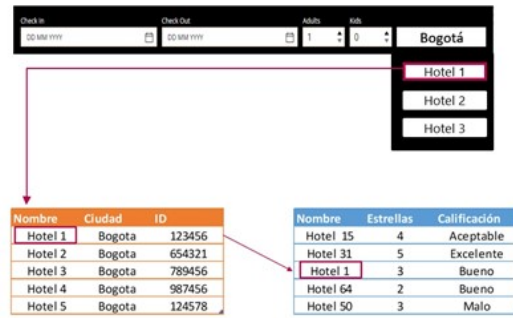
ción de fuentes alternativas de datos y la integración con las fuentes tradicionales, que tienen como objetivo generar información estadística oportuna y relevante que sirva de insumo para la política pública y las decisiones de las empresas. Estudios previos han abordado los impactos de la pandemia de COVID-19 a partir de análisis de creencias del riesgo de viajar de los turistas (Neuburger & Eggber, 2021). De manera complementaria, se han utilizado técnicas de minería de texto para analizar las emociones y preocupaciones de los empleados del sector de alojamiento (Park, 2020; Kim & Kim, 2020; Güliz & Akbıyık, 2020).

## Construcción de la base de datos

1. Se realizó una extracción de información por medio de web scraping sobre facilitadores web de compra de servicios de alojamiento, dónde se realizó la búsqueda con la concatenación del país, ciudad y nombre del hotel, para evitar obtener resultados erróneos al encontrar hoteles homónimos en otra región u otro país. Seguido a esto, se extrajeron atributos tales como estrellas, calificación realizada por usuarios, nombre presentado en la página, y servicios prestados por el hotel (por ejemplo, servicio de piscina, niñera, peluquería, entre otros). Al finalizar la extracción, se efectuó una revisión manual comparando el nombre buscado con el nombre resultado de la búsqueda, corroborando que la información fuera efectivamente la correspondiente al hotel de interés.
2. Para obtener los valores correspondientes a las coordenadas de ubicación de los establecimientos, se realizó una consulta usando la concatenación de país, ciudad y nombre comercial de los 823 hoteles registrados en el RNT para la ciudad Bogotá, en el servicio de API de codificación de geográfica de Google, la cual, devolvía la longitud y latitud del hotel.
3. Finalmente se integra el conjunto de datos del RNT con la Encuesta Mensual de Alojamiento (EMA) utilizando el Número de Identificación Tributaria (NIT) EMA para obtener información sobre el porcentaje de huéspedes no residentes, total huéspedes que pernoctaron y tarifa promedio en pesos habitación doble, entre otras. En esta integración fue posible vincular 126 de los 128 establecimientos de alojamiento incluidos en la muestra de la encuesta.



(a) Esquema de la conformación de la base de datos.



(b) Extracción de información por medio de web scraping sobre facilitadores web de compra de servicios de alojamiento.

Nombre	NIT
Hotel 1	658954-5
Hotel 2	987456-5
Hotel 3	326756-4
Hotel 4	135477-8
Hotel 5	916824-7

NIT	% Huespedes no residentes
123456-5	4
654321-1	5
789456-2	3
987456-5	2
124578-7	3

(c) Cruce entre la base y la EMA a partir del NIT.

Fuente: Elaboración propia, elaborado en base a EMA 2019, 2020, 2021-DANE.

A continuación, se presentan el tratamiento de las variables de diferenciación horizontal y las variables de resultado de los establecimientos hoteleros.

## Diferenciación Horizontal

### *Diferenciación por atributos*

Para cuantificar la diferenciación horizontal por atributos de los servicios hoteleros se construyó una medida de distancia cualitativa. Primero se generaron variables dicotómicas para las características de los servicios ofrecidos por los hoteles y otras particularidades de sus habitaciones. En la construcción de estas variables se tiene en cuenta la divulgación de información de cada uno de los establecimientos en las plataformas facilitadoras. En la siguiente tabla se muestran los atributos seleccionados.

Cuadro 1.1: Atributos de los servicios del hotel y de las habitaciones

Servicios del Hotel		Servicios de la habitación
Acceso a playa	Parqueadero	Aire acondicionado
Asador en exterior	Restaurante	Bañera
Bar	Sauna o turco	Escritorio
Botones	Servicio a la habitación	Habitación para discapacitados
Centro de negocios	Sombrillas	Nevera
Conserje	Tienda o Boutique	Televisión satelital
Gimnasio	trasporte al aeropuerto	
Pago sin efectivo	Viajeros con niños	
Peluquería	Wifi gratis	
Piscina		

Fuente: Elaboración propia.

Posteriormente es calculada la distancia de Gower entre cada par de hoteles<sup>4</sup>. La medida de diferenciación horizontal consolidada a nivel de hotel es la media aritmética entre las distancias obtenidas para cada uno de los pares a partir de la integración con la Encuesta Mensual de Alojamiento.

### ***Diferenciación por distancia geográfica***

El Registro Nacional de Turismo (RNT) tiene 14,422 observaciones únicas de las cuales el 97% son hoteles, hostales y apartahoteles distribuidos en todo el territorio colombiano. En total son 694 municipios en los cuales se distribuyen los establecimientos del RNT. En el caso de Bogotá, en el 2020 se cuenta 823 establecimientos lo que representa el 5.7% del total nacional.

<sup>4</sup>La distancia de Gower, usando ponderaciones homogéneas, se define como:

$$D(ij) = \frac{\sum_l d_l(i, j)}{L}$$

Donde  $D(ij)$  es la distancia entre los establecimientos  $i$  y  $j$ ;  $l$  es un índice de los atributos, y  $L$  es la cantidad total de atributos;  $d_l(i, j)$  es una variable indicadora que toma valor de 1 cuando los hoteles  $i$  y  $j$  son diferentes en el atributo  $l$ .



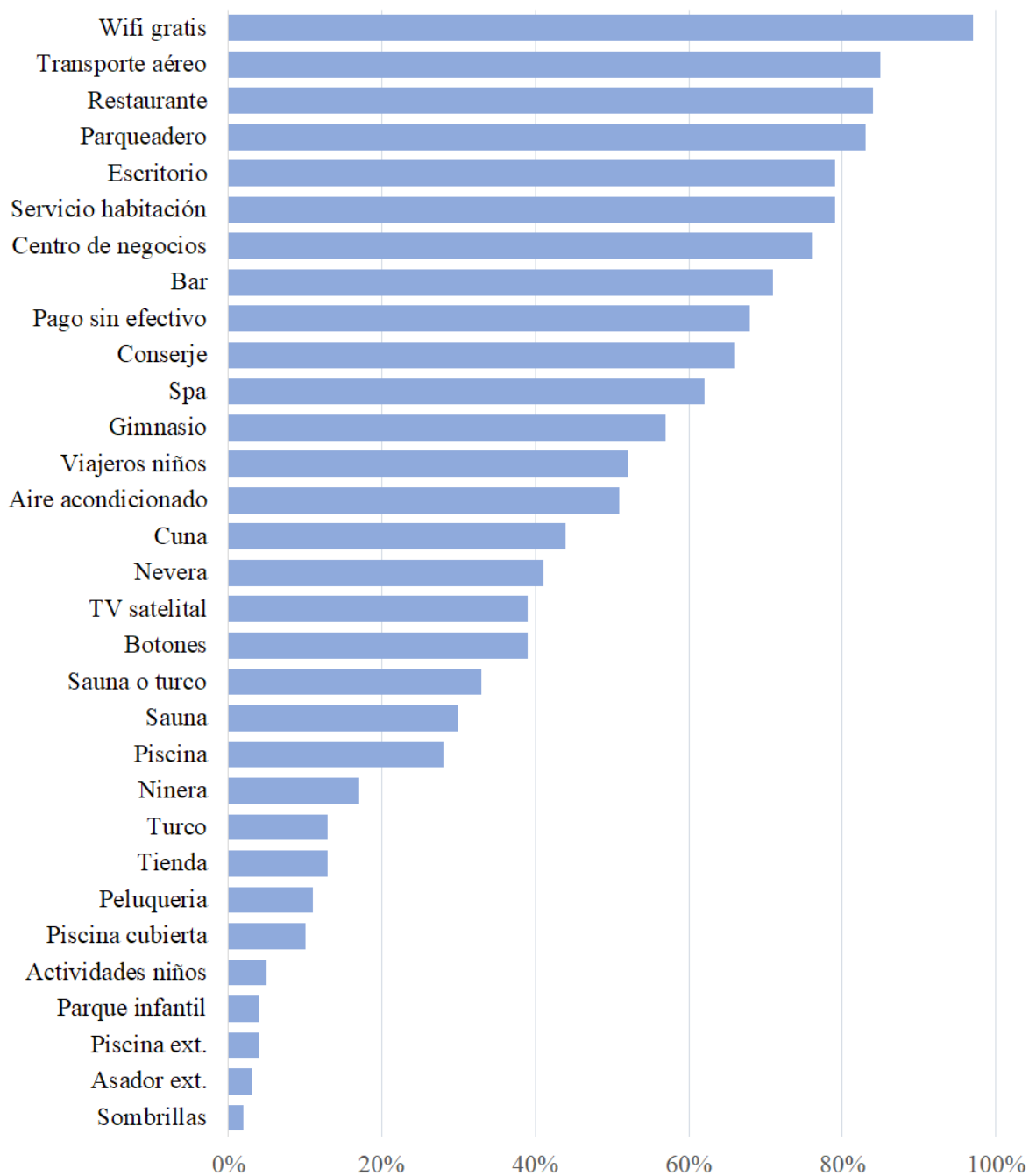
habitaciones ocupadas vendidas.

Cuadro 1.2: Estadísticas descriptivas variables de resultados, de exposición a la pandemia y de diferenciación.

Variable	Media	Des. Est.	Min	Max
Rentabilidad (miles COP)	5,842.44	7,418.03	0	122,987
Precio (Miles COP)	198.06	125.82	20.68	1302.74
Tasa de ocupación	0.39	0.27	0	1
Tasa de huéspedes no residentes	0.48	0.22	0	0.99
Tasa de huéspedes motivo negocios	0.56	0.26	0	1
Distancia por atributos	0.36	0.04	0.28	0.50
Distancia geográfica	8.42	1.11	6.90	11.05
Pertenece a una cadena	0.32	0.47	0	1

Fuente: Elaboración propia con información de la EMA (DANE) y el RNT.

Figura 1.3: Variables dicotómicas para la construcción de las distancias por atributos



Fuente: Elaboración propia.

## 1.4. Estrategia de identificación

Estudios anteriores han establecido la correlación entre diferenciación y desempeño de las firmas en el sector hotelero (Chen et al., 2011; Carlton & Dana, 2008; Lin & Chen, 2022). Sin embargo, estas aproximaciones no pueden identificar si un desempeño superior para un hotel dado es ocasionado por una perturbación de demanda más leve, o por un mayor nivel de resiliencia. Las restricciones sanitarias y los cambios en los comportamientos de los consumidores causados por la pandemia de COVID-19 representan un choque exógeno a las firmas que podemos usar para aproximarnos a un efecto causal. Adicionalmente, identificamos dos variables asociadas con el nivel de exposición de los establecimientos a estos choques: la proporción de consumidores extranjeros y la proporción de consumidores que reportan negocios como motivo de su viaje. En esta sección vamos a explicar cómo usamos estas medidas de exposición, junto a la temporalidad de la pandemia, para identificar efectos causales. El estudio de estos efectos de forma heterogénea por niveles de diferenciación nos va a permitir establecer la relación entre diferenciación horizontal y resiliencia.

La estrategia de identificación explota la exposición diferenciada que tuvieron los hoteles a la reducción exógena de la demanda ocasionada por la pandemia del coronavirus COVID-19. Durante el 2020, restricciones como el cierre de aeropuertos, aislamiento obligatorio, cierre de fronteras<sup>5</sup>, medidas de bioseguridad, entre otros, ocasionó que existieran limitaciones adicionales para los viajeros no residentes en el país. En el caso español, por ejemplo, las pernoctaciones de residentes disminuyeron en 10.9 %, mientras que para los no residentes disminuyó en 77.5 % (Hosteltur, 2020). Para el caso de la ciudad de Bogotá observamos que la caída inicial en el flujo de huéspedes para el mes de abril fue similar entre residentes y no residentes, pero a partir de este momento la evolución de ambos segmentos es distinta; la reactivación para el segmento de viajeros residentes ha sido más rápida (ver gráfica 1.4).

Previo a la pandemia, la proporción de viajeros no residentes en un hotel era estable<sup>6</sup>. Los hospedajes que recibían recurrentemente una proporción mayor de extranjeros estaban, por lo tanto, expuestos en mayor grado a los efectos de una disminución súbita de la demanda que estuviera sesgada hacia los viajeros provenientes del extranjero. De

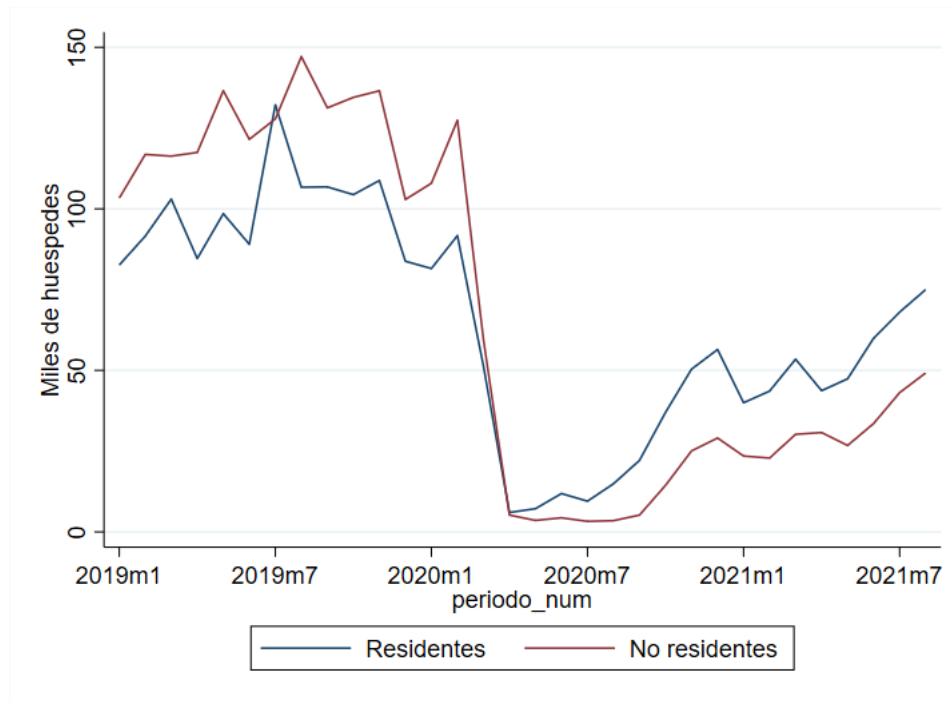
---

<sup>5</sup>En Colombia desde el 23 de marzo de 2020 se cerraron las fronteras aéreas, y solo hasta el primero de septiembre se reactivó la actividad turística en el país, fecha en la que también se reabre el aeropuerto el Dorado de la ciudad de Bogotá.

<sup>6</sup>El promedio de la varianza entre hoteles en la proporción de viajeros no residentes es más de tres veces el promedio de la varianza entre meses para un mismo hotel.

esta forma, la existencia de una variación temporal, dada por el antes y el después de la pandemia, y un canal de exposición diferenciado por la naturaleza de la demanda, bien sea turismo interno, o visitantes no residentes, sugiere el uso de un estimador de diferencias en diferencias<sup>7</sup>.

Figura 1.4: Número total de huéspedes por mes: residentes y no residentes



Fuente: Elaboración propia, elaborado en base a EMA 2019, 2020, 2021-DANE.

La estimación del efecto del choque de demanda se va a realizar de forma separada para los establecimientos con alta o baja diferenciación relativa. La diferenciación se mide de dos formas, usando la distancia de atributos de los servicios o la distancia geográfica – para cada definición se generan los respectivos grupos de alta y baja diferenciación, y se presentan las estimaciones correspondientes. Estos grupos se definen a partir de los establecimientos que tienen una distancia promedio a sus pares por encima y por debajo, respectivamente, de la mediana de la ciudad<sup>8</sup>. En la sección 5.2

<sup>7</sup>Uno de los desafíos que plantea el uso del impacto del COVID-19 como experimento natural es que no solo existió un choque de demanda, sino que este estuvo acompañado por choques de oferta. Esto no es un problema para la metodología propuesta, en la medida en que estos hayan sido homogéneos entre el grupo de tratamiento y control

<sup>8</sup>En la sección de robustez se presenta una estimación alternativa en la que los grupos de dife-

también se presenta una estimación suplementaria en la cual el factor de diferenciación estudiado es la pertenencia a una cadena hotelera.

De esta forma, el siguiente modelo se estima de forma independiente para los hoteles de alta y baja diferenciación:

$$y_{it} = \alpha + \sum_{\tau} \theta_{\tau} E_i T_{\tau} + \sigma_t + \phi_i + \varepsilon_{it} \quad (1.1)$$

Donde  $y_{it}$  representa la rentabilidad, el precio o la tasa de ocupación hotelera;  $E_i$  es el porcentaje de huéspedes del hotel en 2019 que no eran residentes colombianos;  $T_{\tau}$  es un conjunto de variable indicadoras por trimestres;  $\sigma_t$  son efectos fijos de mes;  $\phi_i$  son efectos fijos de establecimiento. Los parámetros de interés son  $\theta_{\tau}$ , donde  $\tau$  es un contador de trimestres<sup>9</sup>.

La estrategia de identificación requiere el supuesto que plantea que las variables de rentabilidad, precio, y tasa de ocupación hotelera tengan tendencias paralelas entre residentes y no residentes. En la gráfica 1.5, en el panel a. se puede observar cómo antes del COVID-19 la rentabilidad tenía una tendencia paralela entre los establecimientos más enfocados hacia los huéspedes no residentes y los más enfocados hacia los huéspedes locales. Así mismo este comportamiento de tendencias paralelas entre estos dos grupos se puede apreciar al observar los precios y la tasa de ocupación hotelera. Las estimaciones presentadas en la sección de resultados proveen también evidencia de tendencias paralelas pre-tratamiento mediante pruebas de significancia de los coeficientes efectos placebo antes del tratamiento.

La gráfica 1.5 muestra que antes del COVID-19 los hoteles con una mayor proporción de no residentes presentaban mayores niveles de rentabilidad (panel a) y que ésta quedo en niveles inferiores durante algunos meses en el periodo posterior a la declaración de la emergencia sanitaria. En el caso de los precios, en el panel b, se puede observar que el promedio de esta variable es superior en los hoteles con una mayor diferenciación de definen a partir de quintiles (primero y segundo en un grupo y cuarto y quinto en el otro).

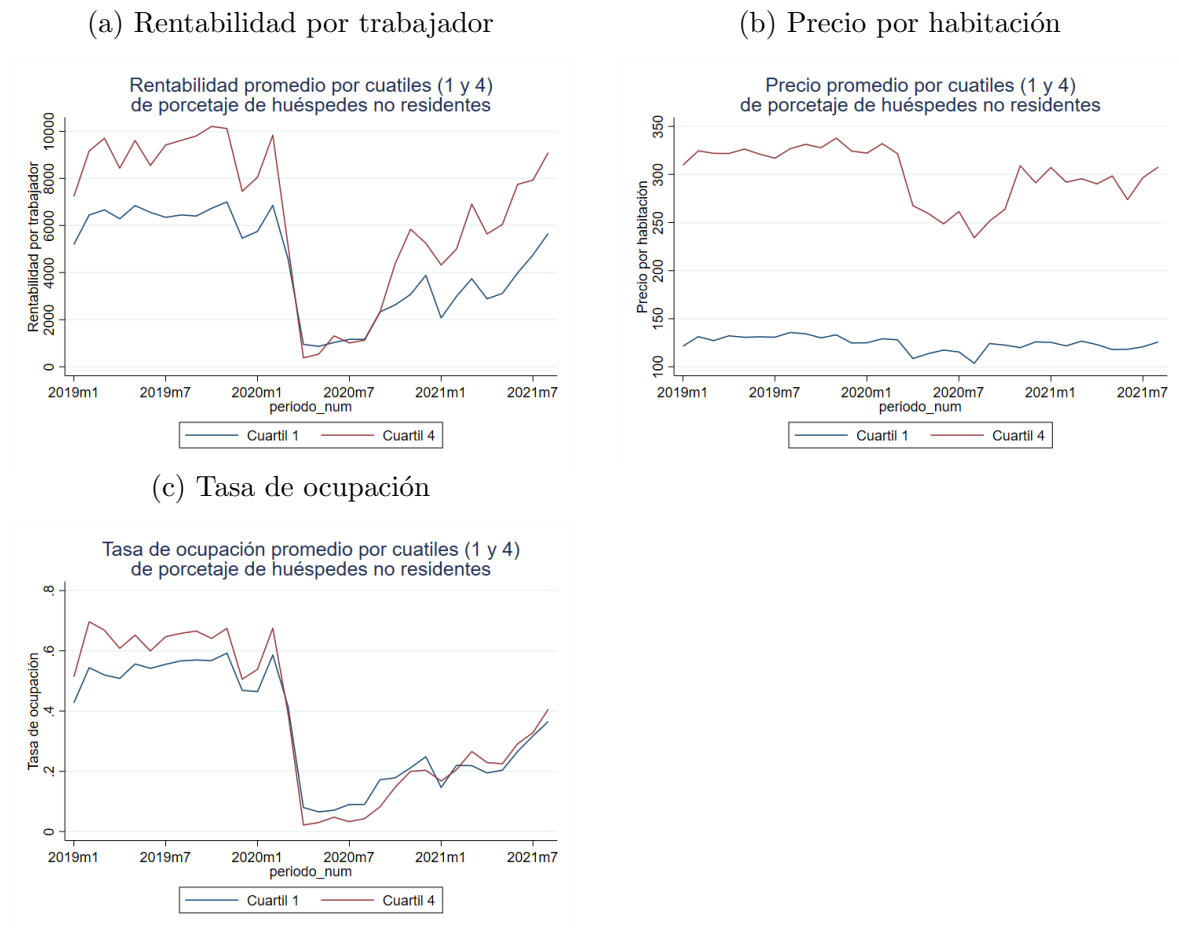
<sup>9</sup>En la especificación favorita se usan efectos fijos a nivel de establecimiento y no se incluye un vector de controles. También se probó la especificación sin efectos fijos y con controles a nivel de establecimiento de la siguiente forma:

$$y_{it} = \alpha + \sum_{\tau} \theta_{\tau} E_i T_{\tau} + \sigma_t + \varphi D_i + \beta X_i + \varepsilon_{it}$$

Donde  $X_i t$  es un vector de controles que incluye las características de las habitaciones y hoteles dentro de la muestra. Los resultados usando esta especificación alternativa son similares a los presentados.

recepción de huéspedes no residentes, y durante unos meses después del inicio de las restricciones hubo un ajuste muy significativo, que había desaparecido para el final del periodo de estudio. Finalmente, en el panel c., se observa que los establecimientos con mayor ocupación de no-residentes tenían tasas de ocupación consistentemente más alta antes del inicio de las restricciones, pero después de estas quedaron aproximadamente al mismo nivel durante todo el resto del periodo de estudio.

Figura 1.5: Rentabilidad, precio, tasa de ocupación promedio hoteles según residentes no residentes



Fuente: Elaboración propia, elaborado en base a EMA 2019, 2020, 2021-DANE.

## 1.5. Resultados

En esta sección se muestran los resultados de la estimación del modelo. En primer lugar, se presentan los resultados de la estimación del modelo de diferencia en diferencias con la muestra completa, sin tener en cuenta si los establecimientos están alta o bajamente diferenciados. La variable de exposición usada para estas estimaciones es el porcentaje de ocupación de viajeros no residentes en 2019<sup>10</sup>. Así, se estudió el impacto de la perturbación de la demanda en las variables de rentabilidad, tasa de ocupación y precios en los hoteles de Bogotá D.C. La gráfica 1.6 presenta los resultados de la estimación con temporalidad trimestral e intervalos de confianza del 95 %.

Consistente con el supuesto de tendencias paralelas, la estimación de efectos puntuales antes de la entrada en vigor de las restricciones sanitarias es no significativa y cercana a 0 para casi todos los trimestres, y es estadísticamente igual a 0 conjuntamente para el periodo pre-tratamiento (desde el primer trimestre de 2019 hasta el primer trimestre de 2020). Se encuentra un efecto significativo en las tres variables de resultado: después de las medidas de confinamiento, hubo una caída en la tasa de ocupación, rentabilidad y precios, y una recuperación posterior que ha tomado varios trimestres. En el caso de los precios se observa una resistencia a recuperar los niveles anteriores al trimestre de tratamiento.

Los resultados anteriores agregan dos poblaciones que se comportan diferente: los hoteles que hacen esfuerzos por diferenciarse horizontalmente de otros y los que no. Las estimaciones de efectos heterogéneos confirman la hipótesis de que el impacto de las restricciones de demanda fue diferenciado para estas dos poblaciones.

La gráfica 1.7 presenta los resultados de estas estimaciones usando dos especificaciones: midiendo la diferenciación por atributos o por distancias geográficas. Bajo ambas especificaciones se encuentran resultados cualitativamente similares. Para los establecimientos menos diferenciados, la rentabilidad cae significativamente en el trimestre de inicio de las restricciones. A partir de entonces empieza un proceso de recuperación que se extiende hasta el final del periodo de estudio en el tercer trimestre de 2021. Para los establecimientos más diferenciados el efecto post-pandemia sobre la rentabilidad fue no significativo<sup>11</sup>.

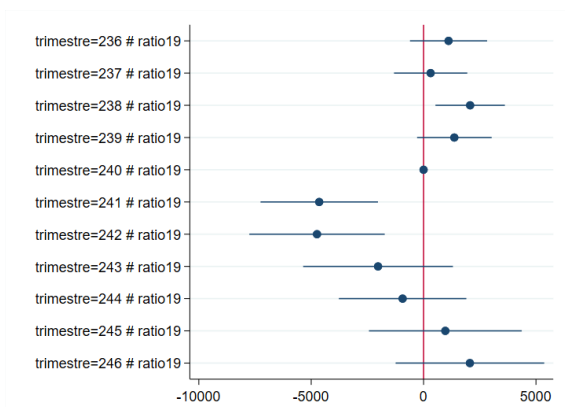
---

<sup>10</sup>La especificación principal mide la exposición de acuerdo al porcentaje de viajeros no residentes, sin embargo, también se evalúa esta exposición según el motivo de viaje. Los resultados de esta estimación son incluidos en el ANEXO 1.

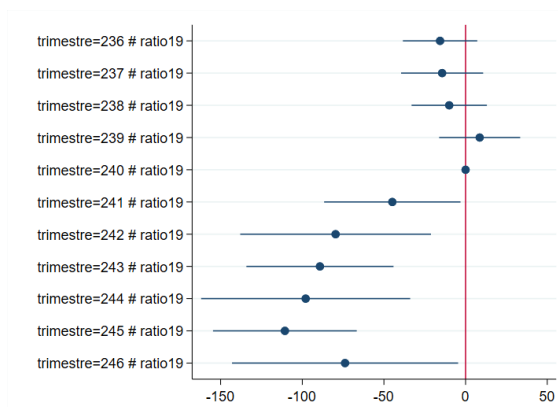
<sup>11</sup>Excepto por los trimestres 2 y 3 de 2021 en las estimaciones que usan diferenciación por atribu-

Figura 1.6: Modelo sin efectos heterogéneos por diferenciación.

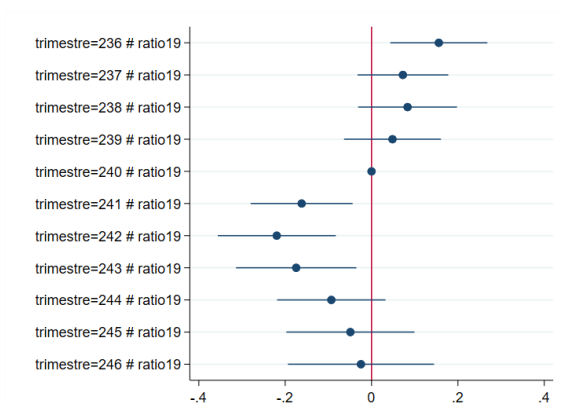
(a) Rentabilidad



(b) Precio



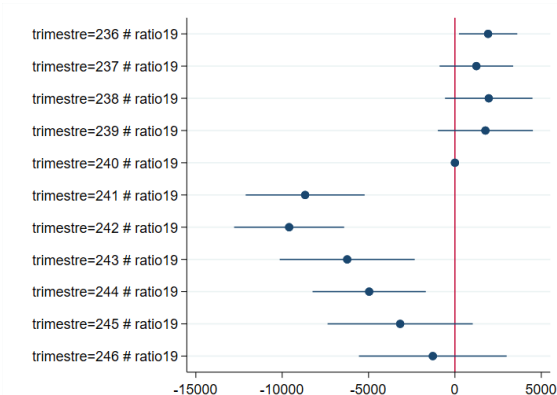
(c) Tasa de ocupación



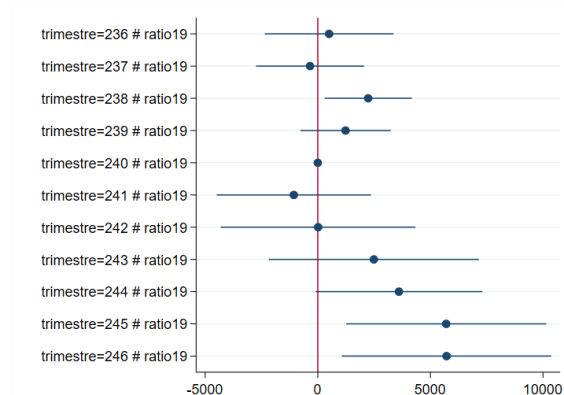
tos, donde encontramos un efecto positivo inesperado. Es importante agregar que en este momento muchas de las restricciones sanitarias ya habían sido levantadas, y el sector ya había experimentado varios meses de recuperación, de forma que estos resultados puede sugerir la existencia de nuevas oportunidades durante la recuperación que los establecimientos diferenciados pudieron explotar.

Figura 1.7: Resultados para rentabilidad por trabajador.

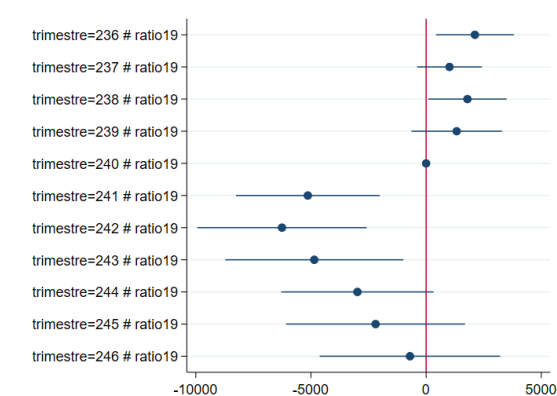
(a) No diferenciados, proximidad por atributos



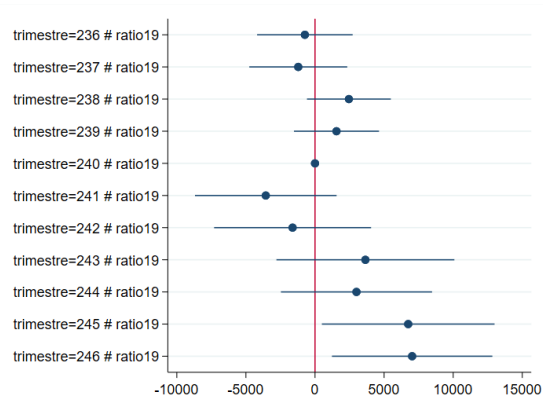
(b) Diferenciados, proximidad por atributos



(c) No diferenciados, distancia geográfica



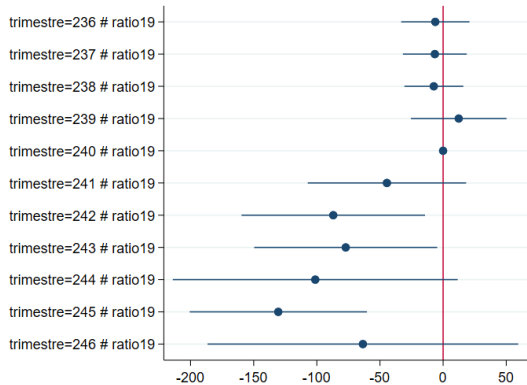
(d) Diferenciados, distancia geográfica



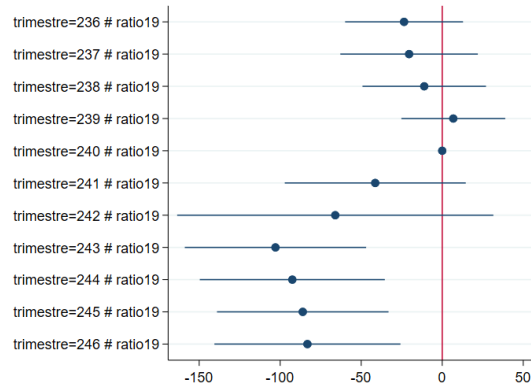
En el caso de la variable precio, en el gráfico 1.15 se presenta el efecto diferenciado en los dos grupos de estudio. Tanto los hoteles diferenciados como los no diferenciados tuvieron que ajustar sus tarifas desde el segundo trimestre de 2020. La dinámica de los precios es bastante similar cuando se analizan los grupos definidos según los atributos del servicio. Al analizar la heterogeneidad por diferenciación geográfica se encuentran efectos más persistentes para el grupo de establecimientos menos diferenciados.

Figura 1.8: Resultados para precio por habitación.

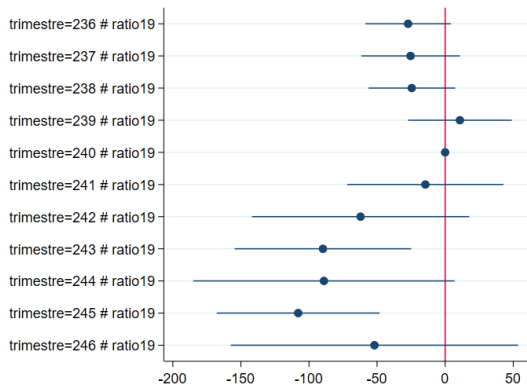
(a) No diferenciados, proximidad por atributos



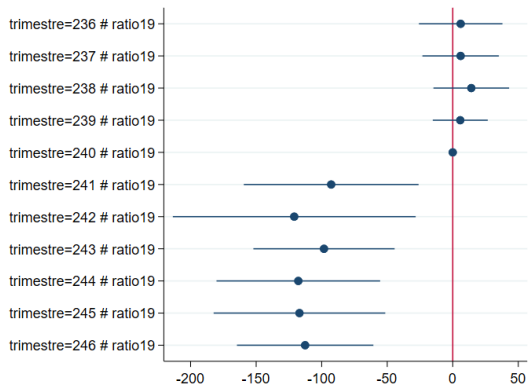
(b) Diferenciados, proximidad por atributos



(c) No diferenciados, distancia geográfica



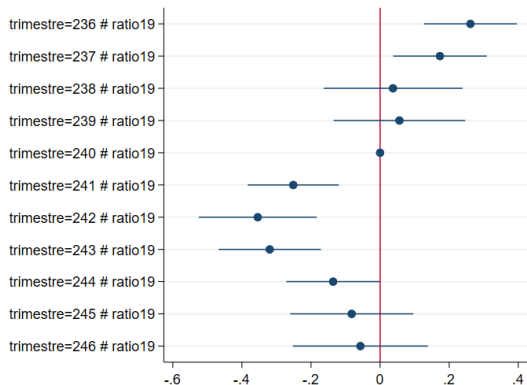
(d) Diferenciados, distancia geográfica



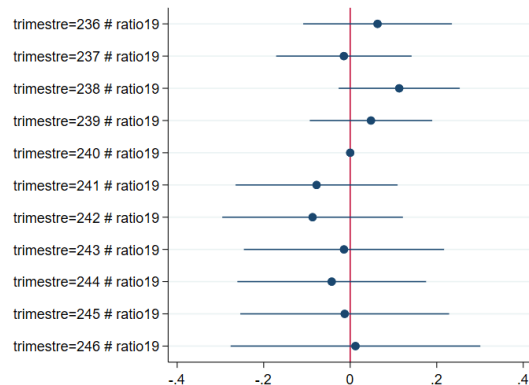
Finalmente, al evaluar estos resultados en la tasa de ocupación, los resultados de la diferenciación horizontal, bien sea por atributos o por distancia espacial, muestra que en los hoteles no diferenciados hay variaciones negativas y significativas en la tasa de ocupación después de la variación exógena de la demanda ocasionada por 19las medidas impuestas para contener la pandemia, una recuperación que dura todo el resto del periodo de estudio. Por su parte, para los hoteles diferenciados se encontraron efectos no significativos, independientemente de la definición de diferenciación utilizada.

Figura 1.9: Resultados para tasa de ocupación.

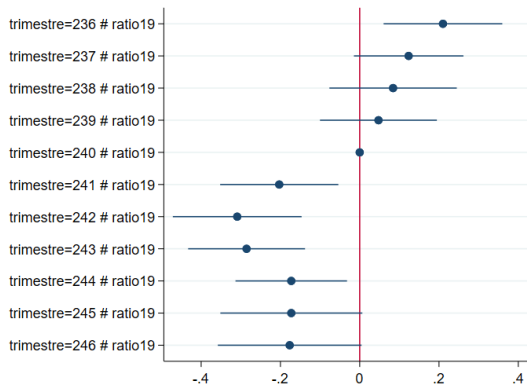
(a) No diferenciados, proximidad por atributos



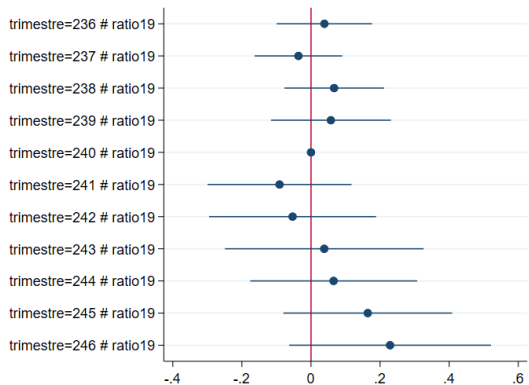
(b) Diferenciados, proximidad por atributos



(c) No diferenciados, distancia geográfica



(d) Diferenciados, distancia geográfica



### 1.5.1. Cadenas de hoteles

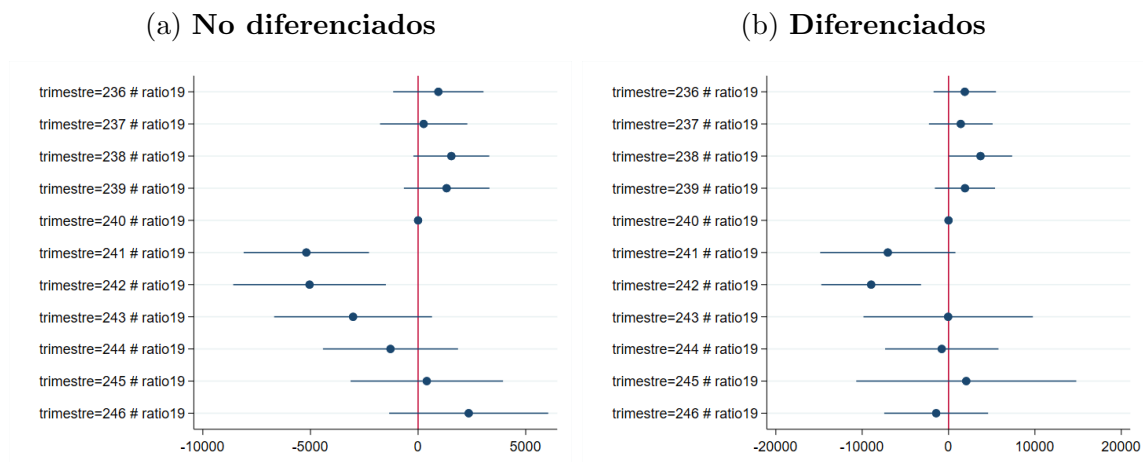
Los hoteles buscan diferenciarse mediante otros atributos, como las alianzas internacionales o la pertenencia a cadenas hoteleras que generalmente tienen un mayor goodwill (O'Neill, et al., 2006), debido a que los consumidores pueden evaluar mejor un servicio cuando este pertenece a una extensión de una marca con la cual tienen lealtad.

Las ventajas competitivas en sinergias financieras, comerciales, gerenciales, tecnológicas y productivas, y las economías de escala asociadas a la integración en cadenas son bases esenciales para la competitividad en el sector turístico debido a su natura-

leza altamente dinámica y su exposición a los movimientos globalizados (Such et. al, 2013). En esta línea, Camisón et. al, (2020) muestran que las empresas pertenecientes a cadenas tienen un desempeño económico superior al de las empresas hoteleras independientes. Este mejor desempeño se mantiene periodos de crisis y auge de acuerdo con O'Neill, J. & Carlbäck (2011). Los autores encontraron explotaron episodios de recesiones para describir (mediante pruebas de diferencias de medias) diferencias entre el rendimiento de los hoteles que pertenecen a una cadena y los que no. Sus resultados sugieren que las tasas de ocupación de los hoteles que pertenecen a cadenas son menos vulnerables durante épocas de recesión.

En la figura XX se presenta el resultado de utilizar la metodología de la sección anterior, usando la pertenencia a una cadena hotelera para encontrar efectos heterogéneos<sup>12</sup>. Los resultados muestran una mayor resiliencia de los hoteles que pertenecen a empresas con más de un establecimiento al evaluarlo en las variables de ocupación y rentabilidad.

Figura 1.10: Resultados para rentabilidad por trabajador, usando pertenencia a cadenas.



<sup>12</sup>La definición corresponde a las empresas en el que la suma de establecimientos por Número de Identificación Tributaria – NIT, es superior a uno.

Figura 1.11: Resultados para precio por habitación, usando pertenencia a cadenas..

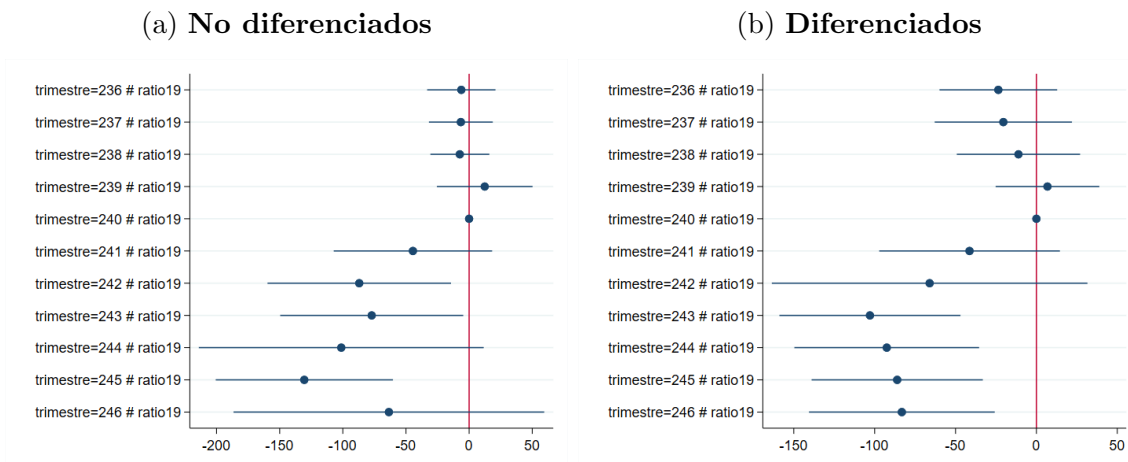
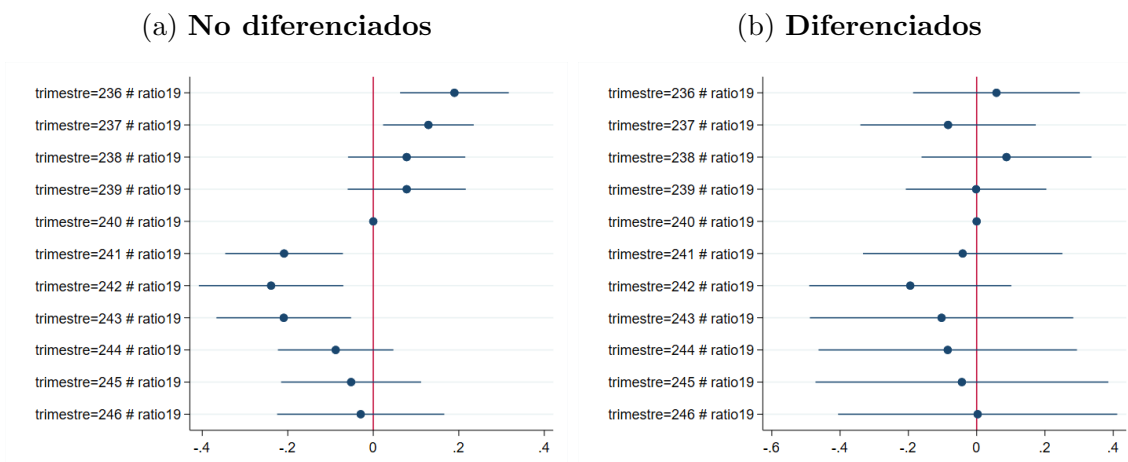


Figura 1.12: Resultados para tasa de ocupación, usando pertenencia a cadenas.



## 1.6. Robustez

Para valorar la sensibilidad de los resultados a cambios la especificación del modelo se implementaron dos conjuntos de ejercicios de robustez. Primero se analizó si las conclusiones principales son sensibles a un cambio en la variable que mide la exposición a las perturbaciones de demanda generadas por la pandemia. En este ejercicio de robustez se identifica otra variable que afecta la exposición del hotel a las restricciones de demanda – la proporción de huéspedes en el hotel en 2019 que viajan con motivos de negocios. Los resultados encontrados son cualitativamente iguales a los generados con

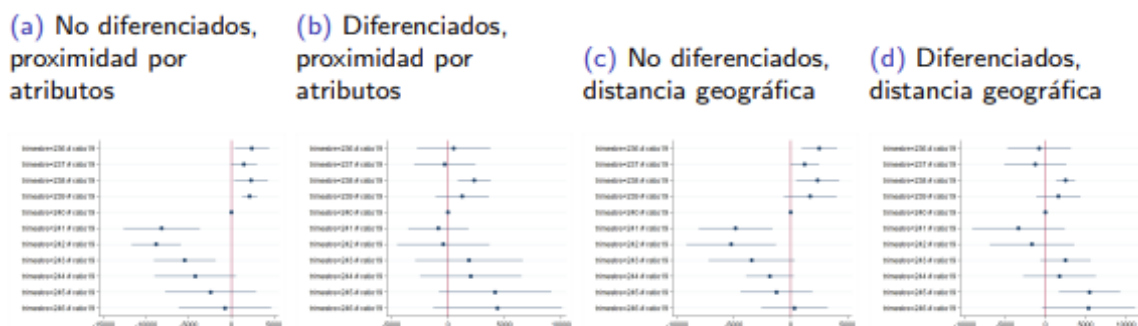
nuestra especificación base. En segundo lugar, se presentan estimaciones en las que los grupos de diferenciación no se definen a partir de la mediana sino a partir de quintiles en la distribución de distancias.

### **1.6.1. Especificación con variables geográficas y controles por diferenciación vertical**

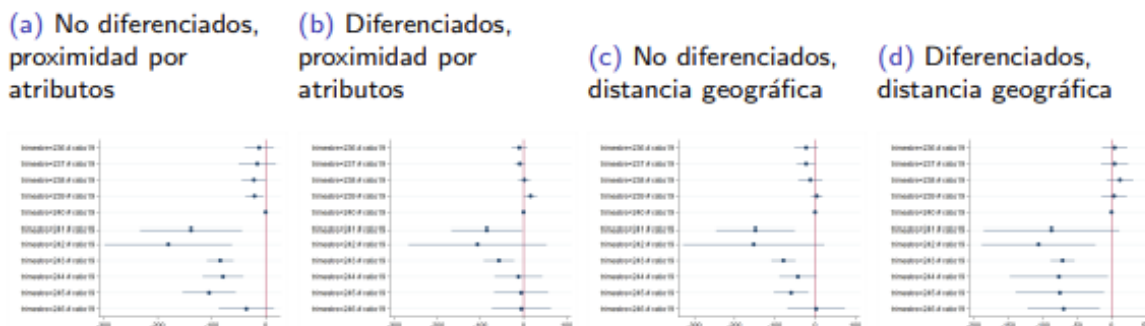
Adicionalmente, se especifico un modelo en el que no se tenían efectos fijos de establecimientos, sino se incorporaron los atributos de los establecimientos de los hoteles, y se adicionaron efectos fijos de localidad en Bogotá y mediante la identificación de clústeres de hoteles, en los que se observaron conglomerados en la zona céntrica, zona financiera y el aeropuerto, se construyeron distancias a de cada establecimiento a los mismos. De igual forma y con el propósito de aislar el posible efecto de diferenciación vertical que puede tener la ubicación geográfica se adicionan variables como el número de estrellas, y los puntajes ponderados por comentarios como una aproximación de diferenciación vertical. Los resultados principales del artículo se mantienen.

Figura 1.13: Resultados para rentabilidad por trabajador, precio por habitación y tasa de ocupación, usando especificación alternativa con efectos fijos geográficos y controles de diferenciación vertical (estrellas).

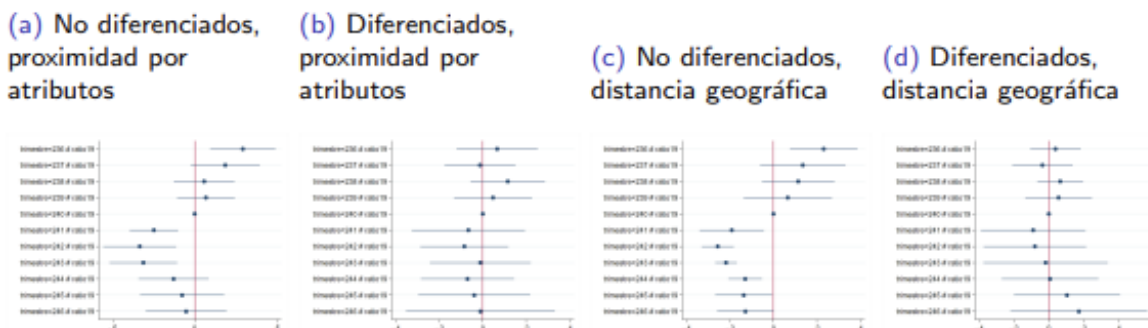
(a) Rentabilidad por trabajador



(b) Precio por habitación



(c) Tasa de ocupación



1.6.2. Medida alternativa de intensidad de tratamiento

La dependencia de la demanda de viajeros de negocios fue considerado un factor de exposición al COVID, por lo cual esperaba que su recuperación fuera desigual, debido

a que las empresas estaban utilizando la tecnología como sustituto de los viajes no esenciales (McKinsey, 2020). En el caso colombiano, las cadenas hoteleras han afirmado que el segmento de negocios es el más rezagado (Portafolio, 2021).

Esta sección presenta los resultados que surgen de una modificación a la estrategia de identificación, explotando que la exposición de los hoteles a la reducción exógena de la demanda ocasionada por las medidas adoptadas para contener la pandemia fue heterogénea de acuerdo al motivo de viaje – siendo los hoteles con mayor ocupación pre-pandemia de viajeros de negocios los más expuestos—. Es importante señalar que la correlación entre la variable de exposición de la especificación original (porcentaje pre-pandemia de usuarios no residentes) y la variable para esta nueva especificación es negativa y significativa (-0,202). Esto demuestra que no se está midiendo la misma exposición a través de dos variables diferentes, sino que se está capturando dimensiones de exposición diferentes. A pesar de esto, se encuentra que los resultados son completamente robustos a este cambio metodológico.

Los efectos en rentabilidad y ocupación son cualitativamente idénticos a los que se presentaron previamente: el grupo de hoteles no diferenciados enfrentó una caída en la rentabilidad y la ocupación, seguido de un proceso lento de recuperación. Estos efectos no fueron significativos en el caso de los hoteles diferenciados. Los resultados de precios son también cualitativamente similares a los expuestos anteriormente, pero ahora muestran una persistencia mayor del choque de precio entre los establecimientos de bajo nivel de diferenciación.

Figura 1.14: Resultados para rentabilidad por trabajador, usando el porcentaje de huéspedes que viajan por negocios como medida de exposición, diferenciación medida por proximidad en atributos.

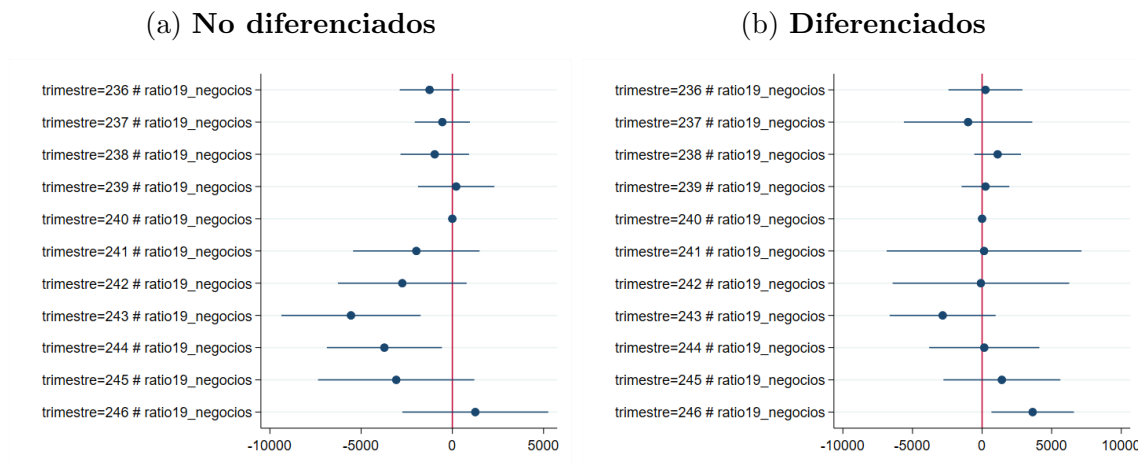


Figura 1.15: Resultados para precio por habitación, usando el porcentaje de huéspedes que viajan por negocios como medida de exposición, diferenciación medida por proximidad en atributos.

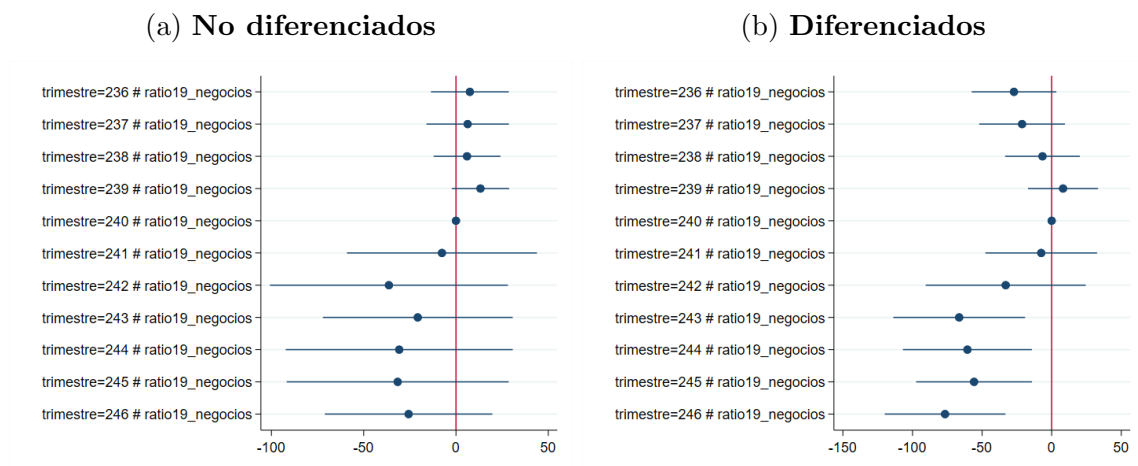
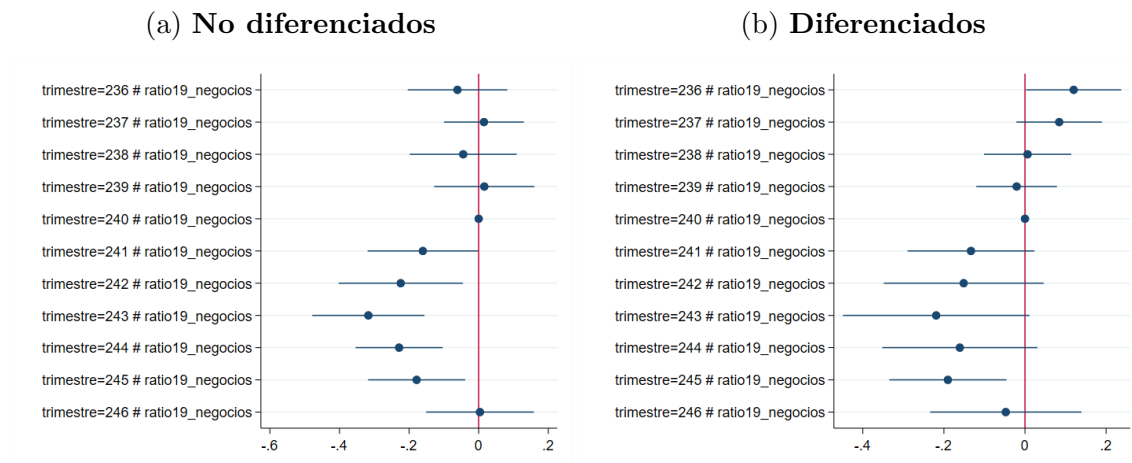


Figura 1.16: Resultados para tasa de ocupación, usando el porcentaje de huéspedes que viajan por negocios como medida de exposición, diferenciación medida por proximidad en atributos..



### 1.6.3. Medida de diferenciación por atributos

A diferencia de la estimación presentada en el capítulo de resultados en el que la variable de diferenciación por atributos estaba construida con la mediana de los índices de diferenciación, en este ejercicio se excluye el quintil de la mitad, y los 2 superiores y los dos inferiores representan respectivamente los hoteles diferenciados y los que no. Los resultados, ilustrados en las figuras 1.17 a 1.19, son cualitativamente iguales a los presentados en el capítulo de resultados.

Figura 1.17: Resultados para rentabilidad por trabajador, diferenciación medida por proximidad en atributos por quintiles (1 y 2 vs. 3 y 4).

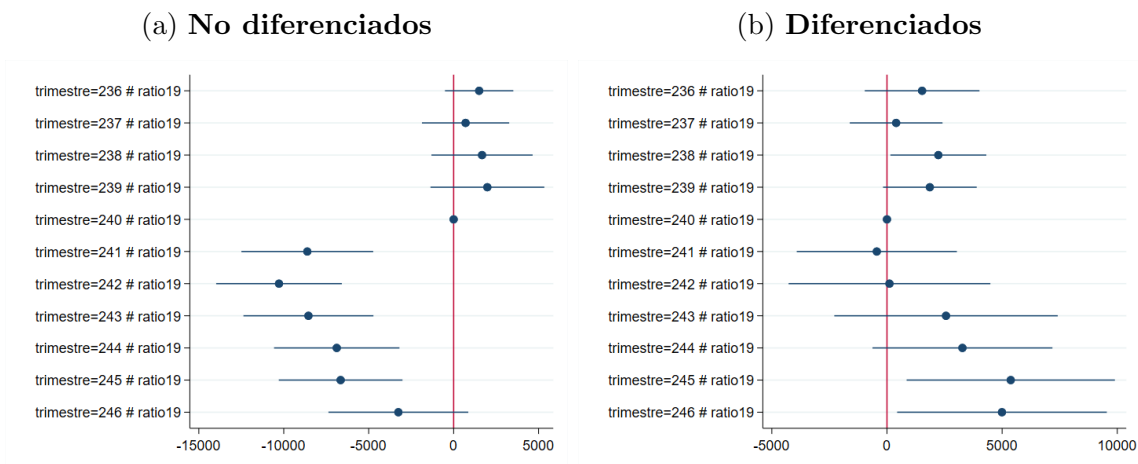


Figura 1.18: Resultados para precio por habitación, diferenciación medida por proximidad en atributos por quintiles (1 y 2 vs. 3 y 4).

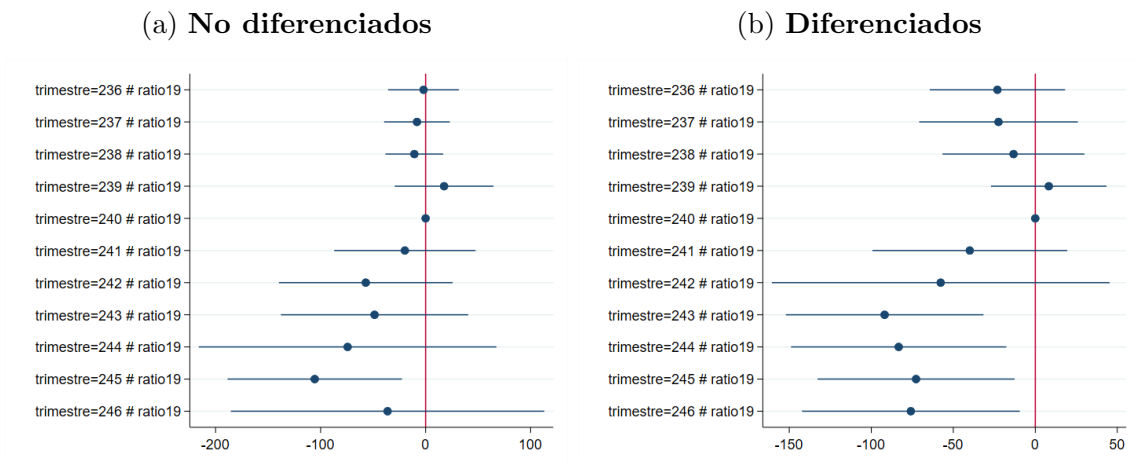
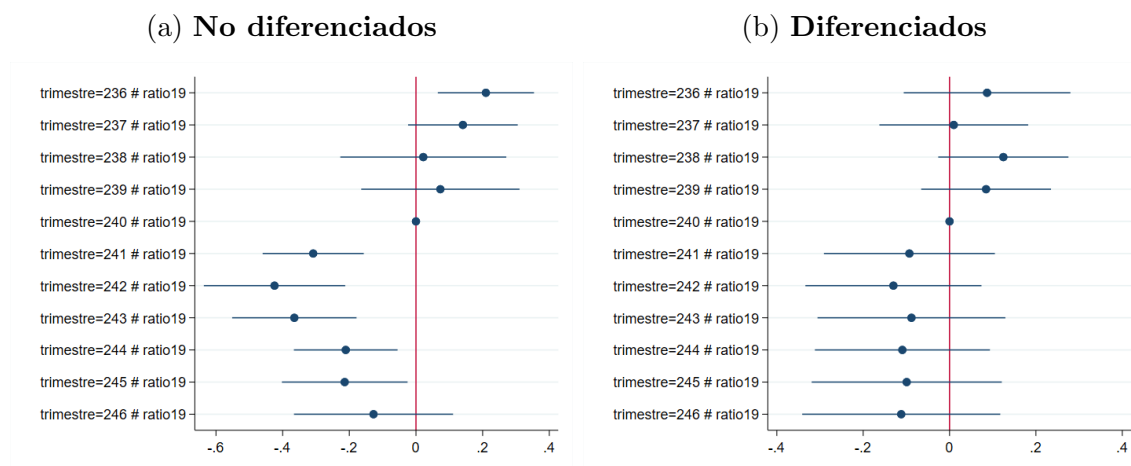


Figura 1.19: Resultados para tasa de ocupación, diferenciación medida por proximidad en atributos por quintiles (1 y 2 vs. 3 y 4).



## 1.7. Conclusiones

Los establecimientos de alojamiento desempeñan un papel fundamental en la oferta turística y por esto es pertinente estudiar sus niveles de resiliencia ante los choques externos. Por este motivo, la metodología utilizada y los resultados encontrados son relevantes, tanto desde la óptica de política pública, como de decisiones empresariales.

Los resultados de este estudio indican que, ante reducciones exógenas de la demanda, como las generadas durante la pandemia por COVID-19, los esfuerzos de diferenciación por atributos o distancia, o la afiliación a una cadena hotelera, se reflejan en mayores niveles de resiliencia, medidos a través de la evolución de las variables de ocupación y rentabilidad laboral. En este proceso de adaptación a nuevas condiciones de demanda, la reducción transitoria de los precios jugó un papel importante. Sin embargo, la flexibilidad de los establecimientos para ejecutar estos cambios estuvo asociada con los niveles de diferenciación y su vinculación con estructuras como las cadenas y marcas.

Una agenda de investigación futura deberá incluir la exploración de mecanismos adicionales que expliquen la mayor resiliencia por parte de los establecimientos más diferenciados. Para este trabajo se podría estudiar la estructura organizativa de las cadenas de hoteles, para analizar los procesos de aprendizaje o de aprovechamiento de economías de escala se presentan en este tipo de conglomerados. La relación entre las

variables de resultados y la diferenciación geográfica sugiere la necesidad de estudiar cómo la ubicación espacial de los hoteles afecta su desempeño, como será abordado en el capítulo 3 del presente documento.

De igual forma, futuros desarrollos de esta misma línea de investigación deberán extender el estudio a otros mercados: a nivel geográfico (otras ciudades) o de sector económico (otras clasificaciones industriales). Esto a su vez permitirá explorar la potencial heterogeneidad entre distintos mercados de los efectos encontrados, así como modelar las relaciones que existen entre estos.

La medición de la diferenciación por atributos es otro punto de la agenda de investigación que demandará nuevos desarrollos en los próximos años. Dado que las fuentes tradicionales de datos en general no tienen el tipo de información necesaria para crear los indicadores requeridos, este estudio introdujo innovaciones en la forma de hacer esta medición para el caso de la industria hotelera, utilizando herramientas de Big Data (web scraping y data mining). Una limitación de la aproximación utilizada es que la descripción de los atributos de los servicios ofrecidos se basa en el auto-reporte por parte de los establecimientos en los portales web facilitadores para la venta de alojamiento. Otros tipos de fuentes alternativas, como el estudio de reseñas por parte de usuarios pasados usando procesamiento de lenguaje natural, pueden servir para validar y complementar la información proveniente de la firma.

Este documento contribuye al estudio de la gestión de crisis y gestión de choques externos en el sector de alojamiento, un tema poco explorado en Colombia. La vulnerabilidad heterogénea que se encontró sugiere la posibilidad de crear instrumentos de política focalizados en función de las características los establecimientos y la composición de sus mercados locales. El diseño de políticas hoteleras locales debe tener en cuenta la vulnerabilidad causada por los niveles de diferenciación. Los indicadores desarrollados para este estudio pueden contribuir a la identificación anticipada de riesgos en mercados hoteleros.

El constante cambio en las necesidades de los turistas, y el reconocimiento de que la pandemia pudo cambiar el perfil de demanda de los clientes (nuevos tipos de turistas, cambios en demanda) puede generar cambios continuos en la estructura del mercado de alojamiento, por lo que este estudio permite diseñar una política pública sectorial para sopesar la heterogeneidad del sector, de modo que se introduzcan estrategias para “personalizar” los instrumentos de la política a las necesidades de los distintos agentes del mercado. De igual forma, tener en cuenta esta dinámica puede indicar al sector

privado la importancia de ubicarse mejor en relación con sus competidores y demás agentes del mercado al tomar decisiones gerenciales basada en la evidencia en relación con las distintas estrategias que pueden adoptarse para asegurar la resiliencia de los indicadores de venta del sector ante choques externos negativos.

Finalmente, este documento contribuye a una agenda poco explorada en la literatura académica y en el diseño de políticas públicas a partir de la construcción de índices de diferenciación. El presente trabajo sugiere algunas líneas de investigación como estudiar la competencia que se puede presentar entre ciudades por complementariedades de los servicios, y la adopción de estas técnicas para otros mercados altamente diferenciados como por ejemplo el de construcción

## 1.8. Referencias

Adhinugroho, Y., Putra, A. P., Luqman, M., Ermawan, G. Y., Mariyah, S., & Pramana, S. (2020). Development of online travel Web scraping for tourism statistics in Indonesia.

Aires Barroso, G., & Nicolau Mota, K. C. (2010). Marketing turístico internacional: la marca Brasil. *Estudios y perspectivas en turismo*, 19(2), 241-267.

Akerlof, G. A. (1978). The market for “lemons”: Quality uncertainty and the market mechanism. In *Uncertainty in economics* (pp. 235-251). Academic Press.

Alonso, A. D., Kok, S. K., Bressan, A., O’Shea, M., Sakellarios, N., Koresis, A., ... & Santoni, L. J. (2020). COVID-19, aftermath, impacts, and hospitality firms: An international perspective. *International journal of hospitality management*, 91, 102654.

Alonso-Almeida, M. D. M., Rodríguez-Antón, J. M., & Rubio-Andrada, L. (2012). Reasons for implementing certified quality systems and impact on performance: an analysis of the hotel industry. *The Service Industries Journal*, 32(6), 919-936.

Amaro, S., & Duarte, P. (2015). An integrative model of consumers’ intentions to purchase travel online. *Tourism management*, 46, 64-79.

American Hospitality and Lodging Association, 2020. Covid-19’s Impact on the Hotel Industry [accessed on 8 May 2020] Available at. <https://www.ahla.com/covid-19s-impact-hotel-industry>.

Arcay, A. O., & Benítez, F. G. (2006). Modelos de elección discreta en transportes con coeficientes aleatorios. *Cátedra abertis*.

- Archer, B., & Shea, S. (1977). Manpower in tourism: the situation in Wales. *Manpower in tourism: the situation in Wales*.
- Audretsch, D. B., & Feldman, M. P. (1996). Innovative clusters and the industry life cycle. *Review of industrial organization*, 11(2), 253-273.
- Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American economic review*, 86(3), 630-640.
- Bagnera, S.M., Stewart, E., Edition, S., 2020. Navigating hotel operations in times of COVID-19. *Boston Hospitality Rev.* ahead-of-print.
- Banerjee, A. V. (1992). A simple model of herd behavior. *The quarterly journal of economics*, 107(3), 797-817.
- Bar-Isaac, H., Caruana, G., & Cuñat, V. (2009). Costly search and design. Available at SSRN 1435257.
- Baum, J. A., & Haveman, H. A. (1997). Love thy neighbor? Differentiation and agglomeration in the Manhattan hotel industry, 1898-1990. *Administrative Science Quarterly*, 304-338.
- Baum, J., & Mezias, S. (1992). Localized competition and organizational failure in the Manhattan hotel industry, 1898–1990. *Administrative Science Quarterly*, 37(4), 580–604.
- Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242-262.
- Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841-890.
- Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100(5), 992-1026.
- Boni, P (2005) *The Distance Between a Hotel’s Keywords and Bookings: How to close the deal after Search Engine Optimization*.
- Borko S., Geerts W. & Wang H. (2020). *The travel industry turned upside down*. McKinsey.
- Branicki, L.J., Sullivan-Taylor, B., Livschitz, S.R., 2018. How entrepreneurial resilience generates resilient SMEs. *Int. J. Entrep. Behav. Res.* 24 (7), 1244–1263.
- Brown, N.A., Orchiston, C., Rovins, J.E., Feldmann-Jensen, S., Johnston, D., 2018. An integrative framework for investigating disaster resilience within the hotel sector. *J. Hosp. Tour. Manag.* 36, 67–75.

Browning, V., So, K. K. F., & Sparks, B. (2013). The influence of online reviews on consumers' attributions of service quality and control for service standards in hotels. *Journal of Travel & Tourism Marketing*, 30(1-2), 23-40.

Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. *Tourism management*, 29(4), 609-623.

Cabello Beltrán, E. A., & Pulido Arbeláez, L. N. (2020). Análisis del impacto económico del coronavirus en el sector turismo en Colombia.

Camison, C., Fores, B., Boronat-Navarro, M., & Puig-Denia, A. (2020). The effect of hotel chain affiliation on economic performance: The moderating role of tourist districts. *International Journal of Hospitality Management*, 87, 102493.

Carlbäck, M. (2012). Strategic entrepreneurship in the hotel industry: The role of chain affiliation. *Scandinavian Journal of Hospitality and Tourism*, 12(4), 349-372.

Carlton, D. W., & Dana Jr, J. D. (2008). Product variety and demand uncertainty: Why markups vary with quality. *The Journal of Industrial Economics*, 56(3), 535-552.

CBRE, 2020. Insights and Implications of Covid-19 Pandemic [accessed online on 8 May 2020] Available at: <https://www.cbre.com/covid-19>.

Chathoth, P.K., Olsen, M.D., 2003. Strategic alliances: a hospitality industry perspective. *Int. J. Hosp. Manag.* 22 (4), 419–434.

Chen, F., Federgruen, A., Zheng, Y.S., 2001. Coordination mechanisms for a distribution system with one supplier and multiple retailers. *Manag. Sci.* 47 (5), 693–708.

Chen, M. H. (2011). The response of hotel performance to international tourism development and crisis events. *International Journal of Hospitality Management*, 30(1), 200-212.

Chen, Y., & Xie, J. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management science*, 54(3), 477-491.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3), 345-354.

Choi, T. Y., & Chu, R. (2000). Levels of satisfaction among Asian and Western travellers. *International Journal of Quality & Reliability Management*.

Christopher, M., 2011. *Logistics & Supply Chain Management*. Pearson education limited, United Kingdom.

Chu, R. K., & Choi, T. (2000). An importance-performance analysis of hotel selection factors in the Hong Kong hotel industry: a comparison of business and leisure

travellers. *Tourism management*, 21(4), 363-377.

Chung, L. L. W. (1994). The economics of land-use zoning: a literature review and analysis of the work of Coase. *The Town Planning Review*, 65(1), 77–98.

Coenders, G., Espinet, J., & Saez, M. (2003). Predicting random level and seasonality of hotel prices: A latent growth curve approach. *Tourism Analysis*, 8(1), 15-31.

Datta, S., & Sudhir, K. (2013). Does reducing spatial differentiation increase product differentiation? Effects of zoning on retail entry and format variety. *Quantitative Marketing and Economics*, 11(1), 83-116.

Davis, B. (2021). Why accommodation is important to hospitality and tourism industry? – MVOrganizing. [www.mvorganizing.org](http://www.mvorganizing.org). <https://www.mvorganizing.org/why-accommodation-is-important-to-hospitality-and-tourism-industry/>

Debo, L. G., & Veeraraghavan, S. K. (2009). Models of herding behavior in operations management. In *Consumer-Driven Demand and Operations Management Models* (pp. 81-112). Springer, Boston, MA.

Dellarocas, C., & Wood, C. A. (2008). The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management science*, 54(3), 460-476.

Devesa, M. J. S., & Peñalver, L. F. M. (2013). Research Note: Size, Efficiency and Productivity in the Spanish Hotel Industry — Independent Properties versus Chain-Affiliated Hotels. *Tourism Economics*, 19(4), 801–809. <https://doi.org/10.5367/te.2013.0320>

Dubé, L., & Renaghan, L. M. (2000). Creating visible customer value: How customers view best-practice champions. *Cornell hotel and restaurant administration quarterly*, 41(1), 62-72.

Dube, L., & Renaghan, L. M. (2000). Marketing your hotel to and through intermediaries. *Cornell Hospitality Quarterly*, 41(1), 73.

Dubé, L., Enz, C. A., Renaghan, L. M., & Siguaw, J. A. (2000). Managing for excellence. *Cornell Hotel and Restaurant Administration Quarterly*, 41(5), 30-39.

Duranton, G., & Puga, D. (2020). The economics of urban density. *Journal of Economic Perspectives*, 34(3), 3-26.

Duro, J. A., Perez-Laborda, A., Turrion-Prats, J., & Fernández-Fernández, M. (2021). Covid-19 and tourism vulnerability. *Tourism Management Perspectives*, 38, 100819.

Eggers, F., 2020. Masters of disasters? Challenges and opportunities for SMEs in times of crisis. *J. Bus. Res.* 116, 199–208.

Eisenhardt, K. M., & Martin, J. A. (2000). Dynamic capabilities: what are they?. *Strategic management journal*, 21(10-11), 1105-1121.

Elshaer, Abdallah M., and Asmaa M. Marzouk. *Labor in the tourism and hospitality industry: skills, ethics, issues, and rights*. Apple Academic Press, 2019.

Espinet, J. M., Saez, M., Coenders, G., & Fluvà, M. (2003). Effect on prices of the attributes of holiday hotels: a hedonic prices approach. *Tourism Economics*, 9(2), 165-177.

European Parliament, 2020. Covid-19 and the Tourism Sector. Available at: <https://www.europarl.europa.eu/media/default.do?app=press&lang=en&docId=13222> (2020) 649368\_EN.pdf

EUROSTAT, 2017, Demunter, C. (2017). *Tourism statistics Early adopters of big data?: 2017 edition*.

Evans, D., 2020. *Airbnb's Future Is Uncertain as It Continues to Struggle Through Its Covid-19 Response*. CNBC [accessed online 11 May 2020] Available at: <https://www.cnbc.com/2020/05/06/can-airbnb-survive-the-coronavirus-pandemic.html>.

Farmaki, A., Miguel, C., Drotarova, M. H., Aleksić, A., Časni, A. Č., & Efthymiadou, F. (2020). Impacts of Covid-19 on peer-to-peer accommodation platforms: Host perceptions and responses. *International journal of hospitality management*, 91, 102663.

Feldman, M. P., & Audretsch, D. B. (1999). Innovation in cities:: Science-based diversity, specialization and localized competition. *European economic review*, 43(2), 409-429.

Feldman, P., Papanastasiou, Y., & Segev, E. (2019). Social learning and the design of new experience goods. *Management Science*, 65(4), 1502-1519.

Fluvà, M., Rigall-I-Torrent, R., Espinet, J. M., Garriga, A., & Saló, A. (2011). Precios implícitos de los atributos de los productos turísticos: ¿Qué esconde el efecto de la localización?. *Estudios de Economía Aplicada*, 29(3), 781-801.

Gabszewicz, J. J., Shaked, A., Sutton, J., & Thisse, J. F. (1986). Segmenting the market: The monopolist's optimal product mix. *Journal of economic theory*, 39(2), 273-289.

Glusac, E., 2020. *Hotels Vs Airbnb: Has Covid-19 Disrupted the Disrupter?* [accessed online on 14 June 2020]

Gonzalez. T. (2020). *La demanda extrahotelera cae 45,4% entre julio y agosto.*, Hosteltur, <https://www.hosteltur.com/139598-la-demanda-extrahotelera-cae-454-entre-julio-y-agosto.html>

González-Torres, T., Rodríguez-Sánchez, J. L., & PelechanoBarahona, E. (2021). Managing relationships in the Tourism Supply Chain to overcome epidemic outbreaks: The case of COVID-19 and the hospitality industry in Spain. *International Journal of Hospitality Management*, 92, 102733.

Guillet, B. D., & Law, R. (2010). Analyzing hotel star ratings on third-party distribution websites. *International Journal of Contemporary Hospitality Management*.

Güliz, N. & Akbıyık, A. (2020). Impacts of COVID-19 on global tourism industry: A cross-regional comparison. *Tourism Management Perspectives*, 36.

Hafsa, S. (2020). Impacts of COVID-19 Pandemic on Tourism & Hospitality Industry in Bangladesh. Available at SSRN 3659196.

Hao, F., Xiao, Q., & Chon, K. (2020). COVID-19 and China's hotel industry: Impacts, a disaster management framework, and postpandemic agenda. *International journal of hospitality management*, 90, 102636.

Haroutunian, S., Mitsis, P., & Pashardes, P. (2005). Using brochure information for the hedonic analysis of holiday packages. *Tourism economics*, 11(1), 69-84.

Hong, Y., & Pavlou, P. A. (2014). Product fit uncertainty in online markets: Nature, effects, and antecedents. *Information Systems Research*, 25(2), 328-344.

<https://www.mckinsey.com/industries/travel-logistics-and-infrastructure/our-insights/the-travel-industry-turned-upside-down-insights-analysis-and-actions-for-travel-executives>

ILO. (2021). Hotels, catering and tourism sector. International Labour Organization. <https://www.ilo.org/global/industries-and-sectors/hotels-catering-tourism/lang-en/index.htm>

Jaffe, A. B., & Trajtenberg, M. (2002). Patents, citations, and innovations: A window on the knowledge economy. MIT press.

Jaffe, A. B., Newell, R. G., & Stavins, R. N. (2002). Environmental policy and technological change. *Environmental and resource economics*, 22(1), 41-70.

Jaipuria, S., Parida S., & Ray, P. (2021). The impact of COVID-19 on tourism sector in India. *Tourism Recreation Research*, 46 (2), 245-260. DOI: 10.1080/02508281.2020.1846971

Jamal, T., Budke, C., 2020. Tourism in a world with pandemics: local-global responsibility and action. *J. Tour. Futures*.

Jannach, D., Gedikli, F., Karakaya, Z., & Juwig, O. (2012). Recommending hotels based on multi-dimensional customer ratings. In ENTER (pp. 320-331).

Jayawardena, C., Tew, P.J., Lu, Z., Tolomiczenko, G., Gellatly, J., 2008. SARS: lessons in strategic planning for hoteliers and destination marketers. *Int. J. Contemp.*

Hosp. Manage. 20 (3), 332–346.

Jiang, Y., Ritchie, B.W., Benckendorff, P., 2019. Bibliometric visualisation: an application in tourism crisis and disaster management research. *Curr. Issues Tour.* 22 (16), 1925–1957.

Kopczuk, W., 2020. *Some Thoughts on Economics and Policy at the Time of the Pandemics*. Department of Economics and SIPA, Columbia University, New York.

Kuckertz, A., Brandle, L., Gaudig, A., Hinderer, S., Reyes, C.A.M., Prochotta, A., Steinbrink, K.M., Berger, E.S., 2020. Startups in times of crisis—a rapid response to the COVID-19 pandemic. *J. Bus. Ventur. Insights*, e00169.

Kuksov, D. (2004). Buyer search costs and endogenous product design. *Marketing Science*, 23(4), 490–499.

Kuo, H.I., Chen, C.C., Tseng, W.C., Ju, L.F., Huang, B.W., 2008. Assessing impacts of SARS and Avian Flu on international tourism demand to Asia. *Tour. Manag.* 29 (5), 917–928.

Kwark, Y., Chen, J., & Raghunathan, S. (2014). Online product reviews: Implications for retailers and competing manufacturers. *Information systems research*, 25(1), 93-110.

Leung, R., Au, N., Liu, J., & Law, R. (2018). Do customers share the same perspective? A study on online OTAs ratings versus user ratings of Hong Kong hotels. *Journal of Vacation Marketing*, 24(2), 103-117.

Li, X., Chen, H., Lu, L., Chen, L. L., Chan, B. P. C., Wong, S. C., ... & To, K. K. W. (2022). High compliance to infection control measures prevented guest-to-staff transmission in COVID-19 quarantine hotels. *Journal of Infection*, 84(3), 418-467.

Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism management*, 68, 301-323.

Liu, M. T., Wong, I. A., Tseng, T. H., Chang, A. W. Y., & Phau, I. (2017). Applying consumer-based brand equity in luxury hotel branding. *Journal of business research*, 81, 192-202.

Liu, X., Schuckert, M., & Law, R. (2015). Can response management benefit hotels? Evidence from Hong Kong hotels. *Journal of Travel & Tourism Marketing*, 32(8), 1069-1080.

Lord, P., Saad, L., 2020. Tackling the COVID-19 Pandemic. <https://doi.org/10.1016/j.arcmed.2020.04.012>

Luca, M. (2016). Reviews, reputation, and revenue: The case of Yelp. com. *Com*

- (March 15, 2016). Harvard Business School NOM Unit Working Paper, (12-016).
- Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science*, 62(12), 3412-3427.
- Marco-Lajara, B., Claver-Cortés, E., Úbeda García, M., & Zaragoza Sáez, P. D. C. (2016). Do Hotels Benefit from Agglomeration?.
- Marco-Lajara, B., Claver-Cortés, E., Úbeda-García, M., & Zaragoza-Sáez, P. D. C. (2016). Hotel performance and agglomeration of tourist districts. *Regional Studies*, 50(6), 1016-1035.
- Marshall, A. (1920). *Principles of economics*. London: Macmillan.
- McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics*, 3(4), 303-328.
- Milgrom, P., & Roberts, J. (1986). Price and advertising signals of product quality. *Journal of Political Economy*, 94(4), 796–821.
- Monczka, R.M., Petersen, K.J., Handfield, R.B., Ragatz, G.L., 1998. Success factors in strategic supplier alliances: the buying company perspective. *Decis. Sci.* 29 (3), 553–577.
- Montoya, Y. M., & Pérez, D. A. P. (2016). Turismo en Colombia ¿Es el turismo una alternativa real para enfrentar el déficit externo? *Ploutos*, 6(2), 48-55.
- Naciones Unidas. (2008). *Recomendaciones internacionales para estadísticas de turismo 2008*.
- Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78(2), 311–329.
- Neuburger, L., & Egger, R. (2021). Travel risk perception and travel behaviour during the COVID-19 pandemic 2020: a case study of the DACH region. *Current Issues in Tourism*, 24 (7), 1003-1016. DOI: 10.1080/13683500.2020.1803807
- Neuburger, L., & Egger, R. (2021). Travel risk perception and travel behaviour during the COVID-19 pandemic 2020: A case study of the DACH region. *Current Issues in Tourism*, 24(7), 1003-1016.
- Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, M., Agha, R., 2020. The socio-economic implications of the coronavirus and COVID-19 pandemic: a review. *Int. J. Surg.* 78, 185–193.
- Nicolaidis, C., Avraam, D., Cueto-Felgueroso, L., Gonzalez, ´ M.C., Juanes, R., 2019. Hand hygiene mitigation strategies against global disease spreading through the air transportation network. *Risk Anal.* 40 (4), 723–740.

Nosko, C., & Tadelis, S. (2015). The limits of reputation in platform markets: An empirical analysis and field experiment (No. w20830). National Bureau of Economic Research. Norris, F.H., Stevens, S.P., Pfefferbaum, B., Wyche, K.F., Pfefferbaum, R.L., 2008. Community resilience as a metaphor, theory, set of capacities, and strategy for disaster readiness. *Am. J. Commun. Psychol.* 41 (1–2), 127–150.

Novelli, M., Burgess, L.G., Jones, A., Ritchie, B.W., 2018. No Ebola... still doomed—the Ebola-induced tourism crisis. *Ann. Tour. Res.* 70, 76–87.

O'Connor, P., & Frew, A. J. (2002). The future of hotel electronic distribution: Expert and industry perspectives. *Cornell Hotel and Restaurant Administration Quarterly*, 43(3), 33-45.

O'Neill, J. W., & Carlbäck, M. (2011). Do brands matter? A comparison of branded and independent hotels' performance during a full economic cycle. *International Journal of Hospitality Management*, 30(3), 515-521.

O'Neill, J. W., & Xiao, Q. (2006). The role of brand affiliation in hotel market value. *Cornell Hotel and Restaurant Administration Quarterly*, 47(3), 210-223.

OECD, 2020. Tourism Policy Responses to the Coronavirus (COVID-19). Obtenido de: <https://www.oecd.org/coronavirus/policy-responses/tourism-policy-responses-to-the-coronavirus-covid-19-6466aa20/#endnotea0z14>. (Acceso: enero, 2021).

Oh, R.S., Raven, P. & DeKay, F. (2011) Selling rooms: hotels vs. third-party websites. *Cornell Hospitality Quarterly* 52(2): 181–189.

Onur, I., Bruwer, J., & Lockshin, L. (2020). Reducing information asymmetry in the auctioning of non-perishable experience goods: The case of online wine auctions. *Journal of Retailing and Consumer Services*, 54, 102060.

Pal, R., Torstensson, H., Mattila, H., 2014. Antecedents of organizational resilience in economic crises—an empirical study of Swedish textile and clothing SMEs. *Int. J. Prod. Econ.* 147, 410–428.

Park, E., Kim, W. & Kim, S. (2020). Tracking tourism and hospitality employees' real-time perceptions and emotions in an online community during the COVID-19 pandemic. *Current Issues in Tourism*. DOI: 10.1080/13683500.2020.1823336

Peña. J. (2021). Turismo de negocios y reuniones se reactiva a menor velocidad. *Portafolio*. <https://www.portafolio.co/negocios/empresas/turismo-de-negocios-y-reuniones-se-reactiva-a-menor-velocidad-559383>

Phillips, P., Zigan, K., Silva, M. M. S., & Schegg, R. (2015). The interactive effects of online reviews on the determinants of Swiss hotel performance: A neural network

analysis. *Tourism Management*, 50, 130-141.

Pizam, A. (1982). Tourism manpower: the state of the art. *Journal of Travel research*, 21(2), 5-9.

Pizam, A., Fleischer, A., 2002. Severity versus frequency of acts of terrorism: which has a larger impact on tourism demand? *J. Travel. Res.* 40 (3), 337–339.

Pizam, Abraham. "Tourism manpower: the state of the art." *Journal of Travel research* 21.2 (1982): 5-9.

Pogodzinski, J. M. (1991). The effects of fiscal and exclusionary zoning on household location: a critical review. *Journal of Housing Research*, 2(2), 145–160.

Polemis, M. (2020). The impact of COVID-19 on hotel performance: Evidence from a Difference-in-Differences approach.

PRAMANA, S., PARAMARTHA, D. Y., ADHINUGROHO, Y., & NURMALA-SARI, M. (2020). Air pollution changes of Jakarta, Banten, and West Java, Indonesia during the first month of COVID-19 pandemic. *Asian Journal of Business Environment*, 10(4), 15-19. Pramana, S., Paramartha, D. Y., Ermawan, G. Y., Deli, N. F., & Srimulyani, W. (2021). Impact of COVID-19 pandemic on tourism in Indonesia. *Current Issues in Tourism*, 1-21.

Rhee, H. T., & Yang, S. B. (2015). Does hotel attribute importance differ by hotel? Focusing on hotel star-classifications and customers' overall ratings. *Computers in Human Behavior*, 50, 576-587.

Rigall-I-Torrent, R., & Fluvià, M. (2011). Managing tourism products and destinations embedding public good components: a hedonic approach. *Tourism Management*, 32(2), 244-255.

Rigall-I-Torrent, R., Fluvià, M., Ballester, R., Saló, A., Ariza, E., & Espinet, J. M. (2011). The effects of beach characteristics and location with respect to hotel prices. *Tourism Management*, 32(5), 1150-1158.

Rodríguez-Antón, J. M., del Mar Alonso-Almeida, M., Celemín, M. S., & Rubio, L. (2012). Use of different sustainability management systems in the hospitality industry. The case of Spanish hotels. *Journal of Cleaner Production*, 22(1), 76-84.

Roger-Monzó, V., Martí-Sánchez, M., & Guijarro-García, M. (2015). Using online consumer loyalty to gain competitive advantage in travel agencies. *Journal of Business Research*, 68(7), 1638-1640.

Runfola, A, Rosati, M, Guercini, S (2013). New business models in online hotel distribution: emerging private sales versus leading IDS. *Service Business* 7(2): 183–205.

Scaglione, M., Schegg, R. (2015) The impact of attribute preferences on adoption timing of hotel distribution channels: are OTAs winning the customer race? In: Tussyadiah, I., Inversini, A. (eds) *Information and Communication Technologies in Tourism 2015*. London: Springer International Publishing, pp. 681–693.

Seabra, C., Reis, P., Abrantes, J.L., 2020. The influence of terrorism in tourism arrivals: a longitudinal approach in a Mediterranean country. *Ann. Tour. Res.* 80, 102811.

Serrano Antón, F., 2020. Medidas fiscales terapéuticas y COVID 19, 7. Instituto Complutense de Estudios Internacionales, pp. 1–4.

Shaked, A., & Sutton, J. (1987). Product differentiation and industrial structure. *The Journal of Industrial Economics*, 131-146.

Shapiro, C. (1983). Optimal pricing of experience goods. *The Bell Journal of Economics*, 497-507.

Sharma, A., and Nicolau, J.L. (2020). An open market valuation of the effects of COVID-19 on the travel and tourism industry, *Annals of Tourism Research*, <https://doi.org/10.1016/j.annals.2020.102811>.

Sigala, M., 2004. Collaborative supply chain management in the airline sector: the role of global distribution systems (GDS). *Adv. Hosp. Leis.* 1 (1), 103–121.

Sigala, M., 2017. Collaborative commerce in tourism: implications for research and industry. *Curr. Issues Tour.* 20 (4), 346–355.

Tapper, R., Font, X., 2004. *Tourism Supply Chains*. Report of a Desk Research Project for The Travel Foundation, Leeds Metropolitan University, Leeds.

Teece, D. J., Pisano, G., & Shuen, A. (1997). (1997). Dynamic capabilities and strategic management. *Resources, Firms, and Strategies: A Reader in the Resource-Based Perspective*, 268.

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic management journal*, 18(7), 509-533.

Thrane, C. (2005). Hedonic price models and sun-and-beach package tours: The Norwegian case. *Journal of Travel Research*, 43(3), 302-308.

Tsang, E. W., & Yip, P. S. (2009). Competition, agglomeration, and performance of Beijing hotels. *The Service Industries Journal*, 29(2), 155-171.

Tunc, M. M., Cavusoglu, H., & Raghunathan, S. (2019). Two-Sided Adverse Selection and Bilateral Reviews in Sharing Economy. Available at SSRN 3499979.

Úbeda-García, M., Claver-Cortés, E., Marco-Lajara, B., & Zaragoza-Sáez, P. (2016). Toward organizational ambidexterity in the hotel industry: The role of human resour-

ces. *Cornell Hospitality Quarterly*, 57(4), 367-378.

United Nations. (2016). Labour intensive industries. United Nations Economic and Social Commission for Western Asia (ESCWA). <https://archive.unescwa.org/labour-intensive-industries>

UNWTO. (2020). COVID-19 RESPONSE: 96 % OF GLOBAL DESTINATIONS IMPOSE TRAVEL RESTRICTIONS. UNWTO REPORTS. Madrid. 17 de abril de 2020.

UNWTO. (2021). Conceptual Guidance on Tourism Statistics in the COVID-19 Context. Madrid.

Voltes-Dorta, A., & Inchausti-Sintes, F. (2021). The spatial and quality dimensions of Airbnb markets. *Tourism Economics*, 27(4), 688-702.

Wang, C. L., & Ahmed, P. K. (2007). Dynamic capabilities: A review and research agenda. *International journal of management reviews*, 9(1), 31-51.

Wu, E.H., Law, R., Jiang, B., 2010. The impact of infectious diseases on hotel occupancy rate based on independent component analysis. *Int. J. Hosp. Manag.* 29 (4), 751–753.

Wu, F. (1999). Intrametropolitan FDI firm location in Guangzhou, China A Poisson and negative binomial analysis. *The Annals of Regional Science*, 33(4), 535-555.

Yehia, Y. (2019). The Importance of Tourism on Economies and Businesses. *GlobaLEDGE*. <https://globaledge.msu.edu/blog/post/55748/the-importance-of-tourism-on-economies-a>

Ying, T., Wang, K., Liu, X., Wen, J., Goh, E., 2020. Rethinking game consumption in tourism: a case of the 2019 novel coronavirus pneumonia outbreak in China. *Tour. Recreat. Res.* 1–6.

Zenker, S., and Kock, F. (2020). The coronavirus pandemic – A critical discussion of a tourism research agenda, *Tourism Management*, 81, <https://doi.org/10.1016/j.tourman.2020.104164>.

Zhang, X., Song, H., Huang, G.Q., 2009. Tourism supply chain management: a new research agenda. *Tour. Manag.* 30 (3), 345–358.

## Capítulo 2

# Impacto del modelo de recolección sobre la captura de ingresos. Un estudio en los tiempos del COVID-19

### 2.1. Introducción

Durante el segundo trimestre de 2020, Gobiernos de todo el mundo impusieron medidas de confinamiento y distanciamiento social con el objetivo de contener la propagación de la pandemia causada por el coronavirus SARS-CoV-2. Esta situación trajo consigo retos importantes a los Institutos Nacionales de Estadística (INE) para monitorear las principales variables económicas y sociales, uno de los insumos más relevantes para la toma de decisiones en políticas públicas.

En este contexto, operaciones estadísticas continuas que suelen recogerse mensualmente mediante entrevistas presenciales, como las encuestas de hogares, se vieron afectadas, por lo cual tuvieron que ser rediseñadas para mantener el proceso de recolección de datos. De hecho, la Comisión Estadística de Naciones Unidas (UNSTATS, por sus siglas en inglés) (2020) documentó que el 96 % de los INE dejaron de recopilar datos de manera presencial, parcial o totalmente.

En un primer momento, las restricciones para implementar las entrevistas presenciales obligaron a realizar las encuestas telefónicamente, a través del método denominado

*Computer-Assisted Telephone Interviews* (CATI, por sus siglas en inglés) y a reducir la longitud de los cuestionarios, en línea con las recomendaciones que fueron expedidas por la Organización de Naciones Unidas (ONU) y la Organización Internacional del Trabajo (OIT)<sup>1</sup>, cuya finalidad era mantener la calidad de los datos.

Este cambio en el proceso de recolección de datos fue temporal, debido a que el grado de severidad de las políticas de confinamiento y distanciamiento social fue disminuyendo progresivamente. No obstante, a pesar de que los INE normalizaron sus operaciones, se requiere evaluar el impacto de estos cambios, particularmente sobre los indicadores del mercado laboral (UNSTATS, 2020).

La variable de ingresos es probablemente uno de los datos más relevantes que se recogen a través de las encuestas de hogares, ya que proporciona información sobre la situación socioeconómica y el nivel de vida de los hogares, y sirve como insumo para la medición de la pobreza (Glewwe, 2007; Lohmann, 2011; Burton et al., 2020). De hecho, hay estudios en esta línea que reconocen que los errores de medición asociados a la recolección de datos de ingresos afectan las medidas de desigualdad y pobreza (van Praag et al., 1983; Ravallion, 1988; Chesher y Schluter, 2002; Figari et al., 2012; Angel et al., 2018; y Fessler et al., 2018). Por este motivo, es crucial comprender el impacto del proceso de recopilación de datos sobre la medición de los ingresos.

El Departamento Administrativo Nacional de Estadística de Colombia (DANE), implementó modos de recolección mixtos, de acuerdo con las regiones donde realizaba sus operaciones. Por ejemplo, durante los cierres estrictos (de marzo a julio del 2020) se mantuvieron las entrevistas presenciales en las zonas menos urbanizadas y rurales, mientras que los datos en las principales ciudades y áreas metropolitanas se recogieron de manera remota, mediante CATI. La naturaleza de esta distinción constituye un experimento natural para entender el impacto generado por el cambio en el modo de recolección.

Motivado por lo anterior, este trabajo se propone estudiar el impacto del cambio temporal en la recolección de datos de la encuesta de hogares de Colombia, Gran Encuesta Integrada de Hogares (GEIH), en la medición de los ingresos en el periodo correspondiente de mayo a julio del 2020.

---

<sup>1</sup>Por ejemplo, la OIT elaboró una serie de documentos de trabajo llamados COVID-19: Orientación para la recopilación de datos sobre estadísticas del trabajo (Guidance for labour statistics data collection), en los que se analizan elementos clave para la realización de encuestas rápidas, se definen cuestionarios más breves que contienen la información esencial y se ofrecen consejos para captar datos pertinentes durante la crisis de COVID-19, como la prevalencia del trabajo a distancia y el teletrabajo, la inactividad, la ausencia de empleo y la pérdida de empleo, entre otros.

De igual modo, es pertinente tener en cuenta que el estudio de los sesgos en la medición de los ingresos es un instrumento crucial para validar la comparabilidad de las variables relevantes recogidas en las encuestas de hogares, así como para determinar diseños metodológicos que combinen diferentes modos de entrevista para reducir costos operativos. En ese sentido, este trabajo propone explotar la variación temporal y espacial en el proceso de recolección de datos, así como integrar a la encuesta datos administrativos para cuantificar el sesgo en los ingresos auto-reportados. Específicamente, se estudiará si la medición de los ingresos es sensible a la implementación de métodos de recolección tipo CATI, en lugar del modo presencial habitual, también conocido como *Computer-Assisted Personal Interviewing* (CAPI, por sus siglas en inglés).

Para ello, se utilizará un diseño cuasi-experimental, en particular especificaciones tipo *event study*, que permite calcular si existe una diferencia en los ingresos declarados como resultado de la aplicación de CATI durante los meses de mayo a julio del 2020 en Colombia, partiendo de la hipótesis de que el uso de información proveniente de registros administrativos permite establecer un valor de referencia de los ingresos laborales y controlar por posibles heterogeneidades no observadas asociadas a la coyuntura del COVID-19. En el estudio también se exploran posibles efectos heterogéneos por grupos de población, como género, educación y grupos de edad y chequeos de robustez de la estrategia de identificación.

Un elemento que servirá como insumo fundamental para este trabajo es el registro estadístico de relaciones laborales desarrollado por el DANE, con el fin de monitorear los efectos en el mercado laboral. Esta fuente fue construida a partir de la integración de registros administrativos de seguridad social que no se vieron alterados por los cambios en el modo de recolección de la encuesta, y por tanto sirven como contrafactual para una medición de posibles sesgos debidos al cambio en el modo de recolección.

A partir de lo anterior, este documento se organiza en seis secciones adicionales a esta introducción. En la primera se presentará el marco teórico que sirve como fundamento para el desarrollo de las hipótesis y metodologías del estudio. En la segunda se describen los cambios ocasionados en los métodos de recolección de la encuesta de hogares durante el 2020 así como una descripción del uso y la integración de registros administrativos y las encuestas. En la tercera se describen los datos y la estrategia empírica. En la cuarta se presentan los resultados principales y la quinta los chequeos de robustez. Finalmente, en la sexta sección se evidencian las conclusiones.

Existe una amplia literatura que examina las fuentes de errores de medición de los

ingresos en las encuestas de hogares. En esta línea, Angel et al. (2019) presentan un marco general de las principales fuentes de error en la medición de ingresos, en particular: comportamientos socialmente deseables, características socioeconómicas, diseño de la encuesta y efectos de aprendizaje (que se refiere a las estrategias de los encuestados para recordar la información).

En general, todas las fuentes pueden producir tanto sobre-reporte como sub-reporte de ingresos. Por ejemplo, en lo que respecta a los comportamientos socialmente deseables, se han observado sesgos, posiblemente relacionados con características de los hogares y los individuos: las personas con ingresos más bajos declaran en exceso sus ingresos para evitar revelar sus condiciones económicas reales, mientras que las que tienen ingresos elevados pueden declarar valores menores. Estos patrones se conocen como errores de reversión de la media (véase e.g. Bound et al., 1994, Moore et al., 2000; Kim y Tamborini, 2014).

Otros factores asociados a la variable ingreso, e.g., la variable objetivo de medición (ingreso bruto o ingreso disponible), son los montos, la redacción de las preguntas, el periodo de referencia y el modo de la encuesta, también afectan su medición (véase Moore et al., 2000; Canberra Group, 2011; Kuhn, 2019). Para un mayor debate sobre el error de medición en los datos de las encuestas, ver Horowitz y Manski (1995) y Bound et al. (2001)<sup>2</sup>

La literatura existente dedicada a comparar procesos de recolección de información basados en CAPI y CATI afirma que el modo CATI tiende a causar un sesgo en la cobertura y la representatividad. Ellis y Krosnick (1999) y Bowling (2005) muestran que una encuesta basada en CATI podría tener una mayor representación de individuos de altos ingresos, más educados y jóvenes (véase también Jordan et al., 1980; De Leeuw, 1992), lo cual está relacionado con la falta de cobertura de los servicios telefónicos en

---

<sup>2</sup>En la práctica, la rápida respuesta durante el periodo de confinamiento estricto para habilitar la implementación de las encuestas de hogares usando CATI requirió superar retos tanto en la recopilación de datos, como a nivel metodológico. En cuanto a los retos en el proceso de recolección, la OIT (2020a) sostiene que el cambio de CAPI a CATI requiere un registro administrativo comprensivo de los números de teléfono que permita preservar la estructura de muestreo, diseñar protocolos para sustituir la no-respuesta, entrenar rápidamente a los entrevistadores y ajustar la batería de preguntas del cuestionario de la encuesta. En cuanto a las cuestiones metodológicas, se requieren protocolos para corregir datos incompletos (e.g., la corrección de los factores de expansión) y un mayor entendimiento del sesgo de medición posterior a la recogida de datos para preservar la comparabilidad en el tiempo. En consecuencia, los INE han aplicado estrategias (e.g., reutilizar las muestras antiguas o contactar previamente con los hogares usando cartas introductorias para promover la participación) para mantener el diseño del muestreo (OIT, 2020b).

los otros grupos de población. Estudios en esta línea han documentado otras diferencias desde el punto de vista operativo. En particular, CAPI ofrece algunas ventajas, debido al contacto físico y visual entre entrevistado y entrevistador, permitiendo utilizar herramientas de comunicación adicionales para guiar a los encuestados y crear confianza y legitimidad para el encuestado (ver Statistics Canada, 2010).

Por otro lado, CAPI resulta un método más eficaz para los cuestionarios largos, al tiempo que se reduce la tasa de no respuesta (Roberts et al., 2010; Klausch et al., 2013, UK Data Service, 2019), lo que también podría mejorar la calidad los datos sobre ingresos Canberra Group (2011). Mientras que CAPI podría sobre-reportar comportamientos socialmente deseables como menores tasas de consumo de alcohol o mayor proporción de votantes (Holbrook et al., 2003). Jäckle et al. (2010), Schräpler et al. (2010), Caeyers et al. (2010), Statistics Canada (2010), Lynn y Kaminska (2013) realizan estudios amplios que comparan el impacto del uso de CAPI, CATI y otros modos de entrevista.

Es importante anotar que la integración de datos provenientes de encuestas y datos administrativos permiten comprender mejor las fuentes de error de medición y contrastar los patrones de reversión de la media, como han analizado previamente Bollinger (1998), Bound y Krueger (1991) Kreiner et al. (2013), Meyer y Mittag (2021), entre otros. De este modo, los resultados que se derivan de este trabajo son también relevantes para diseñar marcos metodológicos que permitan mejorar los datos de las encuestas, a partir del uso de registros administrativos (véase c.f. Abowd y Stinson, 2013), así como para ampliar el conocimiento sobre la comparabilidad y complementariedad de la unidad de observación de estas dos fuentes de información (ver Kapteyn e Ypma, 2007; Meyer et al., 2015, para una discusión)

## **2.2. Cambios en los métodos de recolección en la encuesta de hogares**

Durante la pandemia del coronavirus COVID-19 y las restricciones de movilidad, principalmente en los centros urbanos, en 2020, el DANE implementó diferentes medidas que permitieron garantizar la continuidad de la producción de los principales indicadores de mercado laboral que se derivan de la GEIH. Este proceso de adaptación se dio entre marzo y julio del 2020 y consistió en la implementación de la GEIH a

través de un operativo telefónico y un cuestionario más corto al habitual<sup>3</sup> para las 23 principales ciudades y Áreas Metropolitanas (AM). Para los dominios conocidos como resto urbano (cabeceras municipales diferentes a las 23 principales ciudades y AM) y los centros poblados y rural disperso (zona rural), se mantuvo un operativo presencial y se aplicó el formulario completo de la GEIH.

El cambio de método de recolección podría causar en primera instancia una selección de la muestra, entendida como XXX<sup>4</sup> asociada a la ausencia de información de contacto telefónico (CEPAL, 2020). No obstante, el DANE realizó un esfuerzo logístico importante que permitió no restringir la selección a los hogares que aparecieran en un directorio telefónico. Por el contrario, el muestreo se mantuvo de forma estratificada y por conglomerados geográficos. Luego de que las unidades geográficas fueron seleccionadas, se usaron diferentes canales de comunicación para obtener los números telefónicos. En caso de no ser posible la obtención de números telefónicos, se desplegaban funcionarios en campo para obtenerlos. Si el hogar no podía ser contactado de esta forma, otro hogar del mismo conglomerado geográfico era seleccionado. Por lo tanto, es posible suponer que la representatividad de la muestra no haya cambiado de forma significativa durante el periodo de aplicación de operativo CATI.

Para validar esta suposición, la Tabla 1 presenta la proporción de mujeres calculada a partir de las muestras expandidas (i.e., usando factores de expansión) para diferentes periodos y desagregaciones geográficas.

Tabla 1: Proporción de mujeres en la muestra de la GEIH 2019-2021

Periodo	Áreas metropolitanas	Otras ciudades	Rural	Total
Ene-Mar 2019	51,85	51,34	47,15	50,66
Abr-Jun 2019	51,85	51,34	47,14	50,66
Jul-Sep 2019	51,85	51,34	47,14	50,66
Oct-Dic 2019	51,85	51,34	47,14	50,66
Ene-Feb 2020	51,85	51,34	47,14	50,66
May-Jul 2020	51,84	51,34	47,14	50,66
Ago-Dic 2020	51,84	51,34	47,14	50,66
Ene-Mar 2021	51,83	51,35	47,14	50,66
Abr-Jun 2021	51,83	51,36	47,14	50,66

Fuente: Cálculos propios. Áreas metropolitanas hace referencia a las 23 principales ciudades, Otras ciudades es el resto de población urbana.

En estas estadísticas se observan variaciones no significativas, evidenciando que no

<sup>3</sup>El formulario de la GEIH se recortó de 200 a 39 preguntas entre marzo y abril del 2020 (DANE, 2020), y aumentó a 60 preguntas de mayo a julio del 2020.

<sup>4</sup>DANE (2020) aclara que el procedimiento de selección de la muestra entre marzo julio de 2020, para las 23 ciudades principales y áreas metropolitanas, no se modificó.

se presentaron problemas asociados al muestreo, lo cual es fundamental para reducir posibles efectos que confundan el impacto en el método de recolección. En particular, la proporción de mujeres permanece estable alrededor del 50.7% a nivel nacional. En los tres dominios de la encuesta (áreas metropolitanas, otras ciudades y zonas rurales) y las proporciones también se mantuvieron estables.

Este mismo ejercicio se presenta en la Tabla 2 considerando la composición de la población por grupos etarios:

Tabla 2: Proporciones por grupos de edad en la muestra de la GEIH 2019-2020

Periodo	Menor de 25	Entre 25 y 54	Mayor de 54
Ene-Mar 2019	42,49	40,16	17,35
Abr-Jun 2019	42,37	40,17	17,45
Jul-Sep 2019	42,26	40,18	17,56
Oct-Dic 2019	42,14	40,20	17,66
Jan-Feb 2020	42,05	40,21	17,75
Mar-Jul 2020	41,87	40,22	17,91
Ago-Dic 2020	41,72	40,24	18,04
Ene-Mar 2021	41,57	40,25	18,18
Abr-Jun 2021	41,45	40,26	18,29

Fuente: Cálculos propios.

En este ejercicio también se observa que no hubo una variación significativa. Las personas de menos de 25 años consistentemente corresponden al 42% de la población durante los dos años analizados, aquellos entre 25 y 54 corresponden al 40%, y las personas de más 55 años en adelante corresponden a alrededor del 18%.

## 2.3. Datos y estratégica empírica

### 2.3.1. Datos

Para estimar el impacto causal del cambio en el método de recolección sobre los registros de ingreso laboral, se utilizan dos fuentes de información: la encuesta de hogares y un registro estadístico construido a partir de datos de pagos de seguridad social. La primera fuente de información consiste en la GEIH, la cual recoge mensualmente las principales variables que permiten caracterizar el mercado laboral colombiano. En 2020 esta encuesta recogió información promedio de 19.152 hogares al mes y generó información representativa para dominios geográficos de las principales 23 ciudades y sus áreas metropolitanas, así como para el resto de la cobertura geográfica del país. La GEIH tiene una estructura modular, en la que se incluye información sociodemográfica,

así como variables de características de empleo como el ingreso de los ocupados. Para los ejercicios propuestos se utilizan los datos correspondientes al periodo enero de 2019 a diciembre de 2020.

Adicionalmente, con el fin de tener una medición de referencia del ingreso que no haya sido afectada por el cambio en el método de recolección, la segunda fuente de información es aquella proveniente del Registro Estadístico de Relaciones Laborales (RELAB) producido por el DANE. En este registro es posible observar la relación empleado-empendedor de los dependientes<sup>5</sup> a través de los aportes que se realizan en el sistema de seguridad social. El RELAB incluye información sobre un subconjunto de las relaciones laborales, que, al ser integrado con otras operaciones estadísticas del DANE, permite caracterizar las relaciones laborales por características demográficas como sexo o edad.

El DANE utiliza RELAB como fuente de información para el análisis de coyuntura del mercado laboral de Colombia y el análisis de consistencia de diferentes operaciones estadísticas de mercado laboral y actividades económicas. El insumo principal del RELAB es el registro administrativo<sup>6</sup> Planilla Integrada de Liquidación de Aportes (PILA) del Ministerio de Salud y Protección Social de Colombia. La variable ingreso se construye a partir de la información que reportan los empleadores o aportantes para el pago de subsistemas de seguridad social, como salud o pensión. Además, se aplican reglas de calidad sobre la información cruda que trata información duplicada o redundante que solo tiene fines administrativos o legales.

La integración de la información a nivel individual de la encuesta de hogares GEIH y el registro estadístico RELAB se hace a través de un cruce determinístico con las llaves de identificación disponibles en las dos fuentes de información. Esta integración permite controlar las posibles heterogeneidades que se pueden encontrar entre dominios geográficos<sup>7</sup> y evaluar los potenciales sesgos del cambio en el modo de recolección de la GEIH al pasar del operativo presencial al no presencial (telefónicamente) en los meses de marzo a julio de 2020. Vale anotar que la muestra resultante de esta integración

---

<sup>5</sup>En el RELAB, los dependientes corresponden a las posiciones ocupacionales de GEIH de obreros o empleados de empresa particular o del gobierno y los jornaleros o peones. Estas posiciones ocupacionales también son conocidas como asalariados, dependientes o trabajadores por cuenta ajena.

<sup>6</sup>Los datos administrativos se han convertido en una fuente valiosa de información para generar estadísticas y desarrollar análisis técnicos aun cuando su diseño no responde originalmente a las especificaciones técnicas de propósitos estadísticos (Naciones Unidas 2007; Wallgren y Wallgren, 2007 ; Zhang 2011; Kuhn, U. 2019).

<sup>7</sup>Como las 23 ciudades, el resto urbano y rural disperso.

corresponde a un grupo de trabajadores con relaciones laborales formales. Si bien no corresponde una muestra representativa del total de mercado laboral colombiano, debido a sus altas tasas de informalidad, la consistencia en los datos permite evaluar los impactos en el uso de diferentes métodos de recolección.

### 2.3.2. Metodología

Teniendo en cuenta la variación temporal en la implementación de los modelos de recolección, CATI y CAPI, y la heterogeneidad geográfica, la estrategia de identificación propuesta explota a nivel de corte transversal y temporal. La variación de corte transversal se debe a que la estrategia de recolección telefónica de datos fue usada solo en las 23 principales y áreas metropolitanas, mientras que la variación temporal viene de la aplicación de este procedimiento solo entre los meses de marzo y julio de 2020. Finalmente, el diseño metodológico se aprovecha gracias a que una fuente de datos tuvo los cambios previamente descritos y a que otra no sufriera modificaciones,, ya que el proceso mediante el cual se generan los datos del registro administrativo RELAB no cambió durante el 2020. De esta manera, el efecto de la implementación de CATI se deriva del cambio en la discrepancia en el reporte de ingresos entre la encuesta y el registro administrativo.

Es natural encontrar discrepancias entre el ingreso laboral auto-reportado y el registro del RELAB. Abowd y Stinson (2013) argumentan que las diferencias entre fuentes no deben confundirse con errores de medición, ya que ambas fuentes son sensibles a errores de medición. En la literatura Nordberg et al. (2001) encontraron que las estimaciones de ingresos derivadas de los registros administrativos son confiables y generalmente más altas que los ingresos derivados de encuestas, excepto en el caso de hogares de bajos ingresos.

A través de la comparación de encuestas y registros administrativos se ha intentado caracterizar los sesgos de reporte de ingresos en encuestas. En esta línea se ha encontrado un efecto predominante en los extremos de las distribuciones del ingreso. Por ejemplo, se observan reportes sobrestimados en la población de ingresos bajos y un sub-reportaje en los hogares de altos ingresos (ver Bound et al. 2001; Gottschalk y Huynh 2010; Kim y Tamborini 2012, 2014; Consolini y Donatiello, 2013).

En la medida que estas diferencias sean sistemáticas, un cambio en el proceso de recolección de la GEIH configura un choque exógeno que permite identificar su impacto. Esto quiere decir que, si todas las demás variables se mantienen constantes, los ingresos

reportados en las fuentes mantendrán una diferencia constante que pudo verse afectada por la recolección a través de CATI, debido a las características propias de este método. En ese sentido, la variación en la discrepancia constituye una estimación del efecto causal.

En este escenario una estrategia de *event study* es adecuada para inferir el impacto de interés. En particular, se utiliza la siguiente especificación:

$$\log(y_i^{GEIH}) - \log(y_i^{RELAB}) = \alpha_0 + X_i\beta + \sum_t \theta_t D_i T_t + \sum_t \sigma_t T_t + \sum_z \phi_z zona_i + \epsilon_i \quad (2.1)$$

Donde  $y_i^{GEIH}$  es el salario o ingreso laboral mensual de la actividad principal del individuo  $i$  reportado en la encuesta de hogares;  $y_i^{RELAB}$  es el ingreso calculado a partir de lo reportado en el RELAB. Así, la parte izquierda de la ecuación mide la discrepancia (en unidades logarítmicas) entre las fuentes de información.  $D_i$  es una variable indicativa que se activa para las personas residentes en las 23 áreas metropolitanas (por ejemplo, el grupo de tratamiento);  $T_t$  es una variable indicativa que se activa para el periodo de tiempo correspondiente a la implementación de métodos de recolección mixto. El efecto de interés estará medido por el parámetro  $\theta_t$  correspondiente al periodo marzo y julio de 2020. Es decir,  $\theta_t$  mide los cambios en la discrepancia en la variable de ingreso respecto a un periodo de referencia que se establece al inicio en el cambio de operativo.

Adicionalmente, se contrala por una batería amplia de variables de localización que se construye a partir de la interacción entre departamento y zona geográfica (área metropolitana, otras ciudades o zonas rurales), cuyo efecto se mide por  $\phi_z$ . Finalmente,  $X_i$  es un vector de controles que incluye: sexo, grupo de edad ( $< 25$ ,  $25 - 54$ ,  $> 55$ ), máximo nivel educativo alcanzado, número de horas trabajadas durante la última semana (al momento de la encuesta), tasa de desempleo <sup>8</sup>, sector de actividad del trabajador (CIUO revisión 4 a nivel de sector), clasificación de oficio del trabajador (CIUO a un dígito).

Esta especificación se estima para el periodo 2019 y 2020 tomando como referencia

---

<sup>8</sup>Para las áreas metropolitanas usamos la tasa de desempleo para cada área; para otras ciudades usamos la tasa de desempleo a nivel nacional para todas las otras ciudades; de forma similar, para las zonas rurales tomamos la tasa de desempleo rural a nivel nacional. Estos últimos indicadores no los tomamos a nivel de ciudad o municipio ya que la encuesta no tiene suficientes observaciones en estos dominios para hacer una buena medición de desempleo por municipio.

periodos trimestrales y como periodo de referencia enero y febrero de 2020. Marzo excluye dado que fue un periodo de adaptación de la encuesta. De hecho, el proceso de recolección se interrumpió en este mes. El periodo de aplicación de CATI (de abril a julio de 2020), se definió como un único periodo y mide el impacto del cambio en el método de recolección.

Cabe anotar que la integración entre GEIH y RELAB genera una muestra de trabajadores no representativa del total del mercado laboral, dado a que corresponde a trabajadores asalariados vinculados al sistema de seguridad social. Sin embargo, esta muestra tiene interpretación económica de importancia, en tanto se asocia al componente formal del empleo. A su vez, la muestra resulta representativa del total del RELAB ya que las características socioeconómicas de esta sub-muestra corresponden con el total del registro, como se evidencia en la Tabla 7 y Tabla 8 en el Anexo.

La validez de la estrategia de identificación propuesta recae sobre el supuesto de tendencias paralelas, que para este caso establece que entre los grupos de tratamiento y control, i.e., entre áreas metropolitanas y el resto, la discrepancia entre la respuesta del empleado en la encuesta y el reporte del empleador en el registro administrativo es constante antes de la intervención.

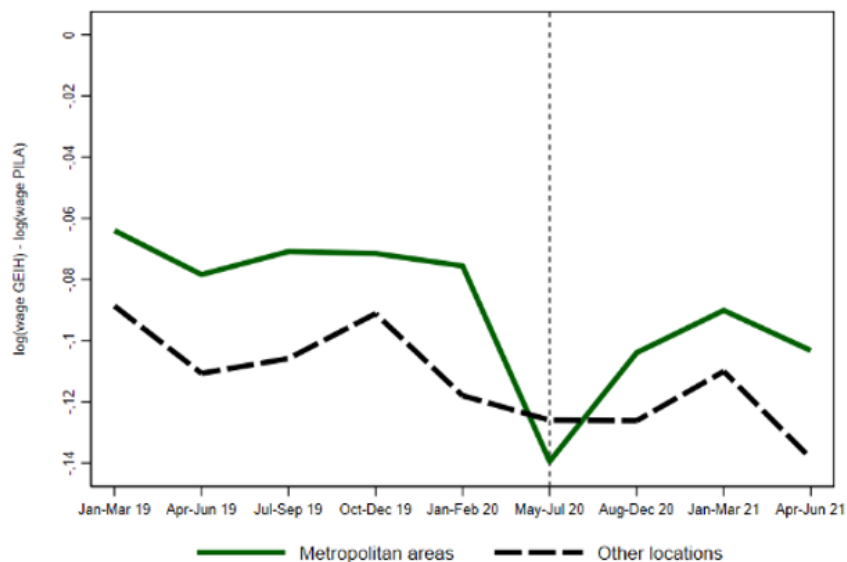
Para validar este supuesto se requiere que la evolución de la discrepancia en el ingreso sea similar, independiente de los posibles sesgos que contenga cada una de las fuentes de información y de los patrones regionales. Por otro lado, es posible que el impacto macroeconómico causado por el Covid-19 y las medidas asociadas sea mayor en las áreas metropolitanas. Esto no es un problema para la metodología propuesta, siempre y cuando ambas fuentes de información reflejen este impacto macroeconómico<sup>9</sup>. La Figura 1 presenta la evolución de la discrepancia promedio en los reportes.

La estrategia de identificación implementada no depende de la existencia de tendencias paralelas en salarios entre las áreas metropolitanas y el resto del país, lo cual podría ser un supuesto más exigente. Inclusive si las tendencias, antes de tratamiento, son paralelas, es posible que las tendencias a partir de entonces no lo sean, dado que las grandes ciudades estuvieron más expuestas a las infecciones de Covid-19 y a las respectivas medidas preventivas. Para dar evidencia de la robustez de resultados, se presentan estimaciones usando este supuesto alternativo de tendencias paralelas

---

<sup>9</sup>Alternativamente, si una de las fuentes de información refleja un mayor impacto de la pandemia que la otra, pero esta discrepancia es similar entre las áreas metropolitanas y en las otras ciudades y zonas rurales, el supuesto continúa corroborándose.

Figura 2.1: Promedio de la diferencia en logaritmos de salarios entre GEIH y RELAB, para áreas metropolitanas, otras ciudades y zonas rurales por mes 2019-2021



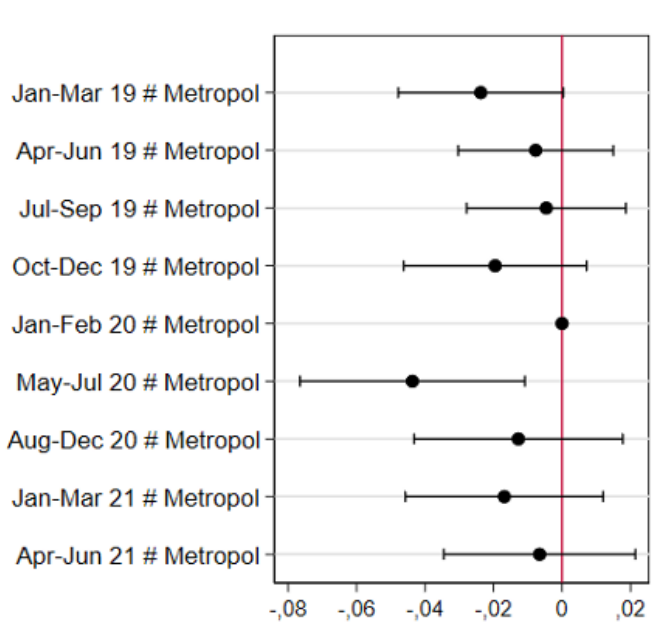
Fuente: Cálculos propios a partir de GEIH y RELAB.

## 2.4. Resultados

La Figura 2 y la Tabla 3 presentan los resultados de la estimación de la ecuación de *event study*. La Tabla 3 muestra el resultado de incluir secuencialmente controles y efectos fijos adicionales, de manera que en la especificación completa (columna 4) se observa que, un cambio en el patrón de la discrepancia en el ingreso obtenido en cada fuente equivale a 4.4 p.p. Para los demás periodos de tiempo, el efecto no es significativo a un nivel de significancia de 5%, lo que además implica que, una vez se restituyó el modo de recolección al CAPI, la discrepancia entre las fuentes retomó los niveles previos a los cambios temporales de recolección de la GEIH. Esto implica que la implementación del método de recolección CATI causó reducciones en el ingreso auto-reportado.

Estos resultados muestran el impacto agregado del método de recolección, sin em-

Figura 2.2: . Impacto estimado del método de recopilación de datos por método CATI en los salarios informados

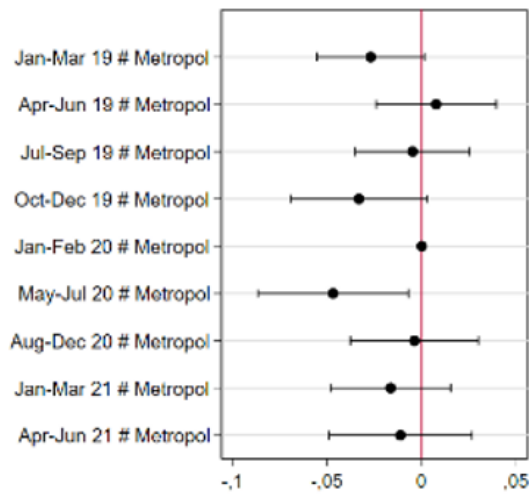


*Fuente:* Cálculos propios. Las líneas representan un intervalo de confianza al 95%.

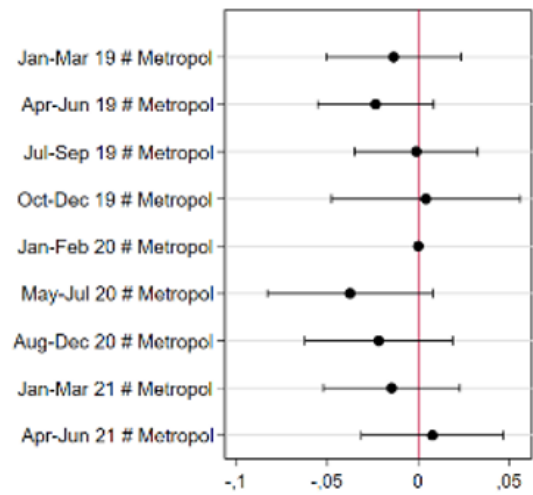
bargo, este cambio en el reporte puede variar de acuerdo con el perfil del entrevistado. De hecho, se ha documentado que el error de reporte de ingresos puede variar con respecto a las características sociodemográficas. En esta literatura se obtiene que los individuos más educados, las mujeres y los de mayor edad tienden a reportar los ingresos de manera más precisa. Esto implicaría que de la misma manera que se generan diferentes patrones en el reporte de ingresos cuando se utiliza el método CAPI, el cambio del modo de recolección podría acentuar dichos patrones. Por tanto, utilizando la misma estrategia de identificación, se exploran los cambios en los patrones de reporte por sexo, edad y nivel educativo.

Los resultados obtenidos, presentados en la Figura 3, muestran que los hombres, y los individuos mayores de 25 años tienden a tener mayores niveles de sub-reportaje de ingreso con la implementación del método CATI. Vale destacar que, en el caso de las mujeres, el impacto es de menor magnitud y no significativo. Por su parte, en el análisis por nivel educativo, se obtiene que el sub-reportaje es de magnitud similar entre los grupos analizados y resulta significativo en ambos casos. Estos resultados

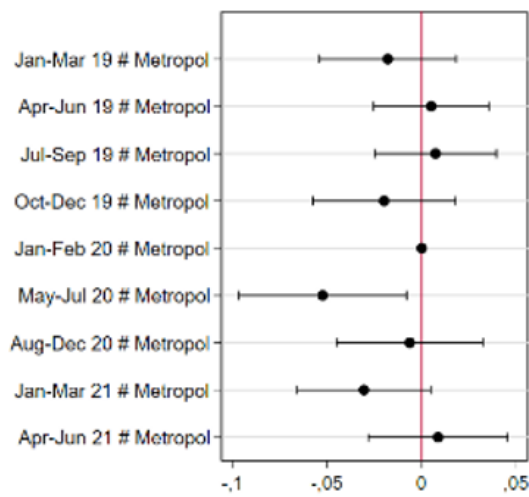
Figura 2.3: Impacto estimado por sexo y nivel educativo



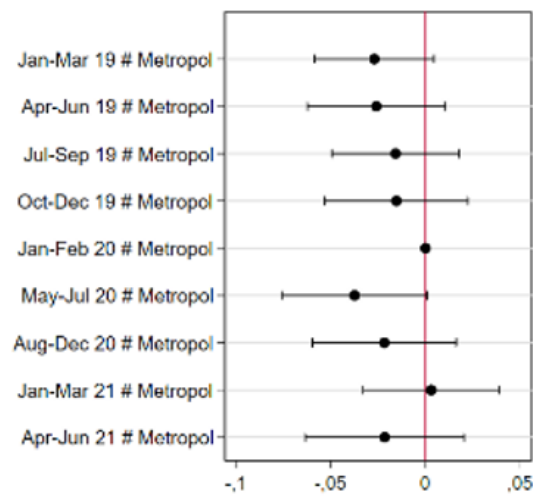
(a) Hombres



(b) Mujeres.



(c) Sin educación superior



(d) Con educación superior

Fuente: Cálculos propios. Las líneas representan un intervalo de confianza al 95%.

Cuadro 2.1: Tabla 3. Impacto estimado del método de recopilación de datos por método CATI en los salarios informados

	1	2	3	4
Ene-Mar 19#Metropol	-0.021 (0.013)	-0.024* (0.012)	-0.024* (0.012)	-0.024* (0.012)
Abr-Jun/19#Metropol	-0.012 (0.012)	-0.011 (0.012)	-0.009 (0.011)	-0.008 (0.011)
Jul-Sep/19#Metropol	-0.009 (0.013)	-0.009 (0.013)	-0.006 (0.012)	-0.005 (0.012)
Oct-Dic/19#Metropol	-0.026* (0.014)	-0.025* (0.013)	-0.022 (0.014)	-0.020 (0.013)
May-Jul/2020#Metropol	-0.055*** (0.015)	-0.054*** (0.015)	-0.052*** (0.015)	-0.044*** (0.016)
Ago-Dic/2020#Metropol	-0.020 (0.014)	-0.020 (0.014)	-0.017 (0.015)	-0.013 (0.015)
Ene-Mar/2021#Metropol	-0.022 (0.015)	-0.019 (0.014)	-0.017 (0.014)	-0.017 (0.014)
Abr-Jun/2021#Metropol	-0.007 (0.014)	-0.007 (0.014)	-0.006 (0.014)	-0.007 (0.014)
Observaciones	178,660	178,660	178,660	178,653
R2	0.004	0.020	0.024	0.044
Controles individuales	NO	SI	SI	SI
Mes FE	SI	SI	SI	SI
Actividad FE	NO	NO	NO	SI
Oficio FE	NO	NO	NO	SI
Dpto-clase FE	NO	NO	NO	SI

implican que, ante el cambio en el método de recolección, en general todos los grupos de población varían sus patrones de respuesta con un marcado patrón de edad.

Un análisis adicional por grupos de edad que se presenta en el Anexo a este documento, evidencia que en los individuos entre 25 y 54 años se tiene un impacto significativo y de similar magnitud ante el cambio en el método de recolección, mientras que para los otros dos grupos de edad este efecto no es significativo (ver Figura 4). Asimismo, por tipologías de hogar (de acuerdo con la presencia de hijos en el hogar), este efecto parece explicarse por el grupo sin hijos en el hogar (ver Figura 5). De esta forma se concluye que el impacto observado parece explicarse por hombres mayores de 25 años sin hijos en el hogar.

## 2.5. Chequeos de robustez

Con el objetivo de presentar evidencia adicional que valide los resultados encontrados, se presentan cuatro ejercicios adicionales que sirven como chequeo de robustez: i) en primer lugar, una prueba placebo en la que se comparan poblaciones cuyo método de encuesta para la GEIH no cambió; ii) una estimación en la que se restringe el conjunto de observaciones de control, iii) una estimación en la que se utilizó únicamente

la información proveniente de la encuesta de hogares y no el registro administrativo; y, finalmente, iv) una estimación que utiliza solo registros de RELAB en la que se espera no obtener ningún efecto, dada la estabilidad en esta fuente de información.

En la prueba placebo se retiraron de la muestra todos los trabajadores que viven en las 23 ciudades y AM, se tomaron los trabajadores del resto urbano como la población tratada y los trabajadores en áreas rurales como el control. De acuerdo con la expectativa, no se encontró ningún efecto significativo (ver Tabla 4, columna 1). Este ejercicio sirve como comprobación del supuesto de tendencias paralelas en los reportes de GEIH y RELAB.

Uno de los desafíos respecto a este supuesto es la coincidencia en el tiempo del operativo telefónico y las medidas de restricción a la movilidad, dado que el desarrollo de la pandemia pudo modificar las tendencias de reporte en las áreas más afectadas<sup>10</sup>, y este efecto pudo confundirse con el efecto de interés de este estudio. Sin embargo, la estimación del efecto placebo compara zonas más afectadas por la pandemia (otras ciudades respecto a zonas rurales) y no encuentra efectos significativos. Esto es consistente con la ausencia de cambios en la tendencia en el reporte.

Otra posible crítica al diseño metodológico implementado es que el conjunto de tratamiento, las 23 ciudades y AM, tiene un mercado laboral muy diferente al del conjunto de control, en particular respecto a las áreas rurales. Gran parte de estas diferencias están capturadas por el conjunto comprensivo de controles. Como comprobación adicional se presenta una estimación en la que se tomaron como control únicamente a los trabajadores en otras ciudades que deberían ser similares a los trabajadores en las áreas metropolitanas. La Tabla 4, columna 2, muestra los resultados de la estimación.

La magnitud del efecto estimado de esta forma es muy cercana a la encontrada en la estimación original de este estudio (4.5 p.p. frente a 4.3 p.p.), lo cual permite concluir que la inclusión de las áreas rurales en la estimación original no genera distorsiones significativas. Vale anotar que existe un efecto significativo, aunque pequeño, en el primer trimestre de 2019.

La columna 3 de la Tabla 4 muestra la estimación de la especificación event study

---

<sup>10</sup>Se debe tener en cuenta que para que este aspecto pueda presentar una dificultad para la estimación de este estudio no basta con que la pandemia haya cambiado las tendencias salariales; sería necesario que haya cambiado las tendencias en la discrepancia de reportes entre las diferentes fuentes de información. Es decir, que haya cambiado los incentivos de reporte de forma diferente para los trabajadores encuestados respecto a los empleadores que informan al registro administrativo, y que además este cambio de incentivos sea diferente en las áreas metropolitanas.

Cuadro 2.2: Tabla 4. Resultado ejercicios de robustez

	1	2	3	4
Ene-Mar 19#Metropol	-0.001 (0.027)	-0.025* (0.013)	0.013 (0.013)	0.031* (0.016)
Abr-Jun/19#Metropol	0.014 (0.025)	-0.014 (0.014)	-0.005 (0.018)	0.016 (0.017)
Jul-Sep/19#Metropol	0.004 (0.027)	-0.008 (0.014)	-0.002 (0.016)	-0.005 (0.019)
Oct-Dic/19#Metropol	-0.024 (0.030)	-0.016 (0.015)	0.019 (0.015)	0.025 (0.016)
May-Jul/2020#Metropol	0.005 (0.025)	-0.044** (0.019)	-0.037* (0.020)	-0.000 (0.017)
Ago-Dic/2020#Metropol	-0.010 (0.026)	-0.011 (0.019)	-0.010 (0.015)	0.008 (0.014)
Ene-Mar/2021#Metropol	-0.030 (0.026)	-0.007 (0.018)	0.013 (0.012)	0.032* (0.017)
Abr-Jun/2021#Metropol	0.034 (0.025)	-0.017 (0.015)	-0.005 (0.018)	0.001 (0.019)
Observaciones	21,998	171,763	327,932	178,641
R2	0.056	0.043	0.571	0.458
Controles individuales	SI	SI	SI	SI
Mes FE	SI	SI	SI	SI
Actividad FE	SI	SI	SI	SI
Oficio FE	SI	SI	SI	SI
Dpto-clase FE	SI	SI	SI	SI

usando solo datos de la GEIH para trabajadores asalariados:

$$\log(y_i^{GEIH}) = \alpha_0 + X_i\beta + \sum_t \theta_t D_i T_t + \sum_t \sigma_t T_t + \sum_z \phi_z zona_i + \epsilon_i \quad (2.2)$$

En este caso el contrafactual del cambio en el método de recolección se construye sobre el nivel del ingreso laboral, en lugar que la discrepancia en el ingreso, lo cual abre la posibilidad a confundir efectos de la pandemia de COVID-19 con los asociados a la implementación de CATI. Sin embargo, los parámetros estimados son similares a los de la estimación original (3.9 p.p. frente a 4.3 p.p.). Es importante resaltar que esta estimación reposa en supuestos diferentes, en particular tendencias paralelas en la evolución de los salarios.

A pesar de requerir supuestos más fuertes, esta estimación permite explorar otra posible objeción a los supuestos de la hipótesis original de este estudio: si los empleadores que informan al registro administrativo tienen incentivos para cambiar su estrategia de reporte durante el periodo comprendido entre marzo y julio de 2020, y estos cambios presentan patrones diferenciados en las áreas metropolitanas, la estimación principal sería la combinación del efecto descrito y el impacto del cambio en el operativo de recolección. Los resultados sugieren que la estimación de (2), sin embargo, no está afectada por el efecto de reporte de los empleadores.

Por último, el cuarto chequeo de robustez que solo utiliza información del RELAB

(La Tabla 4, columna 4) muestra un efecto no significativo de 5% de significancia, lo cual implica que los impactos encontrados en la comparación de la GEIH pueden imputarse al cambio en el método de recolección. Esto se da en la medida en que una fuente sin variación en este proceso no presenta evidencia coherente con los resultados encontrados, de manera que se puede descartar la posibilidad de efectos que generen sesgos adicionales.

La suma de los ejercicios de robustez dan validez al impacto encontrado y presentan evidencia sólida sobre el impacto de los cambios en el método de recolección sobre la variable ingreso, fundamental para análisis de mercado laboral y estimaciones de pobreza.

## 2.6. Conclusiones

Las encuestas de hogares son un instrumento con una gran capacidad para la recolección de información detallada de ingresos de los hogares, que además provee insumos de indicadores relevantes para las políticas públicas. Por tal motivo, estimar el ingreso de manera precisa y comparable en el tiempo es de especial importancia para los Institutos Nacionales de Estadística y los hacedores de política.

El método de recolección de la información es una de las posibles fuentes de sub-reporte y sobre-reporte, lo cual en el contexto del COVID-19 ha sido de particular interés por los cambios en los métodos de recolección de las encuestas de hogares, como consecuencia de las medidas implementadas para la contención de la pandemia.

El caso de Colombia presenta características particulares que constituyen particular para el estudio del impacto en el cambio de método de recolección. En particular, durante la pandemia, se implementaron métodos de recolección diferenciales por dominio geográfico, a lo cual se suma la disponibilidad de datos administrativos que sirven como referencia, debido a que esta fuente de información no se vio afectada por cambios operacionales en la recolección.

Utilizando un diseño cuasi-experimental se estima que el cambio en el método de recolección generó en promedio una reducción de 4% en el reporte ingreso laboral. Este nivel de sub-reporte es mayor para los hombres y los individuos entre 25 y 54 años. Este ejercicio permite cuantificar el impacto del método de recolección en un operativo a gran escala en un contexto no experimental, así como aprovechar la integración de datos administrativos con encuestas para generar aprendizajes que conlleven a di-

seños efectivos basados en métodos mixtos de recolección que garanticen la calidad de variables sensibles como el caso del ingreso.

## 2.7. Referencias

Abowd, J. M., & Stinson, M. H. (2013). Estimating measurement error in annual job earnings: A comparison of survey and administrative data. *Review of Economics and Statistics*, 95(5), 1451-1467.

Angel, S., Disslbacher, F., Humer, S., & Schnetzer, M. (2019). What did you really earn last year?: explaining measurement error in survey income data. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 182(4), 1411-1437.

Angel, S., Heuberger, R., & Lamei, N. (2018). Differences between household income from surveys and registers and how these affect the poverty headcount: Evidence from the Austrian SILC. *Social indicators research*, 138(2), 575-603.

Bell, A., Ward, P., Tamal, M. E. H., & Killilea, M. (2019). Assessing recall bias and measurement error in high-frequency social data collection for human-environment research. *Population and Environment*, 40(3), 325-345.

Bollinger, C. R. (1998). Measurement error in the current population survey: A nonparametric look. *Journal of Labor Economics*, 16(3), 576-594.

Bound, J., Brown, C., Duncan, G. J. and Rodgers, W. L. (1994) Evidence on the validity of cross-sectional and longitudinal labor market data. *J. Lab. Econ.*, 12, 345–368.

Bound, J., Brown, C., & Mathiowetz, N. (2001). Measurement error in survey data. In *Handbook of econometrics* (Vol. 5, pp. 3705-3843). Elsevier.

Bound, John, and Krueger, Alan B. “The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?” *Journal of Labor Economics* 9 (January 1991): 1–24.

Caeyers, B., Chalmers, N., & De Weerd, J. (2010). A Comparison of CAPI and PAPI through a Randomized Field Experiment.

Canberra Group. (2011). *Canberra Group Handbook on Household Income Statistics*. 2a. Geneva: United Nations.

Chesher and Schluter (2002. 2002. “Welfare Measurement and Measurement Error.” *Review of Economic Studies* 69, no. 2:357–78.

DANE (2021), Nota metodológica – Microdatos de ingresos para la medición de pobreza monetaria y desigualdad 2020, Bogotá D.C.

DANE (2020), Comunicado de Prensa – Acciones para garantizar la continuidad de la Gran Encuesta Integrada de Hogares (GEIH), Bogotá D.C.

De Leeuw, E. D. (1992). Data quality in mail, telephone and face to face surveys. TT Publikaties, Plantage Daklaan 40, 1018CN Amsterdam.

ECLAC (Economic Commission for Latin America and the Caribbean) (2020), “Recommendations for eliminating selection bias in household surveys during the coronavirus disease (COVID-19 pandemic)”, May 2020.

Ellis, C. H., & Krosnick, J. A. (1999). Comparing telephone and face to face surveys in terms of sample representativeness: a Meta-Analysis of Demographics Characteristics. Ann Arbor: Universidad de Michigan, NES.

ESRC Question Bank FACTSHEET 6 <https://www.ukdataservice.ac.uk/use-data/guides/methods-and-software-guides>

Fessler, P., Kasy, M., & Lindner, P. (2018). Survey mode effects on measured income inequality. *The Journal of Economic Inequality*, 16(4), 487-505.

Figari, F., Iacovou, M., Skew, A. J., & Sutherland, H. (2012). Approximations to the truth: Comparing survey and microsimulation approaches to measuring income for social indicators. *Social Indicators Research*, 105(3), 387–407. doi:10.1007/s11205-010-9775-4.

Gibson, J., & Kim, B. (2010). Non-classical measurement error in long-term retrospective recall surveys. *Oxford Bulletin of Economics and Statistics*, 72(5), 687–695.

Glewwe, P. (2007). Measurement error bias in estimates of income and income growth among the poor: analytical results and a correction formula. *Economic Development and Cultural Change*, 56(1), 163-189.

Gray, P.G., ”The Memory Factor in Social Surveys,” *Journal of the American Statistical Association*, 50 (June 1955), 344-63

Groves, R. M. (2004). *Survey errors and survey costs* (Vol. 536). John Wiley & Sons.

Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public opinion quarterly*, 67(1), 79-125.

Horowitz, J. and Manski, C. (1995). *Identification and Robustness with Contami-*

nated and Corrupted Data. *Econometrica*, 63, 281-302.

Jäckle, A., Roberts, C., Lynn, P.: Assessing the effect of data collection mode on measurement. *Int. Stat. Rev.* 78(1), 3–20 (2010)

Jäckle, A., Roberts, C., & Lynn, P. (2006). Telephone versus face-to-face interviewing: mode effects on data quality and likely causes: report on phase II of the ESS-Gallup mixed mode methodology project (No. 2006-41). ISER Working Paper Series.

Jordan, L. A., Marcus, A. C., & Reeder, L. G. (1980). Response styles in telephone and household interviewing: A field experiment. *Public Opinion Quarterly*, 44(2), 210-222.

Kamanou, G., Morduch, J., Isidoro, D. P., Gibson, J., Ivo, H., & Ward, M. (2005). Handbook on poverty statistics: Concepts, methods and policy use. The United Nations Statistics Division.

Kapteyn, A., & Ypma, J. Y. (2007). Measurement error and misclassification: A comparison of survey and administrative data. *Journal of Labor Economics*, 25(3), 513-551.

Kim, C. and Tamborini, C. R. (2014) Response error in earnings: an analysis of the survey of income and program participation matched with administrative data. *Sociol. Meth. Res.*, 43, 39–72.

Kreiner, C.T., Lassen, D.D., Leth-Petersen, S. (2013). Measuring the accuracy of survey responses using administrative register data. Evidence from Denmark. NBER Working Paper 19539.

Kuhn, U. (2019). Measurement of income in surveys. FORS Guide No. 02, Version 1.0. Lausanne: Swiss Centre of Expertise in the Social Sciences FORS. doi:10.24449/FG-2019-00002

Lohmann, H. (2011). Comparability of EU-SILC survey and register data: The relationship among employment, earnings and poverty. *Journal of European social policy*, 21(1), 37-54.

Lynn, P., & Kaminska, O. (2013). The impact of mobile phones on survey measurement error. *Public Opinion Quarterly*, 77(2), 586-605.

Meyer, B. D., & Mittag, N. (2021). Combining administrative and survey data to improve income measurement. *Administrative Records for Survey Methodology*, 297-322.

Meyer, B. D., Mok, W. K., & Sullivan, J. X. (2015). Household surveys in crisis.

Journal of Economic Perspectives, 29(4), 199-226.

Moore, J. C., Stinson, L. L., & Welniak, E. J. (2000). Income measurement error in surveys: A review. *Journal of Official Statistics-Stockholm-*, 16(4), 331-362.

NATIONS, U. (2012). Canberra Group Handbook on Household Income Statistics.

OIT 2020a Capturing impacts on employment and unpaid work using Rapid Surveys  
Guidance to data producers to maintain labour force survey data collection COVID-19:  
Guidance for labour statistics data collection

Ravallion, Martin. 1988. "Expected Poverty under Risk-Induced Welfare Vulnerability." *Economic Journal* 98, no. 3:1171-82.

Roberts, C., Eva, G., Allum, N., & Lynn, P. (2010). Data quality in telephone surveys and the effect of questionnaire length: A cross-national experiment (No. 2010-36). ISER Working Paper Series.

Sawada, Y., Nakata, H., & Tanaka, M. (2019). Short and Long Recall Errors in Retrospective Household Surveys: Evidence from a Developing Country. *The Journal of Development Studies*, 55(10), 2232-2253.

Schräpler, J. P., Schupp, J., & Wagner, G. G. (2010). Changing from PAPI to CAPI: Introducing CAPI in a longitudinal study. *Journal of Official Statistics*, 26(2), 239-269.

Sudman, S., & Bradburn, N. M. (1973). Effects of time and memory factors on response in surveys. *Journal of the American Statistical Association*, 68(344), 805-815.

UNSTATS 2020 Planning and Implementing Household Surveys Under COVID-19 van Praag, Bernard, Aldi Hagenaars, and Wim van Eck. 1983. "The Influence of Classification and Observation Errors on the Measurement of Income Inequality." *Econometrica* 51, no. 4:1093-1108.

Withey, S. B. (1954). Reliability of recall of income. *Public Opinion Quarterly*, 18(2), 197-204.

## Capítulo 3

# Competencia a través de plataformas facilitadoras en el mercado de alojamiento: la relación de la diferenciación y la reputación con los beneficios del mercado hotelero en Bogotá-Colombia

### 3.1. Introducción

Las plataformas de agregación de oferta de bienes y servicios por internet han modificado de forma radical el funcionamiento de muchos sectores de la economía, y en particular el sector de alojamiento. Por ejemplo, en Europa la mayor parte de las ventas del sector se realizaron en sitios web o a través de Agencias de Viajes Online (OTAs, por sus siglas en inglés) (Mantovani et al., 2017). Estas plataformas reducen el costo de búsqueda que enfrentan los consumidores a la hora de informarse sobre los precios y las características de los productos propuestos por distintos oferentes y aprovechan los efectos de red entre los distintos agentes para asegurar su posición en el mercado como intermediadores entre la demanda y la oferta. El presente estudio busca identificar la presencia del efecto de externalidad de red asociado a la participación de

establecimientos de alojamiento u hoteles en las plataformas. Para este fin se estudiará la correlación entre los beneficios de los hoteles y las características de sus pares. Sin embargo, no se debe asumir que el efecto va a ser el mismo para cualquier par de hoteles, debido a que los hoteles compiten bajo una lógica de diferenciación. Es decir, además de competir en precios, es de esperar que los establecimientos de alojamiento que atienden segmentos de demanda similares tengan mayores influencias entre sí.

El mercado hotelero tiene características particulares que hacen que estas plataformas tengan un efecto de especial relevancia. En primer lugar, por la naturaleza misma del servicio, los consumidores no tienen contacto constante y cercano con los oferentes, lo cual genera que, en ausencia de estas plataformas, los costos de búsqueda sean muy altos. En segundo lugar, que se tratan de bienes experiencia, es decir, que su calidad sólo puede ser observable o valorada durante o posterior a su consumo (Nelson, 1970, 1974; Dubois & Nauges, 2009).

Si bien los oferentes tienen alguna capacidad de informar a los consumidores sobre estas características a través de la publicidad (o de las mismas plataformas), hay un problema de riesgo moral, en la medida en que los establecimientos hoteleros pueden tener incentivos para reportar información incorrecta. Los esquemas de los sistemas de calificaciones y recomendaciones de los sitios agregadores puede funcionar entonces como mecanismo de verificación de la información provista por el oferente. Como lo indica Bickart & Schindler (2001), los sistemas de puntajes, en donde se presentan las experiencias pasadas de otros consumidores, se han convertido en un elemento crucial dentro del conjunto de variables de decisión de los individuos a la hora de comprar, reduciendo incluso la importancia de la información proporcionada por los proveedores.

En ese sentido, De Los Santos et al. (2017), Chen et al. (2017) y Jiang & Zou (2020) han comprobado que las disminuciones en los costos de búsqueda no solo mejoran el bienestar de los consumidores, sino también de otros actores de la industria. Al aumentar el uso de las plataformas, aumentan los beneficios que ésta percibe, e inclusive es posible que para los oferentes el beneficio de participar en un mercado más grande contrarreste las desventajas del aumento de la intensidad de la competencia.

La información incompleta no se encuentra solo del lado de los consumidores; entre los oferentes también puede haber asimetrías de información. Si la plataforma permite a los hoteles conocer los precios y las características de los productos de sus competidores, como, por ejemplo, a través de comentarios y calificaciones de usuarios anteriores, los competidores modificarán sus estrategias de fijación de precios y de selección de

productos, impactando también los beneficios. No obstante, este efecto de red puede ser difícil de aislar de otras variables del hotel y del mercado, como los rendimientos, especialmente en una industria que, como se comentó anteriormente, se caracteriza por tener particularidades que no son fáciles de observar, no solo para sus participantes, sino también para el econometrista. En cuanto al método, es necesario encontrar una herramienta que permita establecer correlaciones que dependan de una medida de diferenciación. Con este fin se ha adoptado herramientas de econometría espacial, en particular los modelos de auto-correlación espacial Durbin. En lugar de encontrar una correlación que disminuye con la distancia espacial, en este artículo se encuentra una correlación que se reduce a medida que aumenta la distancia en un espacio de características de los productos. De esta forma se puede establecer la correlación entre los beneficios de un hotel y los de sus pares, ponderados por esta distancia cualitativa, y también es posible determinar el efecto que el puntaje promedio de los pares, ponderado por diferenciación, tiene sobre los beneficios del establecimiento.

En materia de datos, para poder aplicar esta metodología se requiere crear una medida de distancia en el espacio particular a los servicios hoteleros. Dado que este tipo de caracterizaciones detalladas no suelen ser provistas en las fuentes tradicionales de datos de sector de alojamiento, fue necesario recurrir a fuentes no tradicionales, en particular web scraping y text mining, en dos de los sitios web facilitadores para la venta de habitaciones de hotel más usados en Bogotá, Colombia. Esto permitió recoger información sobre las características de los servicios ofrecidos por el hotel a partir de las descripciones que los mismos oferentes publican en la plataforma. Con estos insumos fue posible construir una matriz de diferenciación por atributos que permite medir la distancia entre establecimientos en términos de los servicios ofrecidos.

Esta información de Big Data fue integrada con datos provenientes Registros Administrativos y fuentes tradicionales. Por un lado, los datos sobre beneficios, precios y ocupación provienen de la Encuesta Mensual de Alojamiento del Departamento Administrativo Nacional de Estadística de Colombia (DANE). Por el otro, la información de control sobre ubicación geográfica de los establecimientos se calculó a partir de las direcciones incluidas en el Registro Nacional Hotelero, y la posterior georreferenciación con la API de Google. Por lo tanto, este estudio implementa y aprovecha la integración de fuentes tradicionales con fuentes alternativas de información para el estudio de los efectos de red en las plataformas de intermediación hotelera y de alojamiento. De igual modo, este estudio da cuenta de la creciente influencia de internet y nuevos actores,

como las plataformas en línea, sobre la toma de decisiones de los consumidores y oferentes, con el fin de servir de utilidad para la formulación de estrategias de mercado efectivas.

A partir de lo anterior, las estimaciones de este artículo se realizan sobre tres variables de resultado: la tasa de ocupación, la rentabilidad y el precio. Se encuentra una correlación espacial en la que el resultado promedio de los competidores en estas tres variables afecta negativamente el resultado propio. Por otro lado, los aumentos en las calificaciones promedio de los competidores tienen una correlación positiva con los precios y la rentabilidad de otro establecimiento de alojamiento.

El presente documento se divide en cinco secciones: la primera de ellas es la presente introducción, seguido a esto se presentarán los estudios los recientes que se han realizado en la materia y que sirven como marco teórico para el desarrollo conceptual de este documento. El tercer capítulo detalla la construcción del conjunto de datos y la metodología. En cuarto lugar, se presentarán los resultados y finalmente se dará cierre al documento con las conclusiones.

## **3.2. Plataformas y Diferenciación**

La literatura ha documentado que los bienes de experiencia son aquellos cuya calidad es incierta antes del consumo; la disponibilidad de información adicional, bien sea provista por el mismo vendedor en forma de publicidad o por otros participantes del mercado, van a afectar entonces las decisiones de consumo (Nelson, 1970, 1974; Dubois & Nauges, 2009). De este modo, el proveedor del bien o servicio elabora su conjunto de expectativas sobre los bienes o servicios e influye en la decisión de los individuos de comprarlo.

Como consecuencia de las relaciones con múltiples actores de la industria turística, como hoteles, proveedores, distribuidores, competidores, gobiernos y otras empresas, se evidencian una multiplicidad de externalidades potenciales que generan interdependencia en sector turístico (González-Torres, Rodríguez-Sánchez & Pelechano-Barahona, 2021). Shy (2010) se enfoca en los efectos externos entre grupos, definidos como un tipo especial de externalidad en la que la utilidad de los consumidores y / o las ganancias de las empresas se ven directamente afectadas por el número de consumidores y / o productores que utilizan la misma tecnología, y señalan que los efectos de red dependen de la popularidad de las plataformas. De igual modo, se observa que la competencia y

los incentivos de diferenciarse aumentan con el número de participantes.

Estas externalidades han sido estudiadas con diferentes enfoques, en los que el nivel de información de cada uno de los actores juega un rol central en las decisiones de sus pares (Hagiu & Halaburda, 2014) y en el que, dependiendo del nivel de información, se determina el precio. De hecho, Hagiu, A. (2009) ha mostrado cómo, a través de la competencia entre productores y plataformas, se pueden llegar a efectos de reducción de los precios hacia los consumidores.

Además de esto, se ha buscado establecer el modo en el que los efectos de la reputación o divulgación de información a través de plataformas pueden afectar el bienestar de los consumidores (Forman et al., 2008). En el caso de las plataformas facilitadoras de reserva de habitaciones, las ganancias de los hoteles existentes disminuyen por el efecto competencia, pero al entrar más hoteles al mercado incrementan las opciones de los consumidores (Audretsch & Feldman, 1996; Feldman & Audretsch, 1999; Dürr & Jaffe, 2012). Por otra parte, Nunkoo et al. (2020) han documentado que el impacto de la divulgación de información acerca de experiencias pasadas impacta de forma diferenciada a los segmentos de hoteles clasificados por estrellas.

Este impacto, puede materializarse de forma positiva o negativa (Peitz et al, 2019). Para la literatura de los mercados bilaterales, lo negativo surge cuando los miembros de un grupo compiten entre sí para interactuar con el otro grupo.

Por otro lado, podemos ver efectos positivos dentro del grupo, por ejemplo, si un determinado huésped que visitó los hoteles en el pasado, vuelve a visitarlos. Esto permite a los potenciales huéspedes reducir las asimetrías de información (Davis, 2021) y disminuir los costos de búsqueda (De Los Santos et al. (2017); Chen et al. (2017). y Jiang and Zou (2020)).

El desarrollo de este tema se estudió ampliamente en el e-Word-On-Mouth (eWOM) que se define como “cualquier declaración positiva o negativa hecha por clientes potenciales, reales o anteriores sobre un producto o empresa, que se pone a disposición de una multitud de personas e instituciones a través de Internet (Hennig-Thurau et al., 2004), donde estas calificaciones se utilizan como divulgación de terceros y como indicadores de calidad para muchos de los usuarios, así como como insumos para contar con recomendaciones de expertos.

El eWOM en la industria hotelera ha atraído una gran atención entre los profesionales del sector del alojamiento y algunos académicos (Filieri, R., & McLeay, F. (2014) y Reyes-Menendez, A., Saura, J. R., & Martinez-Navalon, J. G. (2019)). En la mayoría

de los trabajos se muestra que el desempeño de los hoteles se ve afectado por cambios en el eWOM. Sin embargo, artículos como (Peitz, 2019) sugieren que el efecto no está bien identificado, ya que depende de la especificación del eWOM, que a su vez no sólo se basa en una valoración específica, sino también en la intensidad o volumen de la valoración.

De igual modo, el aumento de las ventas a través de plataformas facilitadoras en línea ha introducido muchos retos, debido a la naturaleza del bien que allí se ofrece. En primer lugar, existe un reto relacionado con la regulación de precios causado por el beneficio de intermediación que tienen las plataformas (Mantovani, 2017). Por otro lado, en materia de política de competencia, se tienen retos relacionados con la definición del mercado relevante y en muchos casos la caracterización de las externalidades que allí se presentan. Lo anterior, dada la multidimensionalidad de atributos que tienen los establecimientos de hospedaje por su, se enfrentan a retos en materia de regulación.

### **3.2.1. Aglomeración y competencia hotelera**

Una de las características más importantes al estudiar la industria hotelera, incluyendo precios, ingresos, rentabilidad, ocupación, entre otros, es su ubicación espacial. Al respecto, las investigaciones han considerado que ignorar las regulaciones de zonificación espacial cuando se estiman modelos de entrada y ubicación de empresas causa una seria subestimación del impacto en las características del mercado, tales como la población y los ingresos potenciales asociados (Datta & Sudhir, 2013), por lo cual conduce a sesgos en la medición de la verdadera intensidad de la competencia. En ese sentido, es necesario tener en cuenta que cada ciudad tiene un ordenamiento espacial que relaciona la aglomeración de hoteles en zonas específicas y la competencia entre ellos.

En esta misma línea, Datta & Sudhir (2013) encuentran que la zonificación reduce la entrada de nuevas empresas porque la incapacidad de diferenciarse espacialmente aumenta la competencia, reduciendo la rentabilidad y el número de empresas que un mercado puede soportar. Además, para altas restricciones de zonificación, que limitan la capacidad de diferenciarse espacialmente, es posible que no haya cambios en el número de empresas en el mercado, sino solo cambios en la combinación de formatos minoristas de productos (Datta & Sudhir, 2013).

Paralelamente existe un trade-off entre competencia y los efectos de la aglomeración, dado que la proximidad está estrechamente asociada con la intensidad en la

competencia (Tsang & Yip, 2009). De hecho, uno de los principios de la competencia espacial es que las empresas deben ubicarse lejos entre sí para explotar los monopolios locales, dado que cuando las empresas se ubican más cerca unas de otras compiten más intensamente y por lo tanto obtienen menores ganancias (Tsang & Yip, 2009). Sin embargo, la teoría de la aglomeración sostiene que la presencia de competidores en una misma zona suele ser beneficiosa (Marshall, 1920), dado que la proximidad física genera economías de escala por aglomeración que pueden ser lo suficientemente grandes para compensar los efectos negativos de la competencia.

Los beneficios pueden ser de dos tipos: mejoras en la producción (por encadenamientos productivos y economías de escala) y aumento en la demanda (por reducción en los costos de búsqueda de clientes, y mayor frecuencia de visitas de potenciales compradores). A esto deben ser añadidas las ventajas de la infraestructura y otros bienes y servicios alrededor de los hoteles, como lo son restaurantes, bares, atracciones, cines, entre otros, que tienen un impacto sobre los potenciales clientes. En este sentido, la proximidad física entre hoteles favorece a los consumidores con costos de búsqueda más bajos, ayudándoles a evaluar los productos y servicios, y aumentando la probabilidad de visitas y compras, como lo demuestran las aglomeraciones geográficas de teatros, restaurantes, hoteles y tiendas especializadas (Tsang & Yip, 2009).

En esta línea, Baum & Haveman (1997) analizaron las decisiones multidimensionales en las elecciones de ubicación de la industria hotelera de Manhattan desde dos perspectivas distintas que determinan la demanda y la competencia a la que se enfrentan: la elección del producto y la ubicación geográfica de los hoteles. Este estudio encontró evidencia de que los hoteleros ubican nuevos alojamientos lo suficientemente cerca de hoteles ya establecidos que son similares en una dimensión de producto, como el precio, para beneficiarse de economías de aglomeración, pero diferenciándose en otra dimensión del producto, con el fin de evitar la competencia localizada y crear diferencias complementarias (Baum & Haveman, 1997).

Por su parte, Tsang & Yip (2009) y Voltes-Dorta & Inchausti-Sintes (2021) encontraron evidencia de que solo los hoteles de empresas conjuntas de alto rango de estrellas contribuyen a una mayor demanda, mientras que los hoteles de todos los rangos de estrellas se benefician de manera similar de la aglomeración. Así, cuantos más hoteles de alto rango de estrellas estén en una ubicación, mayores serán los beneficios de la aglomeración. Finalmente, este mismo estudio encontró que el aumento de la demanda debería incrementar los ingresos del hotel, como consecuencia de una mayor tasa de

ocupación, un aumento en el precio, o una combinación de ambos.

### **3.2.2. Diferenciación por atributos en el mercado hotelero**

Además de la ubicación física, la proximidad se puede medir en términos de otros dominios, como la calidad y el precio del producto o servicio. Relacionado a esto, la teoría económica predice que las limitaciones de la diferenciación espacial conducirá a una mayor diferenciación de productos (Datta & Sudhir, 2013). Por ejemplo, Kuksov (2004) argumenta que las empresas responderán a la incapacidad de diferenciarse espacialmente en Internet con una mayor diferenciación de productos, a lo cual Bar-Isaac y col. (2009) explican que esto se debe a un efecto de una mayor variedad de productos en los mercados minoristas de internet.

En este sentido, los hoteles buscan diferenciarse mediante otros atributos, para lo cual las empresas hacen uso de otros medios, como la publicidad, para informar a los consumidores de sus características (Milgrom & Roberts, 1986) e informar a los consumidores potenciales sobre la probabilidad de satisfacer o no sus preferencias. Nelson (1970) sostiene que la publicidad es clave para la diferenciación, ya que indica cualidades no observables de un producto o servicio, lo que es especialmente importante para vender bienes de experiencia. Por ejemplo, en el caso de Airbnb, (Voltes-Dorta & Inchausti-Sintes, 2021) encuentran que características diferenciadoras, como la cantidad de dormitorios y baños, así como su proximidad a las atracciones turísticas es un factor clave en el nivel del precio y, por tanto, la rentabilidad.

Es así como los consumidores utilizan con frecuencia sistemas de clasificación de hoteles, como las plataformas, para elegir entre las alternativas de alojamiento, dado que el alojamiento es un bien de experiencia, cuya valoración de calidad se basa en la existencia de información disponible.

## **3.3. Metodología y datos**

### **3.3.1. Metodología**

El propósito de la metodología que se va a plantear es encontrar una manera de estudiar la correlación entre el desempeño de establecimientos hoteleros y algunas características de sus pares en el mismo mercado. Adicionalmente, queremos que esta correlación tenga en cuenta la distancia que existe entre las firmas en el espacio de las

características de los servicios, es decir, que incorpore el hecho de que las interacciones entre firmas serían más probables y más intensas si los productos que ofrecen son parecidos.

Una forma en la que podemos cumplir con estos requisitos metodológicos es mediante la adaptación de metodologías de econometría espacial. En estos modelos de estimación permiten identificar externalidades generadas por la interacción entre unidades. Un elemento distintivo de estos ejercicios es que comúnmente tienen en cuenta que la existencia o intensidad de interacción depende de la distancia geográfica (Anselin 2001). De esta forma, se puede estudiar cuál es la correlación entre una variable a nivel de unidad y una variable que caracterice el promedio ponderado por distancia de los vecinos. Sin embargo, en el ejercicio que se va a proponer a continuación se va a usar una distancia cualitativa, en vez de una distancia geográfica.

La especificación del modelo captura la interdependencia con los pares a través de la inclusión de variables espacialmente rezagadas, y estas relaciones son caracterizadas con una matriz de proximidad. La dependencia entre pares puede presentarse a través de: i) la variable dependiente; ii) las variables explicativas; y/o iii) el término de error (ver Anselin, 1988,2006; Baltagi et al., 2003, 2007; Elhorst, 2003, 2014; Kapoor et al., 2007; Yu et al., 2007).

Dado que el objetivo de este trabajo es identificar el efecto que tiene la reputación y los resultados de los hoteles más cercanos sobre la rentabilidad, se escoge la especificación Durbin, en el cual la interdependencia está determinada por la variable dependiente y la media ponderada por distancia de un conjunto de variables independientes (Elhorst, 2014). Formalmente se representan con la siguiente ecuación:

$$y_i = \alpha + \rho W y_i + \gamma W z_i + \delta z_i + \beta x_i + u_i \quad (3.1)$$

Donde  $y_i$  es la variable dependiente;  $x_i$  es conjunto de variables explicativas, correspondientes a las características propias del establecimiento de alojamiento;  $z_i$  es el puntaje que recibe el establecimiento en la plataforma;  $W$  es una matriz de nxn de ponderaciones construida a partir de las distancias en atributos. Por lo tanto,  $\rho$  captura el efecto spillover directo en la variable de resultados del hotel,  $\delta$  mide el efecto de la calificación sobre el mismo hotel, y  $\gamma$  mide el efecto cruzado de la calificación de los competidores cercanos.

Un aspecto importante a considerar es el de los métodos de estimación apropiados para solventar la dificultad generada por la simultaneidad que se deriva de la depen-

dencia espacial. Los dos procesos de estimación dominantes en econometría espacial son el de máxima verosimilitud (ML) (Ord 1975) y el de método general de momentos (Anselin, 1980 ; Kelejian y Prucha 1998 , 1999). La estrategia de estimación utilizada es la segunda, que es computacionalmente es más eficiente. Ésta consiste en utilizar un método de mínimos cuadrados por dos etapas, usando como instrumentos las distancias de órdenes superiores ( $W^n y_i$ ) y la interacción de las mismas con las variables independientes ( $W^n x_i$ ). De acuerdo con (Kelejian et al. 1999, 2007,2010) los estimadores son asintóticamente eficientes, siempre que los instrumentos que se utilicen no tengan ninguna correlación con la variable dependiente.

La primera etapa corresponde a

$$\begin{aligned} W y_i &= \theta_1 q_i + v_i^1 \\ W z_i &= \theta_2 p_i + v_i^2 \end{aligned}$$

Donde  $p_i$  y  $q_i$  representan los conjuntos de instrumentos. Como lo sugiere Kelejian et.al 1999,  $p_i$  está compuesto por variables independientes que solo presentan relación con  $y_i$  a través de  $W y_i$ . Por otra parte,  $q_i$  está compuesto por variables independientes que solo presentan relación con  $z_i$  a través de  $W z_i$ . Posteriormente con el resultado de las primeras etapas se estiman los parámetros de interés.

$$y_i = \alpha + \rho \widehat{W y_i} + \gamma \widehat{W z_i} + \delta z_i + \beta x_i + u_i \quad (3.2)$$

Los instrumentos utilizados son las relaciones de orden superior, esto es el efecto de los vecinos de mis vecinos (la interpretación geométrica para matrices del inverso de la distancia es el promedio espacialmente ponderado del promedio espacialmente ponderado de mis vecinos – esto es equivalente al promedio de mis vecinos cuando la distancia se toma como el promedio de las rutas que pasan por otros vecinos). El supuesto de relevancia del instrumento se cumple dado que exista el efecto espacial. El supuesto de exclusion se cumple siempre y cuando las relaciones de orden superior no generen spillovers que no sean capturados por las relaciones de primer orden.

### 3.3.2. Datos

Este trabajo aprovecha información disponible en Registros Administrativos y Encuestas, e información extraída a partir de la técnica web-scraping. Además, mediante la Application Programming Interface [API] de geolocalización de Google la API de Google es posible utilizar datos espaciales que permitan determinar distancias geográficas entre establecimientos. Estas fuentes alternativas de datos constituyen una contribución principal y poco explotada en la literatura reciente debido a la utilización de distancias en el análisis estadístico. Además de contribuir a la literatura reciente en torno a las innovaciones metodológicas respecto a la recopilación de fuentes alternativas de datos y la integración con las fuentes tradicionales. Lo cual permite un mejor acercamiento de las empresas y la política pública para la toma de decisiones desde información estadística oportuna y relevante.

Las herramientas de web-scraping han sido utilizadas cada vez más en la producción de estadísticas y en investigación pues permite elaborar indicadores que previamente se consideraban por fuera del alcance de los institutos nacionales de estadísticas (INE). Esta técnica permite recuperar información de páginas web y estructurarla en una base de datos. Estudios previos han abordado los impactos de la pandemia de COVID-19 a partir de análisis de creencias del riesgo de viajar de los turistas (Neuburger & Eggber, 2021). De manera complementaria, se han utilizado técnicas de minería de texto para analizar las emociones y preocupaciones de los empleados del sector de alojamiento (Park, 2020; Kim & Kim, 2020; Güliz & Akbıyık. 2020).

Por otra parte, EUROSTAT (2017) plantea la construcción de un ecosistema estadístico en el largo plazo es que las nuevas fuentes de datos sean integrados y armonizados con encuestas. Estas encuestas continuarán aportando información que no es posible conseguir de otra manera como es el motivo de viaje, los medios de transporte y las características sociodemográficas. A continuación, se encuentra una descripción de los datos utilización. Las siguientes subsecciones muestran cómo se construyó la base de datos utilizada en el presente estudio, así como las metodologías de cálculo de las variables de interés.

Es por lo anterior que los autores del presente estudio diseñaron un conjunto integrado de datos, que incluye variables de desempeño como el porcentaje ocupación hotelera, camas disponibles y precio, variables de ubicación, variables de oferta de servicios e información relacionada con los puntajes otorgados por los diferentes clientes que han visitado los hoteles que será utilizada como proxy de calidad no observada.

Esta integración se realiza entre el Registro Nacional de Turismo, la Encuesta Mensual de Alojamiento y un conjunto de datos construido con herramientas de webscraping de facilitadores de búsquedas web para el sector de alojamiento.

En primer lugar, se utiliza la de información administrativa del Registro Nacional de Turismo y la API de geolocalización de Google, fue posible obtener los valores correspondientes a las coordenadas de ubicación del 98 % de los 823 establecimientos de alojamiento para la ciudad de Bogotá incluidos en este registro con corte al primer trimestre del 2021. En este procedimiento se realizaron búsquedas sobre la API a partir de una variable concatenada de nombres comerciales estandarizados y el código BOG.

En segundo lugar, esta base de datos fue integrada por nombres estandarizados con una extracción de información que utilizo herramientas de web scraping sobre facilitadores web de compra de servicios de alojamiento, dónde se realizó la búsqueda con la concatenación del país, ciudad y nombre del hotel, para evitar obtener resultados erróneos al encontrar hoteles homónimos en otra región u otro país.

El rastreador web seleccionó el nombre comercial de los establecimientos incluidos en la encuesta y buscó obtener una descripción detallada de los servicios ofrecidos por los hoteles, cómo tener o no parques infantiles, centros de convenciones, zonas húmedas, estrellas, calificación realizada por usuarios, nombre presentado en la página, y servicios prestados por el hotel (por ejemplo, servicio de piscina, niñera, peluquería, entre otros). Al finalizar la extracción, se efectuó una revisión manual comparando el nombre buscado con el nombre resultado de la búsqueda, corroborando que la información fuera efectivamente la correspondiente al hotel de interés.

Finalmente utilizando el Número de Identificación Tributaria (NIT) se vinculó con la Encuesta Mensual de Alojamiento EMA, producida por el Departamento Administrativo Nacional de Estadística de Colombia (DANE), que monitorea las principales variables de resultados, como la ocupación hotelera y los ingresos para cada mes y región. Esta encuesta también permite obtener atributos específicos de la demanda, como la proporción huéspedes no residentes, así como el motivo del viaje de los visitantes a los hoteles. En este cruce se logró vincular 160 de los 162 establecimientos incluidos en la EMA en el periodo 2019-2021.

## **Matriz de proximidad**

La matriz W (Tabla 3.1.) fue construida a partir del conjunto de atributos de cada establecimiento de alojamiento a través de las plataformas facilitadoras ([www.trivavago.com](http://www.trivavago.com))

y [www.booking.com](http://www.booking.com)), donde cada elemento  $g_{i,j}$  incluido en la matriz  $W$  corresponde con la distancia de Gower entre cada  $i, j$ . Cada elemento fue expresado como un porcentaje del total de las distancias de todos los hoteles en cada una de las regiones, en donde, en línea con la literatura encontrada, de acuerdo con la literatura, cada fila fue estandarizada con el fin de que la sumatoria de cada una fuese la unidad. Los atributos tenidos en cuenta son:

Cuadro 3.1: Atributos de los servicios del hotel y de las habitaciones

Servicios del Hotel		Servicios de la habitación
Parqueadero	Asador en el exterior	Aire acondicionado
Restaurante	Bar	Bañera
Sauna o turco	Botones	Escritorio
Servicio a la habitación	Centro de negocios	Habitación para discapacitados
Sombrillas	Conserje	Nevera
Tienda o Butique	Gimnasio	Televisión satelital
Trasnporte al aeropuerto	Pago sin efectivo	
Viajeros con niños	Peluquería	
Wifi gartis	Piscina	

*Fuente:* Elaboración propia

## Variables de resultado

Las decisiones estratégicas de los dueños de los establecimientos son, ante todo, fijar un precio y número de habitaciones a ofrecer. Estas decisiones pueden depender de variables resultado de la competencia dentro de un mismo segmento o entre los mismos.

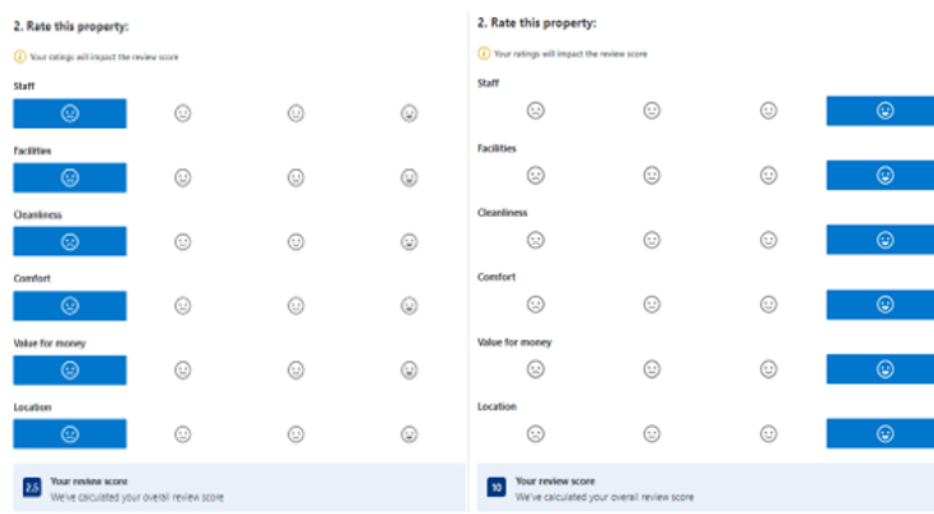
A través de la EMA se obtienen varias medidas de resultados económicos. Primero, se calcula la rentabilidad de los hoteles medida a través de la razón entre el nivel de ingresos y el número de empleados de los hoteles. En segundo lugar, se observan las habitaciones vendidas, que, junto con las habitaciones disponibles, permiten calcular la ocupación y la cuota de mercado de cada uno de los hoteles.

## Características individuales

### a. Puntaje

Un rasgo característico de muchos sistemas de reputación online es que la mayoría de los usuarios dan buenas críticas y, como resultado, hay poca variación en los datos (Nosko y Tadelis (2015), Dellarocas & Wood (2008)). Este también es el caso de [www.Booking.com](http://www.Booking.com), en donde la figura 3.1 evidencia la distribución de la reputación media de los hoteles. Las calificaciones bajas son raras y la gran mayoría de las calificaciones se encuentran dentro del rango de 2.5-10, es decir, entre muy bueno y perfecto.

Figura 3.1



*Fuente:* [www.Booking.com](http://www.Booking.com).

### b. Estrellas

Los hoteles con una clasificación de estrellas más alta tienen más probabilidades de publicitarse que sus contrapartes con una clasificación de estrellas más baja (Tsang & Yip, 2009). Además, los hoteles con un rango de estrellas alto reciben viajeros con mayor poder adquisitivo, y, para captar a estos viajeros, es probable que haya más tiendas y restaurantes.

Entre otras estrategias de mercado se encuentran las alianzas internacionales o cadenas de hoteles, que generalmente tienen un mayor goodwill, o una mejor infraestructura física circundante producto de la aglomeración. Entre los hoteles de alto rango de estrellas, los hoteles de empresas conjuntas aportarán más beneficios de aglomeración que los hoteles que no son empresas conjuntas.

### c. Ubicación geográfica

La distancia promedio en kilómetros de cada hotel se construye a partir del promedio de su distancia, con respecto a la de cada uno de sus competidores. Esta se realiza mediante la distancia euclidiana en kilómetros utilizando cada par de coordenadas de los establecimientos. Estos geo-códigos se rescatan con la del api de Google, mediante consultas de la latitud y longitud a través de la Geocoding API pertenecientes al entorno de Google Cloud Platform.

Figura 3.2



*Fuente:* Elaboración propia, basado en datos del Registro Nacional de Turismo 2020 y geolocalización de Google cloud platform.

### 3.4. Resultados

La tabla 4.1 muestra la estimación del modelo base (ecuación 1) en el que se reporta el efecto de las calificaciones promedio de los vecinos sobre las variables resultado de ocupación hotelera, precio y rentabilidad. Asimismo, para esta tabla se utiliza una matriz que presenta la conexión entre los establecimientos en términos de la proximidad en la oferta de servicios ponderada por los niveles asociados a las distancias de Gower.

Los resultados de la tabla se presentan en columnas: en la primera se tiene en cuenta como variable dependiente la rentabilidad; en la segunda el resultado se presenta en la tasa de ocupación; y finalmente se utiliza como variable-resultado el nivel de precios. En los resultados se incluyen los controles agrupados en características verticales, horizontales y atributos específicos de los establecimientos, tales como contar un bar, piscina, spa, entre otros. En las especificaciones de los modelos se incluyen efectos aleatorios, además del componente autorregresivo en los errores.

La columna 1, cuya variable dependiente es la rentabilidad, evidencia el efecto negativo en el parámetro de interés rho, debido a que incrementos en los precios de los vecinos implica una reducción en la rentabilidad propia. En el caso del coeficiente gamma, que muestra el efecto promedio de los puntajes de los vecinos, se observa que incrementos en la calificación promedio de los vecinos próximos por atributos, implican aumentos en la rentabilidad propia de cada establecimiento. Asimismo, se resaltan los efectos significativos en la rentabilidad de los establecimientos por atributos como el bar, la piscina o el spa; las modalidades de pago sin efectivo; y atributos individuales dentro de las habitaciones, como la nevera.

La columna 2 resalta que hay un efecto de red significativo en la tasa de ocupación hotelera, donde un aumento en las ocupaciones de los vecinos próximos, reduce la venta de habitaciones propias. De las variables firma-específica es importante señalar que, a medida que la distancia promedio a los competidores aumenta, la tasa de ocupación también.

Finalmente, la columna 3, evidencia que al observar el coeficiente gamma, la variable dependiente (el precio) aumenta al momento en que incrementan las clasificaciones de los competidores próximos en atributos. Es decir que el aumento promedio de precios de los vecinos reduce los propios.

### 3.5. Conclusiones

La literatura ha mostrado que los hoteles están inmersos en un entorno muy competitivo por lo que tienen que planificar y rediseñar sus estrategias para mantenerse en la industria de la hospitalidad mientras enfrentan un aumento constante de la competencia. Los hoteleros pueden emplear estrategias de diferenciación horizontal y decisiones de precios para obtener una ventaja competitiva sobre sus competidores.

Cuadro 3.2: Resultados

Variable	Rentabilidad	Tasa de ocupación	Precio
	1	2	3
$\rho$	-13.148*** (2.546)	-21.273*** (2.545)	-9.306*** (1.750)
$\gamma$	100,252.874** (50,417.969)	0.587 (0.829)	1,574.491*** (579.124)
Puntaje general	-2,389.438 (1,884.249)	0.063** (0.030)	61.457*** (17.850)
Promedio razón viajeros no residentes sobre residentes en 2019	-1,863.467 (3,610.809)	0.047 (0.056)	185.751*** (34.109)
Distancia promedio a los competidores (km)	968.494 (603.455)	0.022** (0.010)	-12.362** (5.483)
Centro de negocios	2,864.622 (2,044.135)	-0.031 (0.032)	-5.790 (18.584)
Actividades para niños	-1,937.612 (2,676.322)	-0.024 (0.042)	6.194 (24.171)
Televisión satelital	3,686.703*** (1,275.420)	0.016 (0.020)	35.123*** (12.467)
Nevera en la habitación	5,839.174*** (1,395.332)	0.134*** (0.022)	-17.028 (11.931)
Escritorio en la habitación	1,778.503 (2,152.959)	0.078** (0.033)	28.563 (19.386)
Servicio a la habitación	-538.013 (2,047.952)	-0.007 (0.032)	20.463 (18.814)
Pago sin efectivo	3,836.287** (1,663.177)	0.083*** (0.026)	-53.834*** (15.431)
Restaurante	1,229.645 (2,151.083)	0.018 (0.034)	-1.764 (19.331)
Wifi gratis	-1,003.254 (3,785.633)	0.049 (0.060)	-8.387 (34.501)
Aire acondicionado	10,119.115*** (2,072.750)	0.075** (0.032)	47.986** (19.603)
SPA	5,220.343*** (1,913.169)	0.045 (0.030)	17.345 (17.496)
Bar	3,723.440** (1,891.934)	0.075** (0.030)	6.996 (17.212)
Piscina	2,080.625 (1,647.055)	-0.020 (0.026)	65.498*** (16.667)
Gimnasio	7,884.498*** (2,400.502)	0.026 (0.038)	12.351 (21.871)
Sauna o turco	4,805.486*** (1,790.741)	-0.018 (0.029)	54.910*** (17.324)
Posibilidad de viajar con niños	3,069.019 (3,285.939)	0.067 (0.051)	-8.482 (28.976)
Sombrillas	2,446.464 (4,406.984)	0.001 (0.069)	22.851 (40.111)
Tienda	-4,223.563** (1,652.635)	-0.041 (0.026)	30.151** (15.242)
Niñera	-5,546.710** (2,431.587)	-0.005 (0.038)	8.128 (22.060)
Asador externo	1,262.160 (4,102.795)	-0.008 (0.065)	29.186 (37.549)
Parque infantil	10,943.659*** (3,309.560)	-0.042 (0.050)	26.419 (29.143)
Peluquería	-2,277.584 (2,143.114)	-0.110*** (0.034)	-14.885 (19.532)
Cuna	1,009.035 (3,092.446)	0.007 (0.047)	48.917* (27.138)
Conserje	-497.301 (1,405.146)	0.001 (0.022)	20.090 (12.884)
Botones	-4,769.441*** (1,628.952)	0.159*** (0.027)	-9.300 (14.342)
Parqueadero	7,313.386*** (2,062.357)	0.054* (0.032)	29.401 (18.834)
Transporte al aeropuerto	-2,352.450 (2,662.734)	0.050 (0.041)	-18.641 (24.102)
Observaciones	1240	1240	1240
Efectos fijos del número de estrellas	✓	✓	✓

Notas: Errores estándar en paréntesis. Niveles de significancia: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Illescas-Manzano, et.al, (2022). Este trabajo proporciona información sobre cómo los hoteles incorporando elementos en este proceso de toma de decisiones, deben tener en cuenta los esfuerzos por diferenciarse de sus competidores

Los resultados alcanzados por este estudio permiten identificar que el comportamiento de la competencia de los hoteles, y la presentan una relación de dependencia negativa entre ellos, medido bajo las variables: rentabilidad, tasa de ocupación y precio. En contraste, el impacto positivo de las evaluaciones realizadas por visitantes de la competencia en la rentabilidad y precios propios, y un efecto no significativo en la tasa de ocupación. Contribuye a los resultados de la literatura que en el que el precio de la habitación de hotel en mercados altamente competitivos es un factor importante que influye en el rendimiento propio del hotel (Hiemstra e Ismail (1994) y Qu et. al., 2002; y que esfuerzos por diferenciarse en calidad no observada de los competidores afecta los ingresos por habitación a través del precio de la habitación, no de la cantidad demandada de habitaciones de alojamiento (Chen y Lin, 2013).

Esta metodología novedosa permite ampliar la discusión en la literatura de pricing y diferenciación horizontal de un hotel sobre su reputación online, en la que ya se ha sugerido la relación positiva aquí obtenida con la rentabilidad y los precios Illescas-Manzano, et.al, (2022), e incorpora la dependencia con las decisiones de los competidores en la que se establece que la proximidad por atributos genera procesos de competencia en las plataformas, que se traduce en una divulgación de información medida a través de las opiniones de los consumidores. De igual modo, los resultados sugieren una interdependencia entre el precio y los beneficios y, por su parte, la estructura en la proximidad de los alojamientos muestra un efecto significativo entre las calificaciones de los competidores y las variables resultado de precio y rentabilidad.

Este estudio permitió explotar técnicas econométricas utilizadas comúnmente para medir efectos geográficos, donde los instrumentos que solucionan la endogeneidad presentada por variable omitida están determinados por la interacción de las variables y los rezagos de orden superior. Sin embargo, en este documento no han sido probados los instrumentos relacionados con la proximidad geográfica por atributos. Por este motivo, a futuro pueden ser considerados otro tipo de matrices que determinen la proximidad entre dos establecimientos, así como diseñar un algoritmo que establezca cuáles atributos deben ser excluidos en la construcción de la distancia entre unidades y establecer ponderaciones específicas para cada uno de estos atributos.

Por otro lado, es pertinente continuar teniendo en cuenta que las encuestas con-

tinuarán aportando información que no es posible conseguir de otra manera, como el motivo de viaje, los medios de transporte y las características sociodemográficas (EUROSTAT, 2017)). No obstante, se requiere que sean combinadas con otro tipo de fuentes de información para lograr un mayor entendimiento del sector hotelero. De este modo, este estudio constituye en un avance en las aplicaciones que integran y armonizan los datos obtenidos a través de web-scraping con otras operaciones estadísticas como encuestas y registros estadísticos para solucionar las limitaciones de información en la construcción de estadísticas de turismo.

### 3.6. Referencias

Adhinugroho, Y., Putra, A. P., Luqman, M., Ermawan, G. Y., Mariyah, S., & Pramaana, S. (2020). Development of online travel Web scraping for tourism statistics in Indonesia. Anselin, L. (1980). Estimation Methods for Spatial Autoregressive Structures. Regional Science Dissertation and Monograph Series 8. Field of Regional Science, Cornell University, Ithaca, N.Y.

Anselin, L. (1988). Model validation in spatial econometrics: a review and evaluation of alternative approaches. *International Regional Science Review*, 11(3), 279-316.

Anselin, L. (2001). Spatial econometrics. A companion to theoretical econometrics, 310330.

Anselin, L. (2006). How (not) to lie with spatial statistics. *American journal of preventive medicine*, 30(2), S3-S6.

Belleflamme, P., & Peitz, M. (2019). Platform competition: Who benefits from multihoming?. *International Journal of Industrial Organization*, 64, 1-26.

Bickart, B., & Schindler, R. M. (2001). Internet forums as influential sources of consumer information. *Journal of interactive marketing*, 15(3), 31-40.

Bramoullé, Y., Djebbari, H., & Fortin, B. (2009). Identification of peer effects through social networks. *Journal of econometrics*, 150(1), 41-55.

Cabral, L., Peitz, M., & Wright, J. (2019). Introduction to special issue on platforms. *Journal of Economics & Management Strategy*, 28(1), 3-4.

Chen, Y., & Yao, S. (2017). Sequential search with refinement: Model and application with click-stream data. *Management Science*, 63(12), 4345-4365.

De Los Santos, Babur I., Hortaçsu, Ali, Wildenbeest, Matthijs R. (2017), "Search with Learning for Differentiated Products: Evidence from E-Commerce," *Journal of*

Business & Economic Statistics, 35 (4), 626–41

Dubois, P., & Nauges, C. (2009). Identifying the effect of unobserved quality and expert reviews in the pricing of experience goods: Empirical application on Bordeaux wine. In *World Scientific Reference on Handbook of the Economics of Wine: Volume 1: Prices, Finance, and Expert Opinion* (pp. 459-481).

Elhorst, J. P. (2014). Linear spatial dependence models for cross-section data. In *Spatial econometrics* (pp. 5-36). Springer, Berlin, Heidelberg.

Filieri, R., & McLeay, F. (2014). E-WOM and accommodation: An analysis of the factors that influence travelers' adoption of information from online reviews. *Journal of travel research*, 53(1), 44-57.

Forman, C., Ghose, A., & Wiesenfeld, B. (2008). Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Information systems research*, 19(3), 291-313.

Franck, J. U., & Peitz, M. (2019). Market definition and market power in the platform economy. Centre on Regulation in Europe asbl (CERRE).

Hagiu, A. (2009). Two-sided platforms: Product variety and pricing structures. *Journal of Economics & Management Strategy*, 18(4), 1011-1043.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: what motivates consumers to articulate themselves on the internet?. *Journal of interactive marketing*, 18(1), 38-52.

Jiang, B., & Zou, T. (2020). Consumer search and filtering on online retail platforms. *Journal of Marketing Research*, 57(5), 900-916.

Kapoor, M., Kelejian, H. H., & Prucha, I. R. (2007). Panel data models with spatially correlated error components. *Journal of econometrics*, 140(1), 97-130.

Kelejian, H. H., & Prucha, I. R. (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. H., & Prucha, I. R. (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International economic review*, 40(2), 509-533.

Kelejian, H. H., & Prucha, I. R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of econometrics*, 157(1), 53-67.

Lee, L. F., & Yu, J. (2010). A spatial dynamic panel data model with both time

and individual fixed effects. *Econometric Theory*, 26(2), 564-597.

Mantovani, A., Piga, C. A., & Reggiani, C. (2017). The dynamics of online hotel prices and the EU Booking.com case. *Com Case* (September 30, 2017) .NET Institute Working Paper, (17-04).

Nelson, P. (1974). "Advertising as Information," *Journal of Political Economy*, 83 (July-August), 729-54.

Nunkoo, R., Teeroovengadum, V., Ringle, C. M., & Sunnassee, V. (2020). Service quality and customer satisfaction: The moderating effects of hotel star rating. *International Journal of Hospitality Management*, 91, 102414.

Ord, K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association*, 70(349), 120-126.

Reyes-Menendez, A., Saura, J. R., & Martinez-Navalon, J. G. (2019). The impact of e-WOM on hotels management reputation: exploring tripadvisor review credibility with the ELM model. *IEEE Access*, 7, 68868-68877.

Shy, O. (2010). Consistent bargaining. Available at SSRN 3083747.

Zhu, F., & Zhang, X. (2006). The influence of online consumer reviews on the demand for experience goods: The case of video games. *ICIS 2006 proceedings*, 25.

Akerlof, G. A. (1978). The market for "lemons": Quality uncertainty and the market mechanism. In *Uncertainty in economics* (pp. 235-251). Academic Press.

Alonso, A. D., Kok, S. K., Bressan, A., O'Shea, M., Sakellarios, N., Koresis, A., & Santoni, L. J. (2020). COVID-19, aftermath, impacts, and hospitality firms: An international perspective. *International journal of hospitality management*, 91, 102654.

American Hospitality and Lodging Association, 2020. Covid-19's Impact on the Hotel Industry [accessed on 8 May 2020] Available at.

<https://www.ahla.com/covid-19s-impact-hotel-industry>.

Amaro, S., & Duarte, P. (2015). An integrative model of consumers' intentions to purchase travel online. *Tourism management*, 46, 64-79.

Arcay, A. O., & Benítez, F. G. (2006). Modelos de elección discreta en transportes con coeficientes aleatorios. *Cátedra abertis*.

Archer, B., & Shea, S. (1977). Manpower in tourism: the situation in Wales. *Manpower in tourism: the situation in Wales*.

Audretsch, D. B., & Feldman, M. P. (1996). R&D spillovers and the geography of innovation and production. *The American economic review*, 86(3), 630-640.

Bagnera, S.M., Stewart, E., Edition, S., 2020. Navigating hotel operations in times of COVID-19. *Boston Hospitality Rev.* ahead-of-print.

Banerjee, A. V. (1992). A simple model of herd behavior. *The quarterly journal of economics*, 107(3), 797-817.

Bar-Isaac, H., Caruana, G., & Cuñat, V. (2009). Costly search and design. Available at SSRN 1435257.

Baum, J. A., & Haveman, H. A. (1997). Love thy neighbor? Differentiation and agglomeration in the Manhattan hotel industry, 1898-1990. *Administrative Science Quarterly*, 304-338.

Baum, J., & Mezias, S. (1992). Localized competition and organizational failure in the Manhattan hotel industry, 1898–1990. *Administrative Science Quarterly*, 37(4), 580–604.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 242-262.

Berry, S., Levinsohn, J., & Pakes, A. (1995). Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, 841-890.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, 100(5), 992-1026.

Boni, P (2005) The Distance Between a Hotel's Keywords and Bookings: How to close the deal after Search Engine Optimization.

Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the Internet—The state of eTourism research. *Tourism management*, 29(4), 609-623.

Branicki, L.J., Sullivan-Taylor, B., Livschitz, S.R., 2018. How entrepreneurial resilience generates resilient SMEs. *Int. J. Entrep. Behav. Res.* 24 (7), 1244–1263.

Brown, N.A., Orchiston, C., Rovins, J.E., Feldmann-Jensen, S., Johnston, D., 2018. An integrative framework for investigating disaster resilience within the hotel sector. *J. Hosp. Tour. Manag.* 36, 67–75.

Browning, V., So, K. K. F., & Sparks, B. (2013). The influence of online reviews on consumers' attributions of service quality and control for service standards in hotels. *Journal of Travel & Tourism Marketing*, 30(1-2), 23-40.

Cabello Beltrán, E. A., & Pulido Arbeláez, L. N. (2020). Análisis del impacto económico del coronavirus en el sector turismo en Colombia.

CBRE, 2020. Insights and Implications of Covid-19 Pandemic [accessed online on 8 May 2020] Available at: <https://www.cbre.com/covid-19>.

Chathoth, P.K., Olsen, M.D., 2003. Strategic alliances: a hospitality industry perspective. *Int. J. Hosp. Manag.* 22 (4), 419–434.

Chen, F., Federgruen, A., Zheng, Y.S., 2001. Coordination mechanisms for a distribution system with one supplier and multiple retailers. *Manag. Sci.* 47 (5), 693–708.

Chen, Y., & Xie, J. (2008). Online consumer review: Word-of-mouth as a new element of marketing communication mix. *Management science*, 54(3), 477-491.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of marketing research*, 43(3), 345-354.

Christopher, M., 2011. *Logistics & Supply Chain Management*. Pearson education limited, United Kingdom.

Chung, L. L. W. (1994). The economics of land-use zoning: a literature review and analysis of the work of Coase. *The Town Planning Review*, 65(1), 77–98.

Datta, S., & Sudhir, K. (2013). Does reducing spatial differentiation increase product differentiation? Effects of zoning on retail entry and format variety. *Quantitative Marketing and Economics*, 11(1), 83-116.

Davis, B. (2021). Why accommodation is important to hospitality and tourism industry? – MVOrganizing.

<https://www.mvorganizing.org/why-accommodation-is-important-to-hospitality-and-tourism>

Debo, L. G., & Veeraraghavan, S. K. (2009). Models of herding behavior in operations management. In *Consumer-Driven Demand and Operations Management Models* (pp. 81-112). Springer, Boston, MA.

Dellarocas, C., & Wood, C. A. (2008). The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management science*, 54(3), 460-476.

Dürr, E., & Jaffe, R. (2012). Theorizing slum tourism: Performing, negotiating and transforming inequality. *European Review of Latin American and Caribbean Studies/Revista Europea de Estudios Latinoamericanos y del Caribe*, 113-123.

Eggers, F., 2020. Masters of disasters? Challenges and opportunities for SMEs in times of crisis. *J. Bus. Res.* 116, 199–208.

Elshaer, Abdallah M., and Asmaa M. Marzouk. *Labor in the tourism and hospitality industry: skills, ethics, issues, and rights*. Apple Academic Press, 2019.

European Parliament, 2020. Covid-19 and the Tourism Sector. Available at: [https://www.europarl.europa.eu/RegData/etudes/ATAG/2020/649368/EPRS\\_ATA\(2020\)649368\\_](https://www.europarl.europa.eu/RegData/etudes/ATAG/2020/649368/EPRS_ATA(2020)649368_)

EN.pdf

Evans, D., 2020. Airbnb's Future Is Uncertain as It Continues to Struggle Through Its Covid-19 Response. CNBC [accessed online 11 May 2020] Available at: <https://www.cnbc.com/2020/05/06/can-airbnb-survive-the-coronavirus-pandemic.html>.

Farmaki, A., Miguel, C., Drotarova, M. H., Aleksić, A., Časni, A. Č., & Efthymiadou, F. (2020). Impacts of Covid-19 on peer-to-peer accommodation platforms: Host perceptions and responses. *International journal of hospitality management*, 91, 102663.

Feldman, P., Papanastasiou, Y., & Segev, E. (2019). Social learning and the design of new experience goods. *Management Science*, 65(4), 1502-1519.

Feldman, M. P., & Audretsch, D. B. (1999). Innovation in cities:: Science-based diversity, specialization and localized competition. *European economic review*, 43(2), 409-429.

Glusac, E., 2020. Hotels Vs Airbnb: Has Covid-19 Disrupted the Disrupter? [accessed online on 14 June 2020] Available at: New York Times <https://www.nytimes.com/2020/05/14/travel/hotels-versus-airbnb-pandemic.html>.

González-Torres, T., Rodríguez-Sánchez, J. L., & PelechanoBarahona, E. (2021). Managing relationships in the Tourism Supply Chain to overcome epidemic outbreaks: The case of COVID-19 and the hospitality industry in Spain. *International Journal of Hospitality Management*, 92, 102733.

Guillet, B. D., & Law, R. (2010). Analyzing hotel star ratings on third-party distribution websites. *International Journal of Contemporary Hospitality Management*.

Hafsa, S. (2020). Impacts of COVID-19 Pandemic on Tourism & Hospitality Industry in Bangladesh. Available at SSRN 3659196.

Hao, F., Xiao, Q., & Chon, K. (2020). COVID-19 and China's hotel industry: Impacts, a disaster management framework, and postpandemic agenda. *International journal of hospitality management*, 90, 102636.

HIEMSTRA, S. e ISMAIL, J.A. (2001): «Research Note to Incidence of the Impacts of Room Taxes on the Lodging Industry», Revision to Article Published in Spring 1993 Issue of HR. *Journal of Travel Research*, vol. 39, 319-320. Hong, Y., & Pavlou, P. A. (2014). Product fit uncertainty in online markets: Nature, effects, and antecedents. *Information Systems Research*, 25(2), 328-344.

Illescas-Manzano, M., Martínez-Puertas, S., & Sánchez-Pérez, M. (2022). The Power of Price and Quality to Explain Customer Satisfaction Through Spatial Analysis.

In Contemporary Approaches Studying Customer Experience in Tourism Research (pp. 245-265). Emerald Publishing Limited. ILO. (2021). Hotels, catering and tourism sector. International Labour Organization. <https://www.ilo.org/global/industries-and-sectors/hotels-catering-tourism/lang--en/index.htm>

Jamal, T., Budke, C., 2020. Tourism in a world with pandemics: local-global responsibility and action. *J. Tour. Futures*.

Jannach, D., Gedikli, F., Karakaya, Z., & Juwig, O. (2012). Recommending hotels based on multi-dimensional customer ratings. In ENTER (pp. 320-331).

Jayawardena, C., Tew, P.J., Lu, Z., Tolomiczenko, G., Gellatly, J., 2008. SARS: lessons in strategic planning for hoteliers and destination marketers. *Int. J. Contemp. Hosp. Manage.* 20 (3), 332–346.

Jiang, Y., Ritchie, B.W., Benckendorff, P., 2019. Bibliometric visualisation: an application in tourism crisis and disaster management research. *Curr. Issues Tour.* 22 (16), 1925–1957.

Kopczuk, W., 2020. Some Thoughts on Economics and Policy at the Time of the Pandemics. Department of Economics and SIPA, Columbia University, New York.

Kuckertz, A., Bröndle, L., Gaudig, A., Hinderer, S., Reyes, C.A.M., Prochotta, A., Steinbrink, K.M., Berger, E.S., 2020. Startups in times of crisis—a rapid response to the COVID-19 pandemic. *J. Bus. Ventur. Insights*, e00169.

Kutuzov, D. (2004). Buyer search costs and endogenous product design. *Marketing Science*, 23(4), 490–499.

Kuo, H.I., Chen, C.C., Tseng, W.C., Ju, L.F., Huang, B.W., 2008. Assessing impacts of SARS and Avian Flu on international tourism demand to Asia. *Tour. Manag.* 29 (5), 917–928.

Kwark, Y., Chen, J., & Raghunathan, S. (2014). Online product reviews: Implications for retailers and competing manufacturers. *Information systems research*, 25(1), 93-110.

Leung, R., Au, N., Liu, J., & Law, R. (2018). Do customers share the same perspective? A study on online OTAs ratings versus user ratings of Hong Kong hotels. *Journal of Vacation Marketing*, 24(2), 103-117.

Lord, P., Saad, L., 2020. Tackling the COVID-19 Pandemic. <https://doi.org/10.1016/j.arcmed.2020.04.012>

Liu, X., Schuckert, M., & Law, R. (2015). Can response management benefit hotels? Evidence from Hong Kong hotels. *Journal of Travel & Tourism Marketing*, 32(8), 1069-

1080.

Luca, M. (2016). Reviews, reputation, and revenue: The case of Yelp. com. Com (March 15, 2016). Harvard Business School NOM Unit Working Paper, (12-016).

Luca, M., & Zervas, G. (2016). Fake it till you make it: Reputation, competition, and Yelp review fraud. *Management Science*, 62(12), 3412-3427.

Li, X., Chen, H., Lu, L., Chen, L. L., Chan, B. P. C., Wong, S. C., ... & To, K. K. W. (2022). High compliance to infection control measures prevented guest-to-staff transmission in COVID-19 quarantine hotels. *Journal of Infection*, 84(3), 418-467.

Marshall, A. (1920). *Principles of economics*. London: Macmillan.

McFadden, D. (1974). The measurement of urban travel demand. *Journal of public economics*, 3(4), 303-328

Milgrom, P., & Roberts, J. (1986). Price and advertising signals of product quality. *Journal of Political Economy*, 94(4), 796-821.

Montoya, Y. M., & Pérez, D. A. P. (2016). Turismo en Colombia ¿Es el turismo una alternativa real para enfrentar el déficit externo? *Ploutos*, 6(2), 48-55.

Monczka, R.M., Petersen, K.J., Handfield, R.B., Ragatz, G.L., 1998. Success factors in strategic supplier alliances: the buying company perspective. *Decis. Sci.* 29 (3), 553-577.

Nelson, P. (1970). Information and consumer behavior. *Journal of Political Economy*, 78(2), 311-329.

Neuburger, L., & Egger, R. (2021). Travel risk perception and travel behaviour during the COVID-19 pandemic 2020: A case study of the DACH region. *Current Issues in Tourism*, 24(7), 1003-1016. Nicola, M., Alsafi, Z., Sohrabi, C., Kerwan, A., Al-Jabir, A., Iosifidis, C., Agha, M., Agha, R., 2020. The socio-economic implications of the coronavirus and COVID-19 pandemic: a review. *Int. J. Surg.* 78, 185-193.

Nicolaidis, C., Avraam, D., Cueto-Felgueroso, L., Gonzalez, M.C., Juanes, R., 2019. Hand hygiene mitigation strategies against global disease spreading through the air transportation network. *Risk Anal.* 40 (4), 723-740.

Nosko, C., & Tadelis, S. (2015). The limits of reputation in platform markets: An empirical analysis and field experiment (No. w20830). National Bureau of Economic Research. Norris, F.H., Stevens, S.P., Pfefferbaum, B., Wyche, K.F., Pfefferbaum, R.L., 2008. Community resilience as a metaphor, theory, set of capacities, and strategy for disaster readiness. *Am. J. Commun. Psychol.* 41 (1-2), 127-150.

Novelli, M., Burgess, L.G., Jones, A., Ritchie, B.W., 2018. No Ebola... still doo-

med’–the Ebola-induced tourism crisis. *Ann. Tour. Res.* 70, 76–87.

OECD, 2020. Tourism Policy Responses to the Coronavirus (COVID-19). Obtenido de: <https://www.oecd.org/coronavirus/policy-responses/tourism-policy-responses-to-the-#endnotea0z14>. (Acceso: enero, 2021).

O’Connor, P., & Frew, A. J. (2002). The future of hotel electronic distribution: Expert and industry perspectives. *Cornell Hotel and Restaurant Administration Quarterly*, 43(3), 33-45.

Onur, I., Bruwer, J., & Lockshin, L. (2020). Reducing information asymmetry in the auctioning of non-perishable experience goods: The case of online wine auctions. *Journal of Retailing and Consumer Services*, 54, 102060.

Pal, R., Torstensson, H., Mattila, H., 2014. Antecedents of organizational resilience in economic crises—an empirical study of Swedish textile and clothing SMEs. *Int. J. Prod. Econ.* 147, 410–428.

Park, E., Kim, W. & Kim, S. (2020). Tracking tourism and hospitality employees’ real-time perceptions and emotions in an online community during the COVID-19 pandemic. *Current Issues in Tourism*. DOI: 10.1080/13683500.2020.1823336 Pizam, Abraham. "Tourism manpower: the state of the art." *Journal of Travel research* 21.2 (1982): 5-9.

Pizam, A., Fleischer, A., 2002. Severity versus frequency of acts of terrorism: which has a larger impact on tourism demand? *J. Travel. Res.* 40 (3), 337–339.

PRAMANA, S., PARAMARTHA, D. Y., ADHINUGROHO, Y., & NURMALASARI, M. (2020). Air pollution changes of Jakarta, Banten, and West Java, Indonesia during the first month of COVID-19 pandemic. *Asian Journal of Business Environment*, 10(4), 15-19. Pramana, S., Paramartha, D. Y., Ermawan, G. Y., Deli, N. F., & Srimulyani, W. (2022). Impact of COVID-19 pandemic on tourism in Indonesia. *Current Issues in Tourism*, 25(15), 2422-2442. Pogodzinski, J. M. (1991). The effects of fiscal and exclusionary zoning on household location: a critical review. *Journal of Housing Research*, 2(2), 145–160.

Polemis, M. (2020). The impact of COVID-19 on hotel performance: Evidence from a Difference-in-Differences approach.

Rhee, H. T., & Yang, S. B. (2015). Does hotel attribute importance differ by hotel? Focusing on hotel star-classifications and customers’ overall ratings. *Computers in Human Behavior*, 50, 576-587.

Roger-Monzó, V., Martí-Sánchez, M., & Guijarro-García, M. (2015). Using online

consumer loyalty to gain competitive advantage in travel agencies. *Journal of Business Research*, 68(7), 1638-1640.

Runfola, A, Rosati, M, Guercini, S (2013). New business models in online hotel distribution: emerging private sales versus leading IDS. *Service Business* 7(2): 183–205.

Sánchez-Pérez, M., Illescas-Manzano, M. D., & Martínez-Puertas, S. (2022). Online review ratings: an analysis of product attributes and competitive environment. *Journal of Marketing Communications*, 28(5), 487-505.

Scaglione, M, Schegg, R (2015) The impact of attribute preferences on adoption timing of hotel distribution channels: are OTAs winning the customer race? In: Tussyadiah, I, Inversini, A (eds) *Information and Communication Technologies in Tourism 2015*. London: Springer International Publishing, pp. 681–693.

Seabra, C., Reis, P., Abrantes, J.L., 2020. The influence of terrorism in tourism arrivals: a longitudinal approach in a Mediterranean country. *Ann. Tour. Res.* 80, 102811.

Serrano Antón, F., 2020. Medidas fiscales terapéuticas y COVID 19, 7. Instituto Complutense de Estudios Internacionales, pp. 1–4.

Shapiro, C. (1983). Optimal pricing of experience goods. *The Bell Journal of Economics*, 497-507.

Sharma, A., and Nicolau, J.L. (2020). An open market valuation of the effects of COVID-19 on the travel and tourism industry, *Annals of Tourism Research*, <https://doi.org/10.1016/j.annals.2020.102990>.

Sigala, M., 2004. Collaborative supply chain management in the airline sector: the role of global distribution systems (GDS). *Adv. Hosp. Leis.* 1 (1), 103–121.

Sigala, M., 2017. Collaborative commerce in tourism: implications for research and industry. *Curr. Issues Tour.* 20 (4), 346–355.

Tapper, R., Font, X., 2004. *Tourism Supply Chains*. Report of a Desk Research Project for The Travel Foundation, Leeds Metropolitan University, Leeds.

oh, R.S., Raven, P. & DeKay, F. (2011) Selling rooms: hotels vs. third-party websites. *Cornell Hospitality Quarterly* 52(2): 181–189.

Tunc, M. M., Cavusoglu, H., & Raghunathan, S. (2019). Two-Sided Adverse Selection and Bilateral Reviews in Sharing Economy. Available at SSRN 3499979.

Tsang, E. W., & Yip, P. S. (2009). Competition, agglomeration, and performance of Beijing hotels. *The Service Industries Journal*, 29(2), 155-171.

United Nations. (2016). *Labour intensive industries*. United Nations Economic

and Social Commission for Western Asia (ESCWA). <https://archive.unescwa.org/labour-intensive-industries>

Voltes-Dorta, A., & Inchausti-Sintes, F. (2021). The spatial and quality dimensions of Airbnb markets. *Tourism Economics*, 27(4), 688-702.

Wu, E.H., Law, R., Jiang, B., 2010. The impact of infectious diseases on hotel occupancy rate based on independent component analysis. *Int. J. Hosp. Manag.* 29 (4), 751–753.

Yehia, Y. (2019). The Importance of Tourism on Economies and Businesses. *GlobalEDGE*. <https://globaledge.msu.edu/blog/post/55748/the-importance-of-tourism-on-econom>

Ying, T., Wang, K., Liu, X., Wen, J., Goh, E., 2020. Rethinking game consumption in tourism: a case of the 2019 novel coronavirus pneumonia outbreak in China. *Tour. Recreat. Res.* 1–6.

Zenker, S., and Kock, F. (2020). The coronavirus pandemic – A critical discussion of a tourism research agenda, *Tourism Management*, 81, <https://doi.org/10.1016/j.tourman.2020.104164>.

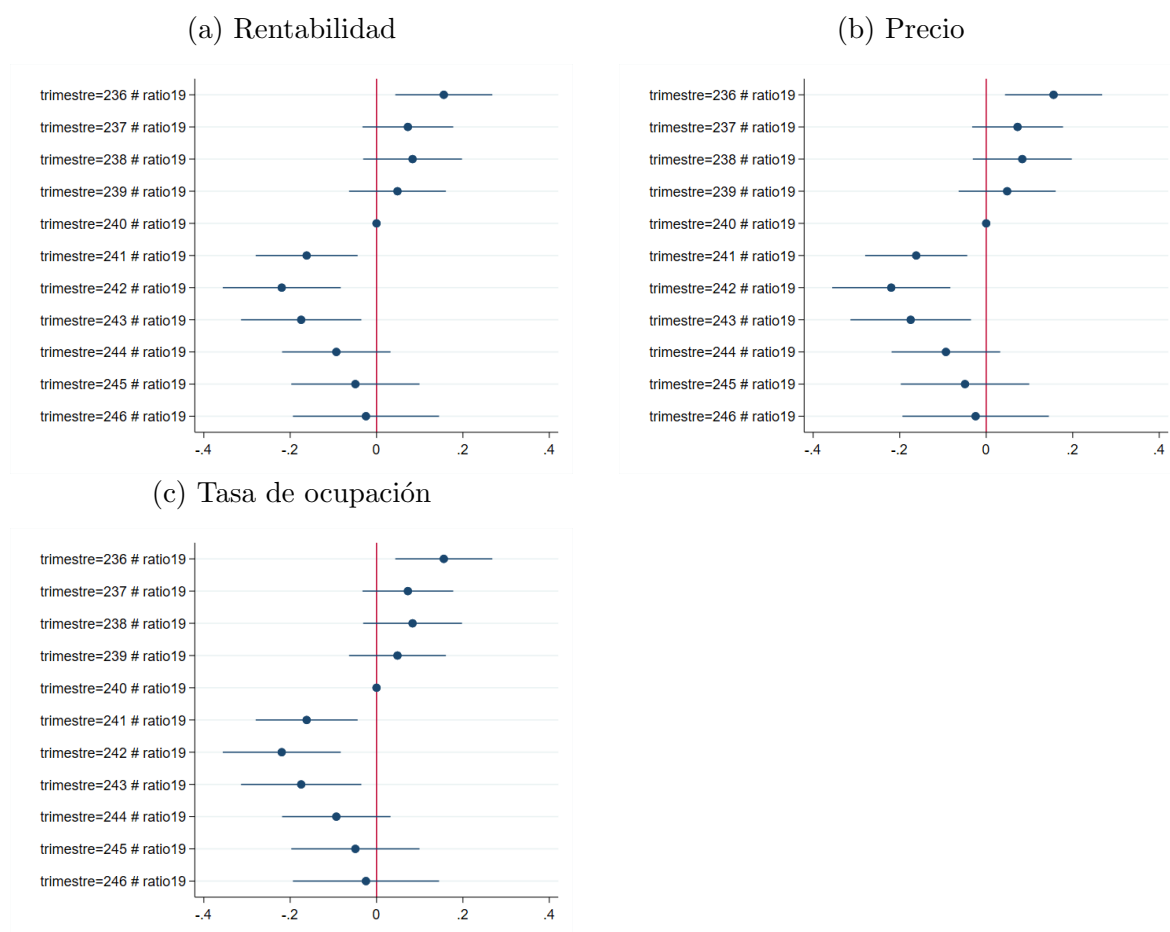
Zhang, X., Song, H., Huang, G.Q., 2009. Tourism supply chain management: a new research agenda. *Tour. Manag.* 30 (3), 345–358.

# Appendices

# Apéndice A

## Chapter 1

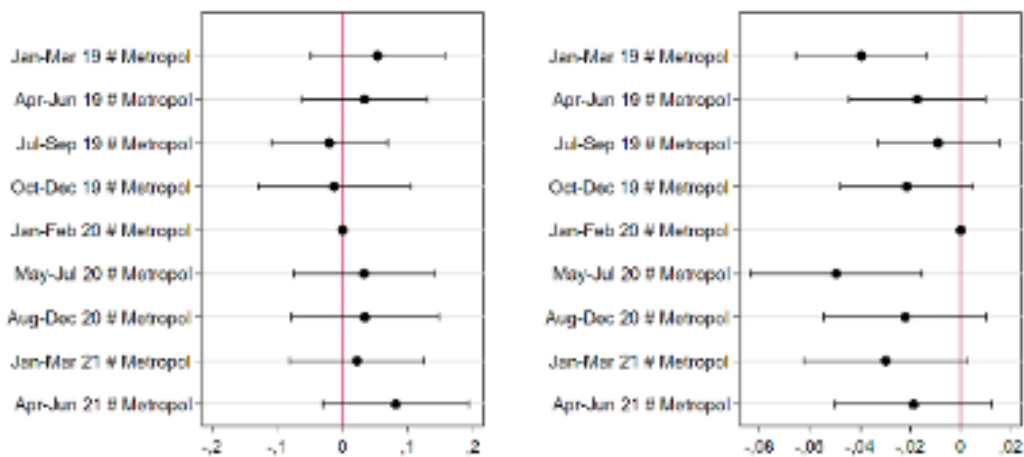
Figura A.1: Modelo sin efectos heterogéneos por diferenciación, usando porcentaje de huéspedes que viajan por negocios como variable de exposición.



# Apéndice B

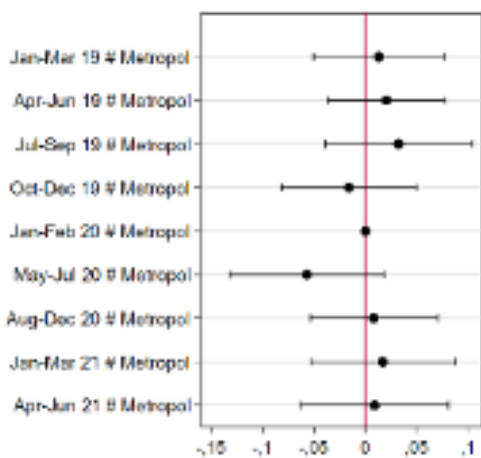
## Chapter 2

Figura B.1: Impacto estimado por rango de edad



(a) Menos de 25 años

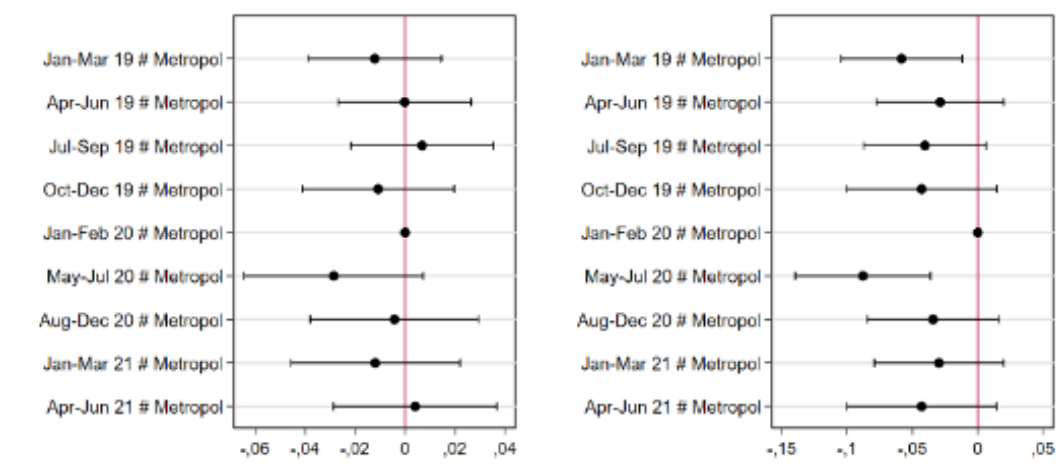
(b) Entre 25 y 55 años



(c) Más de 55 años

Fuente: Cálculos propios. Las líneas representan un intervalo de confianza al 95%.

Figura B.2: Impacto estimado por tipo de hogar



(a) Hogar con niños

(b) Hogar sin niños

Fuente: Cálculos propios. Las líneas representan un intervalo de confianza al 95%.