ISSN 2718-7578

Serie Documentos de Trabajo del CEP XXI

Patrones geográficos del empleo formal en CABA

Evidencia de una nueva base de datos

Martin Trombetta, Agustín Duarte Baracat, Azul Menduiña y Nicolas Sidicaro

Documento de Trabajo N° 16

Diciembre 2022

Cita sugerida: Trombetta, M.; Duarte Baracat, A.; Menduiña, A.; y Sidicaro, N. (diciembre de 2022). Patrones geográficos del empleo formal en CABA: evidencia de una nueva base de datos. Documentos de Trabajo del CEP XXI N° 16, Centro de Estudios para la Producción XXI – Secretaría de Industria y Desarrollo Productivo – Ministerio de Economía de la Nación.

Ministerio de Economía Argentina

Secretaría de Industria y Desarrollo Productivo

CEPXXI CENTRO DE ESTUDIOS PARA LA PRODUCCIÓN

Patrones geográficos del empleo formal en CABA

Evidencia de una nueva base de datos

Martin Trombetta, Agustín Duarte Baracat, Azul Menduiña y Nicolas Sidicaro

Corrección y diagramación: Natalia Rodríguez Simón y Juliana Adamow

ISSN 2718-7578

Centro de Estudios para la Producción XXI (CEP XXI)
Ministerio de Desarrollo Productivo de la Nación

Julio A. Roca 651, Ciudad Autónoma de Buenos Aires

cepxxi@produccion.gob.ar

Los Documentos de Trabajo del CEP XXI son resultado de investigaciones sobre temas productivos, sectoriales, laborales y otros. Los autores son responsables de las opiniones expresadas en estos documentos.

Autoridades

Presidente de la Nación
Alberto Fernández

Vicepresidenta de la Nación
Cristina Fernández de Kirchner

Jefe de Gabinete de Ministros
Juan Luis Manzur

Ministro de Economía
Sergio Tomás Massa

Secretario de Industria y Desarrollo Productivo
José Ignacio de Mendiguren

Directora del Centro de Estudios para la Producción XXI (CEP-XXI)
María Florencia Asef Horno

Resumen

La distribución espacial del empleo es un tema poco estudiado en países en desarrollo, lo que explica la escasa evidencia existente sobre problemas relevantes de la economía urbana, tales como la segregación residencial por ingresos y el costo de la movilidad dentro de las ciudades. Utilizando una nueva base de datos que incluye la georreferenciación del domicilio y del lugar de trabajo de los empleados formales, se exploran los principales atributos de la distribución geográfica del empleo formal en la Ciudad Autónoma de Buenos Aires. Se estiman diversas ecuaciones de salarios que incluyen, además de los determinantes demográficos tradicionales, variables espaciales de interés tales como la distancia al trabajo, el nivel de aglomeración económica, la densidad poblacional y la disponibilidad de transporte público, entre otras. También se recurre a la metodología de modelos con autocorrelación espacial con el fin de capturar potenciales derrames geográficos. Los resultados alcanzados muestran evidencia consistente con las predicciones teóricas para una ciudad monocéntrica que incluye diversas vías de transporte y considerable variabilidad en la distribución del ingreso.

Clasificación JEL: *J61, R32, C21.*

Índice

1. Introducción 6
2. Datos 7
3. Marco teórico y resultados descriptivos 8
4. Metodología 14
5. Resultados 16
6. Comentarios finales 31
Referencias bibliográficas 32

1. Introducción

A pesar del extenso desarrollo de la economía laboral empírica en Argentina y en América Latina en general, todavía es escasa la evidencia existente sobre distribución espacial del empleo y el ingreso al interior de los aglomerados urbanos. La falta de bases de datos georreferenciadas ha limitado seriamente el desarrollo de la economía urbana a nivel empírico en la región y en nuestro país; en particular, casi no existe en la literatura económica evidencia sobre fenómenos que revisten una larga tradición en la producción teórica, como la segregación residencial, las economías de aglomeración y el costo de la movilidad.

Excepcionalmente, Chauvin, *et. al.* (2017) contrastan algunos hechos estilizados de ciudades de Estados Unidos con ciudades de Brasil, China e India. Vargas, *et. al.* (2017) y Roberts, *et. al.* (2017) analizan diferentes patrones para ciudades de la región, mientras que Ferreyra y Roberts (2018) y Quintero y Roberts (2018) estudian sus diferencias de productividad. En particular para la Ciudad de Buenos Aires, González y Anapolsky (2022) analizan patrones de viaje de diferentes usuarios de bajo nivel socioeconómico en la Región Metropolitana, mientras que Mendiola y Gonzalez (2021) estudian la relación entre ciertos factores de desarrollo urbano y el uso de medios de transporte específicos en el *commuting* dentro de esta región.

Este trabajo explota una nueva base de datos georreferenciada por el Centro de Estudios para la Producción (CEP XXI) y el Ministerio de Trabajo, Empleo y Seguridad Social (MTEySS) para ofrecer evidencia sobre la distribución espacial del empleo formal en la Ciudad Autónoma de Buenos Aires (CABA), partiendo del marco teórico que provee la economía urbana. Se realizó un análisis descriptivo de los principales patrones geográficos que caracterizan el mercado de trabajo formal y se estimaron modelos que vinculan los ingresos de los trabajadores con variables tales como la distancia de viaje hasta su trabajo y el nivel de concentración económica del área en que reside, incluyendo la posibilidad de derrames geográficos en la forma de correlación espacial.

Se encuentra una fuerte segregación por ingresos y una marcada tendencia de los trabajadores de menores salarios a residir en la zona sur de la ciudad, donde la concentración de trabajo formal es menor. En línea con esta idea, las estimaciones espaciales encuentran un efecto total negativo para la variable de distancia de *commuting*, indicando que el efecto predominante de la relación es que las personas de mayores ingresos deciden residir cerca de sus trabajos de modo tal de ahorrar en costos de viaje. Por su parte, los resultados no son concluyentes para el caso de la variable de aglomeración, cuyo efecto total difiere notablemente de acuerdo a la especificación del modelo y el método de estimación elegidos. No se encuentra una relación significativa de la remuneración con la distancia al subterráneo, aunque sí con la medida de frecuencia de colectivos y la densidad poblacional. Finalmente, se encuentra una relación positiva entre los salarios pagados por las firmas y la pertenencia a un distrito económico, aunque el signo de la relación se invierte cuando se controla por efectos espaciales; esto, posiblemente se explique por el hecho de que los distritos económicos están situados en torno a zonas generalmente poco favorecidas de la ciudad.

El resto del trabajo está estructurado de la siguiente manera. La sección 2 explica las fuentes de datos utilizadas y la definición de las principales variables de interés. La sección 3 describe el marco teórico y comenta algunos resultados descriptivos. La sección 4 desarrolla la metodología de estimación implementada. La sección 5 contiene los resultados de los modelos estimados. Finalmente, la sección 6 ofrece algunos comentarios finales.

2. Datos

Las principales fuentes de información utilizadas en este trabajo provienen del Sistema Previsional Argentino (SIPA). De esta forma, el análisis desarrollado en este trabajo contempla únicamente el espectro de trabajadores registrados. El período analizado corresponde a noviembre de 2021. Adicionalmente, se utiliza la variable de densidad poblacional extraída del Censo de Población y Vivienda 2010.

Uno de los principales aportes de este trabajo es la georreferenciación de todas las relaciones laborales formales de residentes de CABA que trabajan en la ciudad. Para esto se reconstruyeron las relaciones laborales a partir de información provista por la Administración Federal de Ingresos Públicos (AFIP) respecto a la sucursal de explotación y al domicilio declarado por el empleador (provincia, localidad, código postal, calle y altura donde reside). De esta manera, se utilizó la aplicación abierta de Open Street Map (una base de datos geográfica colaborativa) que permitió identificar de forma precisa la mayor parte de las direcciones informadas. Para el resto de las direcciones se utilizó el servicio de normalización y georreferenciación de calles de ArcGIS.

En base a los datos georreferenciados se construyó la medida de distancia de *commuting* de las personas al trabajo. Esta distancia fue estimada mediante la medida de distancia rectilínea o "Manhattan" entre la manzana de residencia y la manzana de trabajo. Esta distancia se calcula sumando las longitudes de las proyecciones sobre el eje Y y el eje X del segmento que se traza entre los puntos de llegada y de destino. En otras palabras, se toma la distancia que se requiere para llegar de un punto a otro si solamente fueran posibles movimientos verticales u horizontales (y no en diagonal, como en la distancia euclídea).

Por otro lado, se creó una variable de aglomeración de puestos de trabajo que captura la cantidad de personas que trabajan en un punto dado o "cerca" de él. Para esto, se calculó la cantidad de personas que trabajan en cada manzana (P_i); luego, a este número se le suma la cantidad de personas empleadas en las demás manzanas de la ciudad ponderadas por la inversa de su distancia en hectómetros (d_j). De esta manera, todo individuo que trabaja a una distancia de una cuadra (un hectómetro) de i es contabilizado como un trabajador de la manzana i, mientras que otro individuo que trabaja a dos cuadras de i es contabilizado como medio trabajador de la manzana i.

$$A_i = P_i + \sum_{j \neq i} \frac{P_j}{d_j}$$

De forma similar, para cada manzana se creó una variable que mide la densidad de flujo de colectivos en horas pico de la mañana (promedio horario de la franja de 7 a.m. a 11 a.m.). La construcción de la variable es análoga a la de aglomeración: para cada manzana se suma el promedio horario de la cantidad de colectivos que frenan en las paradas de esa manzana más el promedio horario de los colectivos que frenan en las paradas que se encuentran dentro de un radio de un kilómetro, ponderados por la inversa de su distancia en hectómetros.[1] La decisión de realizar el corte en un kilómetro responde a que, según la Encuesta de Movilidad Domiciliaria de 2018, se puede verificar que a más del 99% de los viajes en transporte público se accede caminando menos de 10 cuadras.

Adicionalmente, se construyó una variable de precio promedio ofrecido del metro cuadrado de viviendas en la Ciudad. Para esto utilizamos técnicas de *web scrapping* que permitieron obtener los precios de todas las casas y departamentos publicados a abril de 2022 en los portales de ZonaProp, Properati y Mercado Libre.

[1] Datos Abiertos del GCBA.

Finalmente, no es trivial la elección de la unidad de análisis seleccionada. Si bien la base de datos utilizada en este trabajo permite un análisis a nivel individuos, existe un claro *trade-off* entre la granularidad de los datos y la eficiencia computacional e interpretativa al momento de analizarla. Por estos motivos, aquí se utilizan principalmente radios censales y manzanas para las visualizaciones y solamente manzanas para las estimaciones econométricas. Los radios censales son la mínima unidad geoestadística provista en los Censos Nacionales por el Instituto Nacional de Estadística y Censos de la República Argentina (INDEC). En las ciudades, los radios censales agrupan en promedio 300 viviendas. La utilización de radios tiene además dos ventajas adicionales: (i) permite compatibilizar los datos con los surgidos a partir de los Censos Nacionales, (ii) no pierde el grado de granularidad suficiente como para realizar interpretaciones precisas.

Por último, la desagregación sectorial se realiza en base a la letra del Clasificador de Actividad Económica (CLAE), lo que da lugar a 17 actividades productivas diferentes.

3. Marco teórico y resultados descriptivos

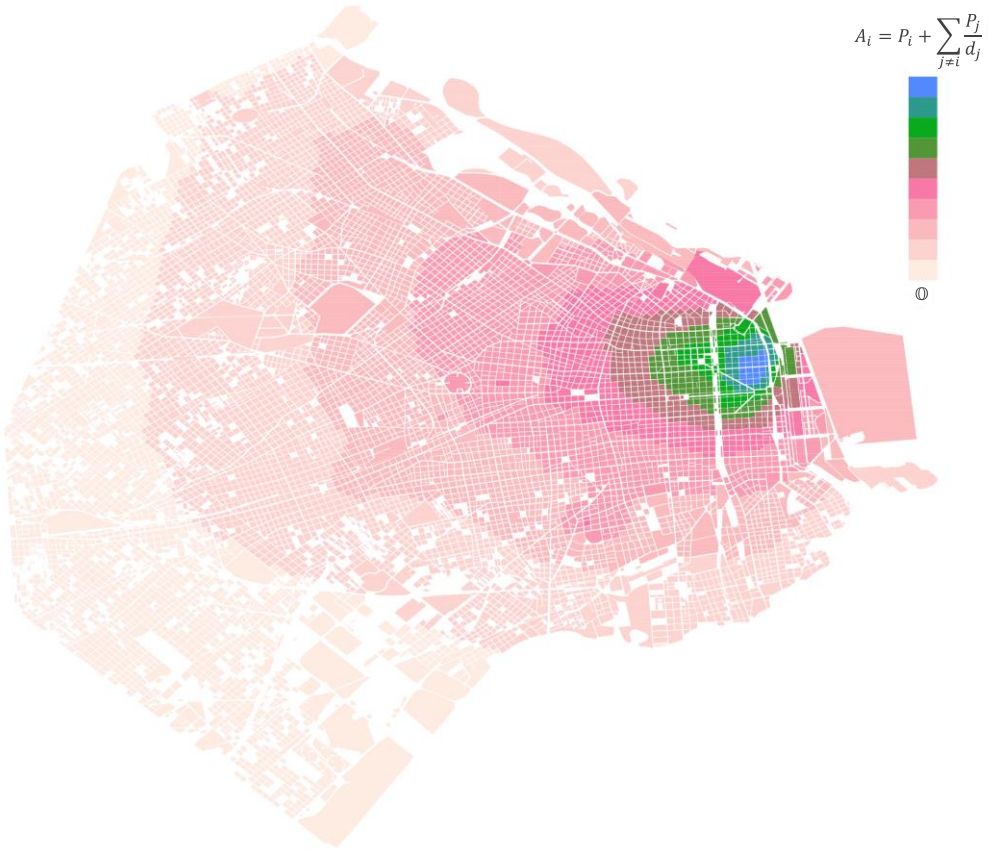
Buenos Aires fue fundada sobre la ribera derecha del Río de La Plata y fue concebida con un diseño simétrico respecto al eje de la Plaza Mayor, hoy Plaza de Mayo. Sin embargo, durante el siglo XIX algunos factores fueron sesgando el crecimiento de la ciudad orientado hacia el norte. A nivel económico, un hecho importante fue el establecimiento de la actividad financiera por encima de lo que hoy es la avenida Rivadavia. Sin embargo, el abandono del sur de la ciudad por parte de la clase aristocrática durante la epidemia de la fiebre amarilla, termina de configurar el sesgo negativo sobre el sur de la ciudad (Scobie,1986). Actualmente, la actividad económica conserva este sesgo y continúa fuertemente concentrada sobre el norte del centro histórico (mapa 1). Si bien la posición fundacional puede haber sido elegida sobre la base de ventajas naturales que ya no son tan relevantes actualmente, las razones para pensar que la dependencia de la trayectoria tiene un papel preponderante no son menores. Principalmente, la infraestructura que implica el trazado urbano de una ciudad supone un costo hundido difícil de ignorar. Sin embargo, la configuración de los patrones urbanos que adquiere la ciudad hoy en día son el resultado de intrincadas dinámicas económicas que determinan las decisiones de asignación de recursos y el grado de concentración de la actividad.

La literatura más relevante sobre estos patrones geográficos es probablemente la Nueva Geografía Económica (NEG). Krugman (1991a, 1991b y 1992), como uno de sus mayores referentes, sostiene que cualquier modelo de economía geográfica debe presentar una tensión entre fuerzas "centrípetas", que tienden a concentrar la actividad económica, y fuerzas "centrífugas", que tienden a dispersarla. De esta forma, el grado de concentración urbana es resultado de la interacción entre estas tendencias contrapuestas. Dentro del primer tipo de fuerzas se encuentran las llamadas economías de escala y economías de aglomeración. Mientras que la primera de estas opera a nivel interno de cada firma, la segunda opera a partir de la interacción entre ellas y su entorno. De acuerdo con este marco teórico, los beneficios derivados de la concentración espacial son, fundamentalmente, la reducción en los costos de producción y de transacción, la mayor oferta de insumos intermedios, el mejor *matching* laboral y el mayor intercambio y difusión de tecnologías. Por otro lado, dentro del segundo tipo de fuerzas se destaca, como principal desventaja de la concentración, el aumento en los costos de vida y otras externalidades negativas derivadas de la congestión.

Dentro de este marco, la economía urbana incluye una amplia gama de modelos que explican el uso del suelo en las ciudades a partir de estas tensiones. En el modelo canónico de Alonso-Mills-Muth (Alonso, 1964; Becker, 1965; Mills, 1967; Mills y Hamilton,1995; Muth, 1968) de ciudades monocéntricas, el centro se determina exógenamente por la localización de la producción. Esta decisión tiene implícitos muchos

factores vinculados a las economías de escala y de aglomeración. La otra fuerza centrípeta del modelo está asociada a los costos de transporte, que incentivan a los ciudadanos a vivir cerca del centro productivo. De esta forma, la menor demanda de tierras apartadas empuja a disminuir sus precios para que estas resulten atractivas. Cabe destacar que, en las diferentes estrategias de modelización, el componente más importante del costo de transporte no es el costo monetario, sino el costo de oportunidad por el tiempo no destinado al trabajo u ocio, lo que es especialmente cierto en Buenos Aires en donde la tarifa de transporte público es menor a USD 0,20. De esta manera, se observa que la mayor parte de la población en Buenos Aires se encuentra en torno al centro productivo y, especialmente, en torno a las líneas de subterráneo que acortan dramáticamente los tiempos de viaje al centro (mapa 2).

Mapa 1. Aglomeración de la actividad económica

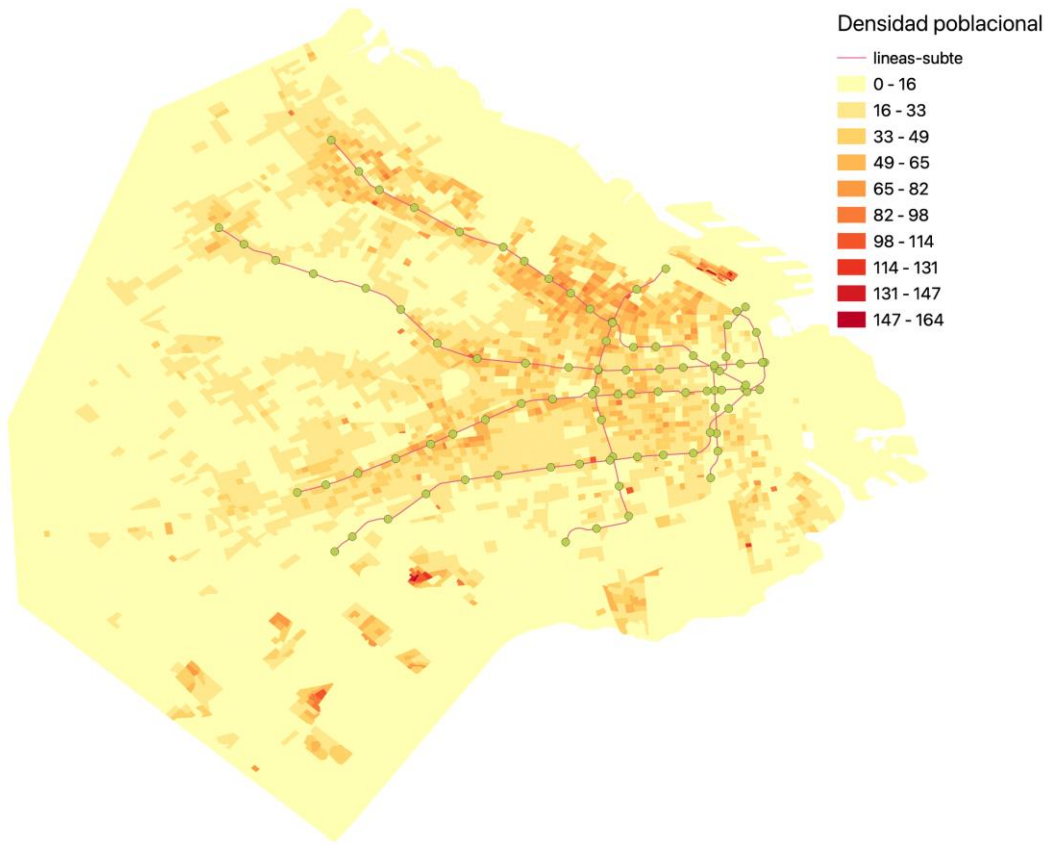


Nota: el gráfico ilustra la variable de aglomeración de puestos de trabajo formales en intervalos iguales. Los valores exactos de la escala fueron omitidos debido a que la unidad no es intuitiva. Fuente: elaboración propia con base en datos provenientes de SIPA y AFIP.

Sin embargo, esto no logra explicar algunos de los patrones que se observan actualmente en la ciudad. Buenos Aires presenta una fuerte segregación por ingresos, en la que el tercio de la población con menores ingresos vive en zonas muy diferentes de las del resto de la población (mapa 3). En línea con esto, LeRoy & Sonstelie (1983) generalizan el modelo AMM incluyendo agentes con distintos niveles de ingreso y varios medios de transporte. Estas extensiones del modelo permiten explicar por qué los trabajadores más pobres pueden situarse más cerca de los centros urbanos, dependiendo de la disponibilidad de transporte público, mientras que los más ricos, que tienen la posibilidad de viajar en medios de transporte más rápidos (aunque más costosos), se alejan, dado que tienen una mayor

demanda de tierra, que es más barata en zonas alejadas del centro. Sin embargo, si el costo fijo del medio de transporte privado es lo suficientemente alto, una parte de los ricos puede decidir vivir cerca del centro de la ciudad.

Mapa 2. Aglomeración de la actividad económica

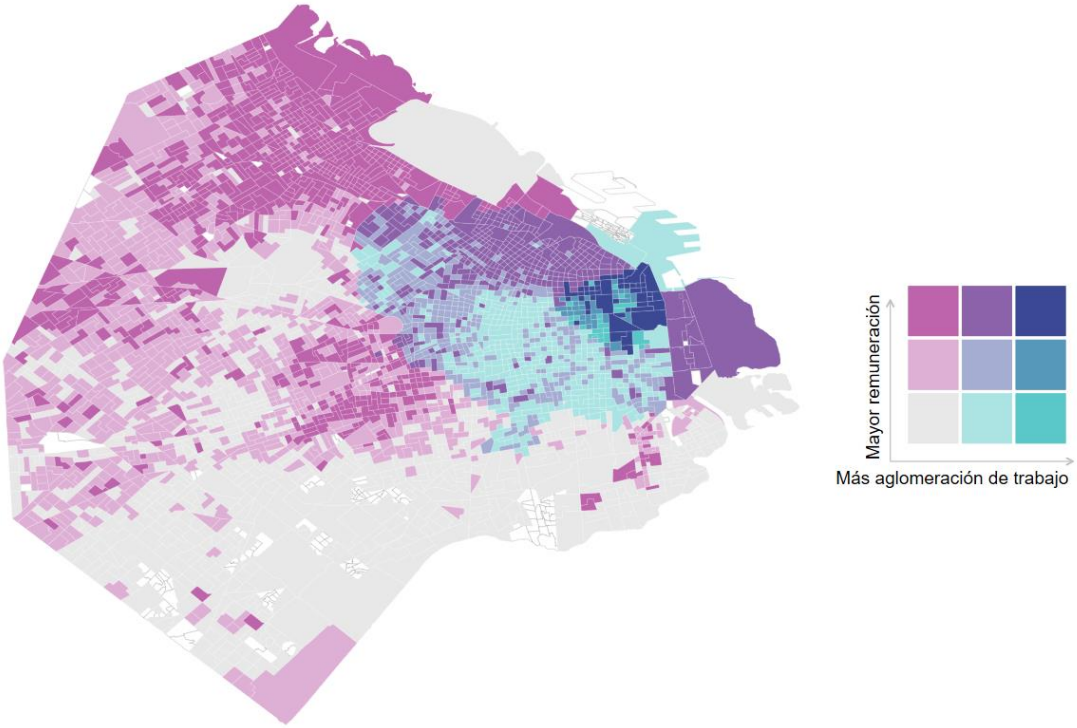


Fuente: elaboración propia con base en datos provenientes de SIPA y AFIP.

En particular, en la Ciudad de Buenos las zonas de mayor aglomeración de la actividad cuentan con mayor proporción de individuos de altos ingresos. Para los otros niveles de aglomeración, la segregación por ingresos es aún más clara, con zonas mejor diferenciadas.[2] El cuadro 1 muestra más claramente cómo se cumplen las predicciones del modelo: mayor proporción de altos ingresos en torno al centro de actividad y mayor proporción de bajos ingresos en las zonas más alejadas. Sin embargo, el mecanismo que plantea el modelo pareciera al menos no ser el único. En las zonas más antiguas de la ciudad, generalmente las propiedades se ofrecen a un precio mayor, mientras que tienden a ser más baratas en la zona sur y suroeste de la ciudad, en línea con la hipótesis aquí planteada sobre el peso de la dependencia de la trayectoria histórica (mapa 4). Por otra parte, las zonas donde el precio de las propiedades es más alto generalmente coinciden también con zonas en las que hay una mejor infraestructura de transporte. De esta manera, es posible que el precio de las propiedades esté reflejando las diferencias en la calidad del servicio y que, por tanto, esta sea la forma en que los ricos "compran" mejor transporte público, aunque explorar esta hipótesis excede los límites de este trabajo.

[2] Glaeser *et al.* (2008) encuentran patrones similares para ciudades de Estados Unidos.

Mapa 3. Aglomeración de la actividad y remuneración, por radio censal



Nota: para la remuneración se utilizaron los terciles de la distribución empírica y para la aglomeración intervalos de igual longitud.
Fuente: elaboración propia con base en datos provenientes de SIPA y AFIP.

Cuadro 1. Porcentaje de personas en cada categoría

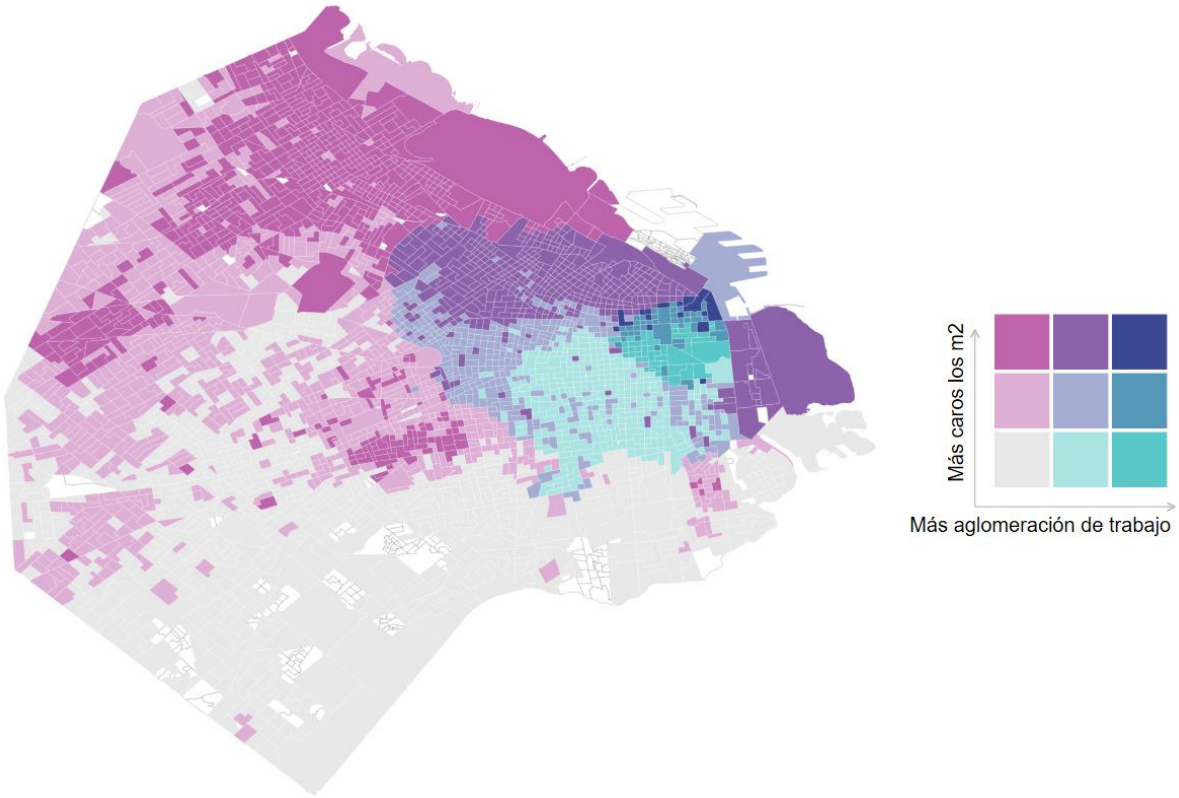
		Aglomeración		
		Baja	**Media**	**Alta**
Remuneración	**Alta**	11,57%	10,90%	10,86%
	Media	15,96%	11,21%	6,16%
	Baja	18,24%	10,04%	5,06%

Nota: para la remuneración se utilizaron los terciles de la distribución empírica y para la aglomeración intervalos de igual longitud.
Fuente: elaboración propia con base en datos provenientes de SIPA y AFIP.

Existen muchos factores que pueden explicar esta tendencia. En el modelo de Arnott y Stiglitz (1979), el valor de las propiedades depende de los bienes públicos que provee la ciudad. No obstante, hay otras variables que alejan la dinámica real de la ciudad de los esquemas estilizados. Como se vio en el mapa 1, los puestos de trabajo no se encuentran completamente concentrados en un punto, sino que el nivel de actividad desciende gradualmente desde el centro hacia el resto de la ciudad. Esto genera otra tensión similar pero diferente, en la que los trabajadores prefieren trabajar lo más cerca posible de donde viven pero a la vez están dispuestos a trasladarse más a cambio de un mayor ingreso. De esta forma, en las zonas en las que se concentra la mayor actividad económica se ofrecen mayores salarios para atraer así la suficiente oferta de fuerza de trabajo. No menos importante, los empleadores deben pagar mayores

salarios para compensar a los trabajadores locales por los mayores costos de vida derivados de la mayor congestión.

Mapa 4. Aglomeración de la actividad y precios del m^2, por radio censal

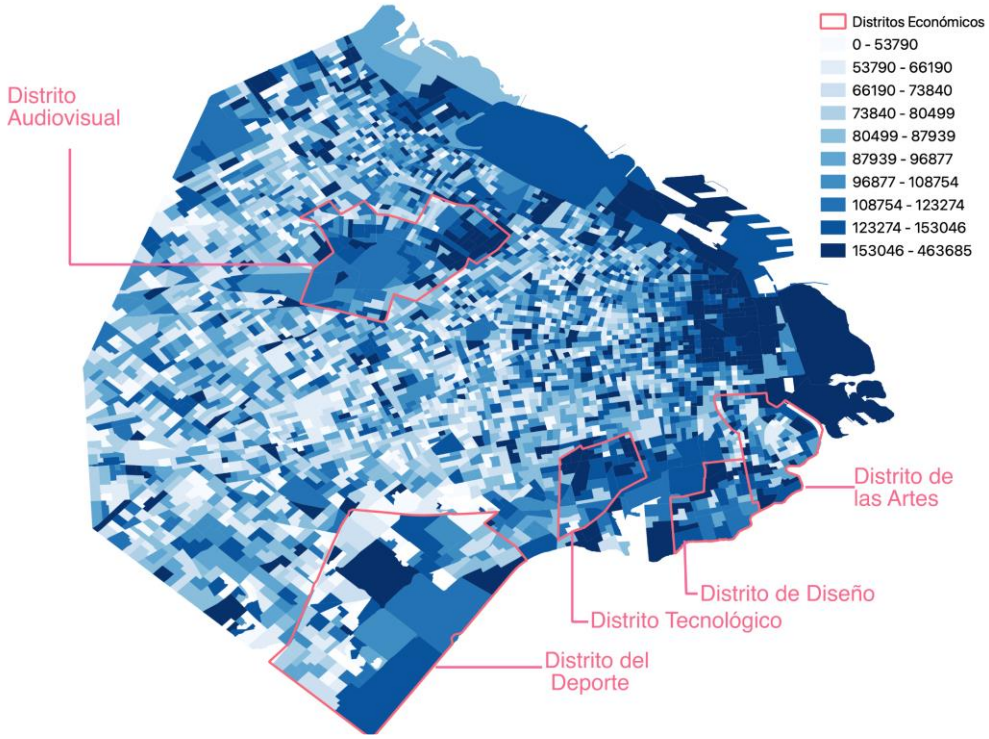


Nota: para los precios se utilizaron terciles y para la aglomeración intervalos iguales.
Fuente: elaboración propia con base en datos provenientes de SIPA y de ZonaProp, Properati y Mercado Libre.

En Buenos Aires, los mayores salarios se pagan en el centro productivo pero también en los diferentes "distritos económicos" (mapa 5). Los distritos económicos pueden ofrecer salarios muy competitivos, a pesar de estar relativamente aislados, por dos motivos. En primer lugar, porque la aglomeración a nivel industria genera importantes externalidades positivas que van más allá de la presencia de otras industrias, por ejemplo la generación de proveedores específicos o la baja en los costos logísticos. En segundo lugar, porque las empresas, a cambio de encontrarse relativamente aisladas, reciben importantes exenciones y financiamiento subsidiado a la vez que pagan costos bajos en el uso del suelo (Carmona, 2017).

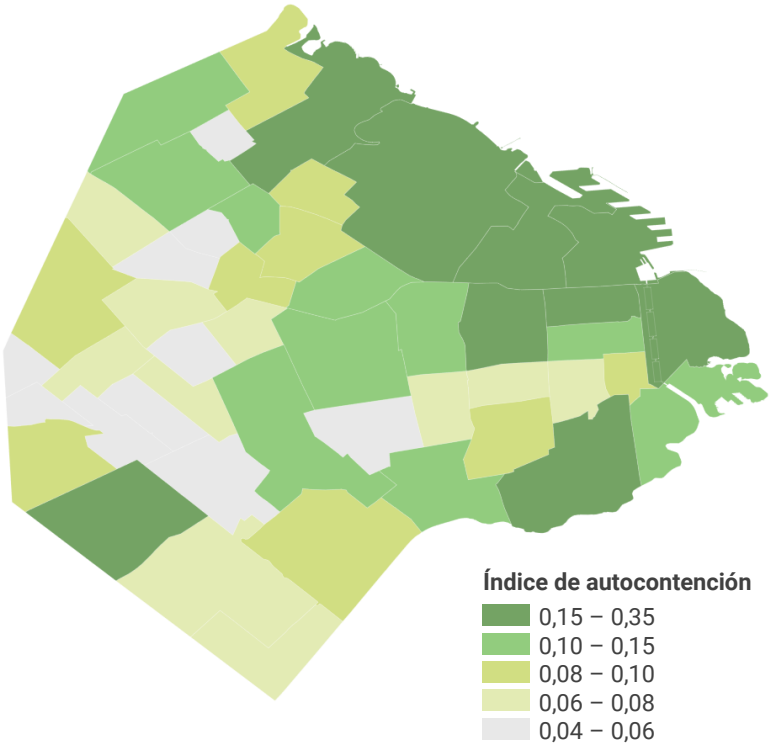
También es interesante analizar el fenómeno inverso a la distancia de *commuting*, el llamado nivel de autocontención, es decir, la cantidad de personas que trabajan en el mismo barrio en el que residen. El mapa 6 muestra que los barrios con mayor capacidad de autocontención del empleo formal se encuentran mayormente en el centro y norte de la ciudad, donde la actividad económica a su vez es mayor: San Nicolás (35,4%), Puerto Madero (31,9%), Retiro (25,5%), Palermo (21,2%) y Recoleta (19,3%). Que la autocontención esté positivamente asociada al nivel de ingreso también es un resultado consistente con el modelo, en la medida en que los trabajadores de mayores ingresos son quienes pueden pagar el costo de la tierra ubicada en zonas cercanas a los principales núcleos de aglomeración.

Mapa 5. Remuneración y distritos económicos



Nota: remuneración mediana en pesos por radio censal para noviembre de 2021. Fuente: elaboración propia con base en datos provenientes de SIPA y AFIP.

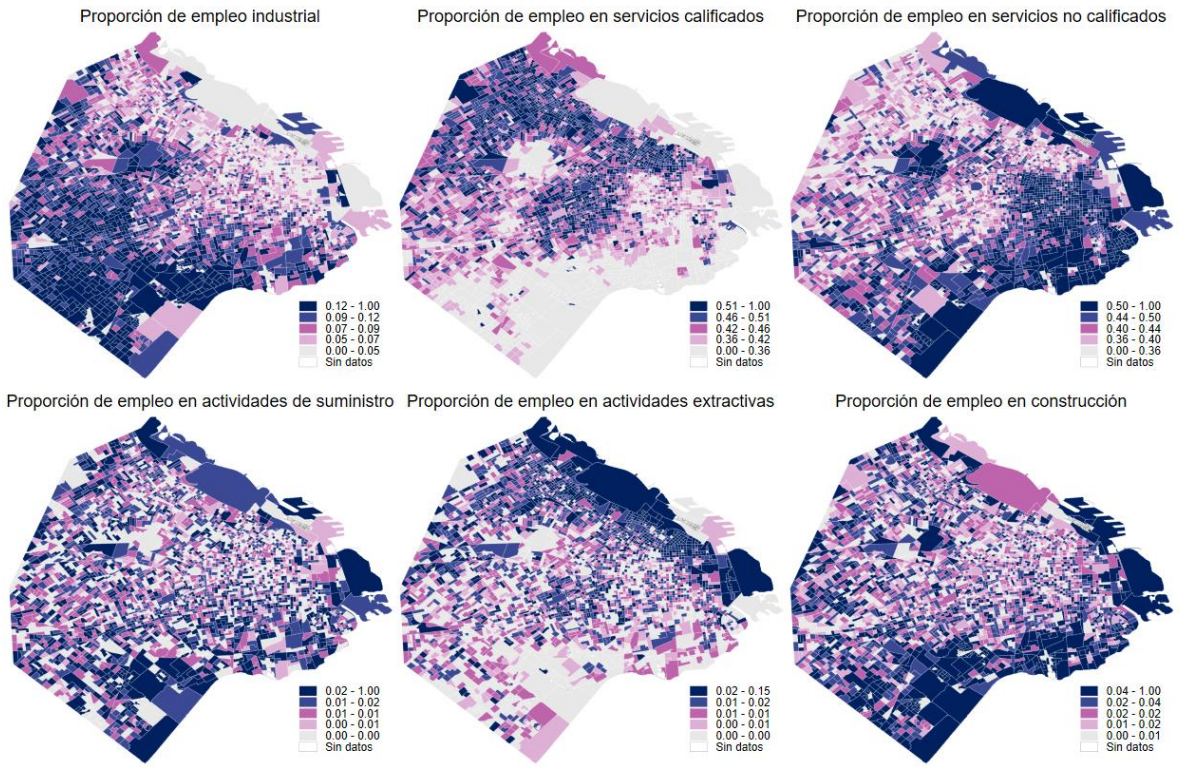
Mapa 6. Medida de autocontención, por barrios



Fuente: elaboración propia con base en datos provenientes de SIPA y AFIP.

Por otro lado, dado que el sector productivo en que cada trabajador se inserta explica una proporción relevante de la variabilidad salarial, es esperable que trabajadores de una misma rama decidan vivir en las mismas zonas de la ciudad. También es posible que exista cierto grado de transmisión intergeneracional de habilidades productivas entre padres e hijos que reproduzca patrones de este tipo en el tiempo. El mapa 7 muestra que los trabajadores de servicios calificados y de actividades extractivas residen mayormente en las zonas de mayores ingresos, mientras que la industria manufacturera, los servicios no calificados y la construcción tienen mayor proporción de trabajadores ubicados en el sur de la ciudad (principalmente, sudoeste en el primer caso, sudeste en el último).

Mapa 7. Empleo por sectores, según radio censal de residencia



Letras en industria: C; en servicios calificados: J, K, M, P y Q; en servicios no calificados: G, H, I, L, R y S; en actividades de suministro: D y E; en actividades extractivas: A y B; y en construcción: F.
Fuente: elaboración propia con base en datos provenientes de SIPA y AFIP.

4. Metodología

Explotando la variabilidad a nivel de manzana, en esta sección se estiman modelos de regresión que contemplan la existencia de autocorrelación espacial. La ventaja de esta estrategia de modelización es que permite la posibilidad de que existan efectos derrame entre las distintas unidades geográficas (en el caso de este trabajo, entre las distintas manzanas). Para incluir estos potenciales efectos, se estiman diferentes versiones del siguiente modelo:

$$y = \delta W y + X\beta + WZ\theta + \lambda W u + \varepsilon \qquad (1)$$

Donde y es el vector que contiene los valores de la variable dependiente, X la matriz que contiene en sus columnas todas las variables explicativas (incluyendo una ordenada al origen), Z la matriz que contiene las variables explicativas para las cuales se evalúan los derrames espaciales, W la matriz de ponderaciones espaciales, y δ, λ y θ los parámetros (o vector de parámetros, en el último caso) que capturan el tamaño de los efectos derrame asociados a la variable dependiente, el término de error y las variables explicativas, respectivamente; u es el vector de errores que sufren correlación espacial y ε es un vector de errores tradicional. La matriz de ponderaciones fue construida en base al criterio de la inversa de la distancia. Siguiendo la convención, esta matriz es diagonalizada de modo tal que cada fila sume uno.

El modelo es estimado mediante dos técnicas: el estimador de Mínimos Cuadrados en Dos Etapas (MC2E), que instrumenta los rezagos espaciales de la variable dependiente con los de la variable independiente, y el estimador de Máxima Verosimilitud (MV), que obtiene los parámetros estimados maximizando el logaritmo de la función de verosimilitud de la forma tradicional (lo que implica incorporar el supuesto de normalidad de ε). Cabe destacar que los resultados de este trabajo no deben interpretarse de forma causal, sino más bien como correlaciones parciales, debido a que probablemente estas estimaciones sufran de endogeneidad derivada de la omisión de variables relevantes y de la potencial causalidad bidireccional entre las variables.

En primer lugar, se utiliza como variable dependiente la remuneración bruta promedio de los trabajadores formales que residen en cada manzana. Lo que se busca con este primer análisis es arrojar luz acerca de los determinantes del salario que reciben los trabajadores en una determinada área de residencia. Tomando como referencia la ecuación de Mincer tradicional, las variables independientes utilizadas en este modelo son la edad promedio de los trabajadores residentes de la manzana (lineal y cuadrática), su experiencia laboral formal (lineal y cuadrática), la proporción de mujeres, la participación de los distintos sectores productivos y la proporción de empleo calificado.

Adicionalmente, se incorporan otras variables explicativas de tipo espacial. En primer lugar, la distancia de *commuting* promedio realizada por los trabajadores. En términos de la NEG, la relación teórica entre la distancia de *commuting* y los ingresos no está definida ya que, por un lado, existe tensión entre fuerzas que van en sentidos contrapuestos y, por otro lado, esta relación es presumiblemente no lineal en la distancia. En sentido positivo, las personas deberían estar dispuestas a viajar más por un mayor salario y a vivir más lejos de los centros productivos para acceder a viviendas más baratas. En sentido contrario, los costos de transporte, incluyendo el costo de oportunidad del tiempo utilizado para movilizarse, suponen incentivos a las personas de mayores ingresos a pagar precios más altos para vivir cerca del trabajo. Es decir que no es posible anticipar qué signo debería tener la relación entre distancia de *commuting* e ingresos. En segundo lugar, se incorpora la distancia promedio a la estación de subterráneo más cercana como una medida del acceso a servicios de transporte público. A su vez, se incluye una medida de frecuencia del servicio de colectivos en la manzana y su cuadrado. La idea detrás de estas variables es que una baja frecuencia dificulta la movilidad dentro de la ciudad, lo que obstaculiza el acceso al centro productivo; sin embargo, una frecuencia de colectivos demasiado alta está relacionada con una mayor congestión del tránsito en estas zonas de la ciudad. También se incluye la densidad poblacional,[3] en tanto es esperable que mayor congestión (que implica mayores costos de vida) sea compensada con mayores salarios. Finalmente, se incorpora una medida de concentración del empleo que está estrechamente relacionada con la distancia de los hogares al centro económico de la ciudad. De ahora en adelante, esta especificación del modelo será llamada Modelo 1.

[3] La variable de densidad es extraída del Censo Nacional de Población y Vivienda (2010) del INDEC. Debido a la falta de datos a nivel de manzana, se utiliza la variable agregada a nivel de radio censal.

En segundo lugar, se utiliza como variable explicada la remuneración bruta promedio de los trabajadores formales que llevan a cabo su actividad productiva en cada manzana. Este modelo (que es el espejo del anterior) apunta a determinar cuáles son las variables que explican el salario promedio que reciben los trabajadores de acuerdo a la zona en la que se desempeñan laboralmente. Nuevamente, se incorporan todas las típicas variables de Mincer. Dentro de las variables espaciales, se incluye la distancia de *commuting* promedio que efectúan los empleados que trabajan en la manzana, la distancia promedio al subte en dicha manzana, la medida de frecuencia de colectivos y su cuadrado, la densidad y, finalmente, la medida de aglomeración. En esta especificación del modelo, se espera que los efectos de las variables de densidad y de aglomeración reflejen las potenciales ganancias de la concentración espacial propuestas por la NEG (economías de escala y economías de aglomeración). De ahora en adelante, esta especificación del modelo será llamada Modelo 2.

5. Resultados

La estrategia estándar en la mayoría de los análisis espaciales es comenzar con un modelo de regresión lineal no espacial y testear, a partir de sus residuos, si es necesario incorporar interacciones espaciales entre las unidades. En este contexto, el primer paso consiste en evaluar la presencia de autocorrelación espacial en los residuos del modelo estimado por MCO. Para eso se aplica el test de Moran, cuyos resultados se resumen en el cuadro 2.

Cuadro 2. Test de Moran sobre los residuos de MCO

	Modelo 1	Modelo 2
Estadístico	10.724,79	568,62
P-valor	0,000	0,000

Nota: la hipótesis nula es que los errores son i.i.d.
Fuente: Elaboración propia con base en datos provenientes de SIPA y AFIP.

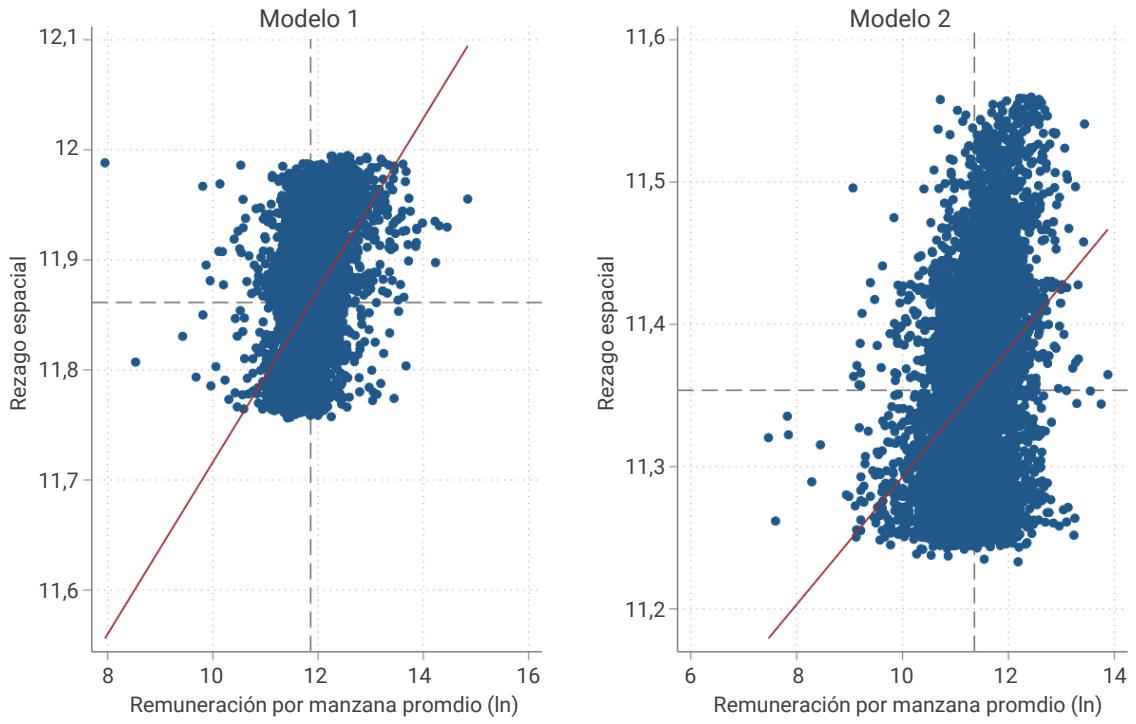
La hipótesis nula de no correlación espacial se rechaza, para cualquier nivel relevante de confianza, en ambos modelos. Estos resultados avalan la hipótesis de que los términos de error asociados a manzanas cercanas entre sí se encuentran correlacionados. Esto es evidencia a favor del uso de una modelización de tipo espacial.

Una forma alternativa de visualizar este fenómeno de dependencia geográfica es usar lo que se conoce como un diagrama de Moran, en el que se representa gráficamente la correlación entre la variable dependiente y su rezago espacial, es decir, la misma variable premultiplicada por la matriz W.

Se observa una considerable correlación espacial positiva entre la remuneración (ya sea promediada por manzana de residencia o por manzana de trabajo) y su rezago espacial, que es deseable incorporar al modelo de forma explícita.

Finalmente, para determinar la especificación del modelo se recurre a un conjunto de tests conocidos en la literatura como tests LM robustos, que sirven para detectar la presencia de autocorrelación espacial tanto en la variable dependiente como en el término de error.

Gráfico 1. Diagrama de Moran



Fuente: elaboración propia con base en datos provenientes de SIPA y AFIP.

Cuadro 3. Test de LM robustos

	Modelo 1	Modelo 2
Hipótesis nula: no existe autocorrelación espacial en la variable dependiente		
Estadístico	606,13	23,58
P-valor	0,000	0,000
Hipótesis nula: no existe autocorrelación espacial en el término de error		
Estadístico	8499,23	316,40
P-valor	0,000	0,000

Nota: el estadístico que se reporta es el Multiplicador de Lagrange Robusto (MLR).
Fuente: Elaboración propia con base en datos provenientes de SIPA y AFIP.

Los resultados de los test robustos permiten rechazar tanto la hipótesis nula de no autocorrelación espacial en el término de error como la correspondiente a la variable dependiente. Dado que no existen test análogos que permitan evaluar la existencia de autocorrelación espacial en las variables independientes, en este trabajo se opta por presentar los resultados de dos versiones distintas del modelo espacial:

- Modelo espacial general (habitualmente conocido como GNS, *general nested spatial model*, o modelo de Cliff-Ord): representado por la ecuación (1). Este modelo contempla la existencia de autocorrelación espacial en la variable dependiente, en el término de error y en ciertas variables explicativas. La variable explicativa incluida en la matriz Z es la aglomeración. Esta elección responde a que, dado el marco teórico discutido anteriormente, la aglomeración resulta ser la variable en la que se espera observar efectos derrame: el salario observado en una manzana factiblemente está influido por el nivel de aglomeración observado en ella, pero también por el observado en las manzanas cercanas.

- Modelo SAC (*spatial autocorrelation model*): representado por la ecuación (1), pero imponiendo la restricción $\theta = 0$. Este modelo solo contempla la posibilidad de autocorrelación en la variable dependiente y en el término de error.

Los modelos espaciales permiten calcular, para cada variable explicativa, un efecto directo (el efecto de un cambio marginal en el regresor sobre la variable dependiente de la propia unidad espacial, ignorando los efectos derrame), un efecto indirecto (el efecto de un cambio marginal en el regresor sobre la variable dependiente producido exclusivamente por los derrames geográficos) y un efecto total (la suma de los dos efectos anteriores).

Para comenzar, se analizan los resultados de la estimación de un modelo simple de regresión lineal no espacial, en dos especificaciones diferentes, según se incluyan o no efectos fijos por comuna. A su vez, se contemplan diversas estrategias para lidiar con la posible heterocedasticidad: errores estándar tradicionales, robustos a heterocedasticidad de forma desconocida, robustos y clusterizados a nivel de comuna y, finalmente, una estimación por Mínimos Cuadrados Ponderados (MCP) en que el ponderador es definido como la cantidad de trabajadores formales de la manzana. Los resultados de estas primeras estimaciones se presentan en los cuadros 4 y 5.

Cuadro 4. Modelo simple de regresión lineal no espacial: Modelo 1

	Base	Con efectos fijos por comuna	Con ponderadores por tamaño
Distancia de *commuting* (ln)	0,154 (0,013)*** [0,028]*** {0,070}**	0,104 (0,013)*** [0,028]*** {0,062}	0,237 (0,015)***
Aglomeración (ln)	0,346 (0,019)*** [0,029]*** {0,042}***	0,177 (0,026)*** [0,035]*** {0,064}**	0,348 (0,017)***
Distancia al subte (ln)	0,005 (0,005) [0,005] {0,017}	-0,010 (0,006)* [0,006]* {0,015}	0,005 (0,004)

Continúa en la página siguiente.

Cuadro 4. Modelo simple de regresión lineal no espacial: Modelo 1 (continuación)			
	Base	**Con efectos fijos por comuna**	**Con ponderadores por tamaño**
Frecuencia de colectivos (ln)	0,916 (0,108)*** [0,123]*** {0,251}***	0,835 (0,111)*** [0,124]*** {0,337}**	0,791 (0,104)***
Frecuencia de colectivos al cuadrado (ln)	-0,069 (0,008)*** [0,009]*** {0,019}***	-0,062 (0,008)*** [0,009]*** {0,027}**	-0,059 (0,007)***
Densidad (ln)	0,027 (0,004)*** [0,007]*** {0,013}**	0,018 (0,004)*** [0,007]*** {0,009}*	0,034 (0,003)***
Prop. de mujeres	-0,243 (0,024)*** [0,043]*** {0,052}***	-0,239 (0,023)*** [0,041]*** {0,051}***	-0,321 (0,030)***
Edad (años)	0,109 (0,011)*** [0,025]*** {0,040}**	0,115 (0,010)*** [0,024]*** {0,040}**	0,050 (0,021)**
Edad al cuadrado	-0,001 (0,000)*** [0,000]*** {0,000}**	-0,001 (0,000)*** [0,000]*** {0,000}**	-0,000 (0,000)
Experiencia formal (años)	0,033 (0,015)** [0,034] {0,050}	0,037 (0,014)*** [0,033] {0,048}	0,069 (0,028)**
Experiencia formal al cuadrado	-0,000 (0,001) [0,002] {0,002}	-0,000 (0,001) [0,002] {0,002}	-0,002 (0,001)
Prop. de empleo de baja calificación	-0,509 (0,039)*** [0,064]*** {0,066}***	-0,455 (0,038)*** [0,060]*** {0,087}***	-0,481 (0,050)***

Continúa en la página siguiente.

Cuadro 4. Modelo simple de regresión lineal no espacial: Modelo 1 (continuación)			
	Base	**Con efectos fijos por comuna**	**Con ponderadores por tamaño**
Prop. de empleo de alta calificación	0,253 (0,026)*** [0,047]*** {0,056}***	0,170 (0,025)*** [0,045]*** {0,043}***	0,578 (0,031)***
Prop. servicios calificados	0,529 (0,039)*** [0,063]*** {0,076}***	0,385 (0,038)*** [0,066]*** {0,072}***	0,687 (0,048)***
Prop. servicios no calificados	-0,028 (0,039) [0,064] {0,055}	-0,076 (0,037)** [0,064] {0,048}	-0,124 (0,049)**
Prop. actividades extractivas	2,564 (0,109)*** [0,317]*** {0,515}***	2,155 (0,106)*** [0,296]*** {0,575}***	3,988 (0,126)***
Prop. actividades de suministro	0,814 (0,094)*** [0,121]*** {0,117}***	0,752 (0,090)*** [0,116]*** {0,114}***	0,923 (0,138)***
Prop. construcción	-0,186 (0,072)*** [0,123] {0,060}***	-0,234 (0,069)*** [0,120]* {0,067}***	-0,290 (0,079)***
Prop. administ. pública	0,343 (0,236) [0,384] {0,356}	0,263 (0,226) [0,344] {0,313}	0,300 (0,269)
N	9452	9452	9452
R2	0,361	0,416	0,537

Nota: Categorías base: proporción de personas que trabajan en industria y proporción de personas que tienen un empleo de media calificación.
***Estadísticamente significativo al 99%, **al 95%, *al 90%; ( ) SE, [ ] SE robustos y { } SE robustos y clusterizados por comuna. Variable dependiente expresada en logaritmo natural.
Fuente: elaboración propia con base en datos provenientes de SIPA y AFIP.

En las tres especificaciones consideradas se encuentra una relación positiva y estadísticamente significativa entre la distancia de *commuting* y la remuneración: un aumento de un 1% en la distancia de *commuting* promedio realizada por los residentes de una manzana está asociado a un aumento de entre un 0,10% y un 0,24% en su remuneración promedio. Este resultado refleja que, en términos netos, predominan los factores que hacen que el salario crezca con la distancia de traslado, como el hecho de que las personas están dispuestas a viajar más por un mayor salario o que quienes tienen mayores salarios eligen vivir más lejos de los centros productivos para acceder a viviendas relativamente más baratas.

En lo que respecta al resto de las variables espaciales, se observa que vivir en una zona de alta aglomeración de trabajo se relaciona con una mayor remuneración promedio, lo cual es razonable considerando que las personas que viven en zonas de alta concentración de la actividad económica se benefician de la aglomeración (por ejemplo, con mejores oportunidades de trabajo, mejores *matches* laborales o posibilidad de inserción en firmas más productivas) sin necesidad de efectuar grandes distancias de *commuting*. La elasticidad que vincula las variables se ubica entre el 0,18% y el 0,35%. Sin embargo, la mayor densidad en estas zonas aumenta los costos de vida, lo que es consistente con la relación positiva entre la densidad poblacional y la remuneración promedio recibida por los habitantes. En otras palabras, los trabajadores que tienen mayor remuneración son los que pueden permitirse vivir en zonas de mayor congestión, que tienen costos más altos.

Finalmente, el coeficiente estimado para la variable de distancia al subte es positivo y estadísticamente significativo en solo una de las especificaciones, aunque con un efecto económicamente poco relevante. Esto se explica, posiblemente, por el hecho de que la cercanía al subte no es una medida informativa de la facilidad para movilizarse dentro de la ciudad. De acuerdo a la Encuesta de Movilidad Domiciliaria de 2018, el 33,4% de las etapas de los viajes se realizan en colectivo, el 26,4% a pie, el 23,2% en auto y solo un 3,5% de las etapas se efectúan a través del subte. En línea con estos datos, se observa que la medida de frecuencia de colectivos de una manzana sí tiene una relación relevante con la remuneración recibida por sus habitantes. Para bajos niveles de frecuencia, un aumento de esta variable está asociado a considerables aumentos en los salarios de los trabajadores. Sin embargo, dicha relación positiva alcanza un límite, a partir del cual aumentar la frecuencia de colectivos se asocia a una reducción en la remuneración promedio, probablemente debido a la congestión de tránsito que genera, lo que vuelve a la zona menos atractiva.

En cuanto a las variables de Mincer se observa que, al controlar por los años de experiencia laboral formal, la relación entre edad y salario presenta la típica forma de U invertida. En cuanto a la participación de las mujeres en la fuerza laboral formal, aquellas manzanas más feminizadas tienen una peor remuneración promedio, un resultado robusto a las distintas especificaciones. Finalmente, como es de esperarse, las manzanas con mayor participación de empleo calificado tienen, en promedio, mayores remuneraciones. También se observan patrones sectoriales conocidos: los salarios son, ceteris paribus, más altos para servicios calificados, actividades extractivas y de suministro y administración pública, mientras que son más bajos en los servicios no calificados y la construcción.

Cuadro 5. Modelo simple de regresión lineal no espacial: Modelo 2			
	Base	**RSE y efectos fijos por comuna**	**RSE y con ponderadores por tamaño**
Distancia de *commuting* (ln)	0,074 (0,008)*** [0,010]*** {0,018}***	0,060 (0,008)*** [0,010]*** {0,017}***	0,328 (0,014)***
Aglomeración (ln)	0,312 (0,022)*** [0,022]*** {0,057}***	0,210 (0,036)*** [0,036]*** {0,089}**	0,363 (0,014)***
Distancia al subte (ln)	0,003 (0,008) [0,008] {0,018}	-0,016 (0,009)* [0,009]* {0,011}	-0,011 (0,006)**
Frecuencia de colectivos (ln)	-0,106 (0,156) [0,144] {0,253}	-0,096 (0,167) [0,160] {0,311}	1,111 (0,102)***
Frecuencia de colectivos al cuadrado (ln)	0,007 (0,011) [0,010] {0,019}	0,005 (0,012) [0,011] {0,023}	-0,087 (0,007)***
Densidad (ln)	-0,034 (0,006)*** [0,007]*** {0,012}**	-0,029 (0,007)*** [0,007]*** {0,010}**	0,003 (0,004)
Dummy de distrito económico	0,051 (0,013)*** [0,014]*** {0,023}**	0,071 (0,017)*** [0,018]*** {0,026}**	0,148 (0,011)***
Prop. de mujeres	-0,120 (0,019)*** [0,023]*** {0,030}***	-0,117 (0,019)*** [0,023]*** {0,031}***	-0,423 (0,028)***

Continúa en la página siguiente.

Cuadro 5. Modelo simple de regresión lineal no espacial: Modelo 2 (continuación)			
	Base	**RSE y efectos fijos por comuna**	**RSE y con ponderadores por tamaño**
Edad (años)	0,051 (0,007)*** [0,010]*** {0,008}***	0,053 (0,007)*** [0,010]*** {0,008}***	-0,136 (0,014)***
Edad al cuadrado	-0,001 (0,000)*** [0,000]*** {0,000}***	-0,001 (0,000)*** [0,000]*** {0,000}***	0,001 (0,000)***
Experiencia formal (años)	0,100 (0,009)*** [0,013]*** {0,009}***	0,099 (0,009)*** [0,013]*** {0,008}***	0,292 (0,019)***
Experiencia formal al cuadrado	-0,002 (0,000)*** [0,001]*** {0,000}***	-0,002 (0,000)*** [0,001]*** {0,000}***	-0,008 (0,001)***
Prop. de empleo de baja calificación	-0,070 (0,024)*** [0,029]** {0,045}	-0,083 (0,024)*** [0,029]*** {0,042}*	-0,434 (0,033)***
Prop. de empleo de alta calificación	0,183 (0,020)*** [0,027]*** {0,040}***	0,178 (0,020)*** [0,026]*** {0,041}***	-0,189 (0,020)***
Prop. servicios calificados	0,140 (0,022)*** [0,026]*** {0,033}***	0,112 (0,022)*** [0,026]*** {0,028}***	0,217 (0,021)***
Prop. servicios no calificados	-0,080 (0,018)*** [0,021]*** {0,016}***	-0,094 (0,019)*** [0,021]*** {0,018}***	-0,167 (0,021)***

Continúa en la página siguiente.

Cuadro 5. Modelo simple de regresión lineal no espacial: Modelo 2 (continuación)			
	Base	**RSE y efectos fijos por comuna**	**RSE y con ponderadores por tamaño**
Prop. actividades extractivas	0,498 (0,078)*** [0,091]*** {0,138}***	0,450 (0,077)*** [0,091]*** {0,132}***	1,075 (0,062)***
Prop. actividades de suministro	0,681 (0,091)*** [0,074]*** {0,062}***	0,673 (0,090)*** [0,073]*** {0,065}***	0,275 (0,047)***
Prop. construcción	-0,258 (0,031)*** [0,040]*** {0,042}***	-0,270 (0,031)*** [0,040]*** {0,039}***	-0,232 (0,044)***
Prop. administ. pública	0,497 (0,206)** [0,128]*** {0,044}***	0,458 (0,205)** [0,117]*** {0,048}***	0,341 (0,090)***
N	9299	9299	9299
R2	0,278	0,288	0,599

Nota: Categorías base: proporción de personas que trabajan en industria y proporción de personas que tienen un empleo de media calificación.
***Estadísticamente significativo al 99%, **al 95%, *al 90%; ( ) SE, [ ] SE robustos y { } SE robustos y clusterizados por comuna.
Variable dependiente expresada en logaritmo natural.
Fuente: elaboración propia con base en datos provenientes de SIPA y AFIP.

Al agrupar por domicilio de trabajo, se observa una prima asociada a la distancia de *commuting* menor a la obtenida agrupando por domicilio de residencia. Un aumento de un 1% en la distancia de *commuting* promedio realizada por los trabajadores de una manzana está asociado a un aumento de entre un 0,06% y un 0,33% en su remuneración promedio. Este resultado es un "espejo" del discutido para el cuadro 4: a igualdad de otras características, una manzana que "paga" mayores remuneraciones va a lograr captar trabajadores que residen más lejos; si este efecto es el efecto predominante, entonces es razonable que la elasticidad obtenida sea positiva.

Por otro lado, se encuentra una relación positiva y significativa entre la aglomeración de una manzana y la remuneración promedio recibida por los trabajadores que allí se desempeñan. Una posible interpretación de este resultado es que la mayor demanda de trabajo, derivada de la mayor concentración de la actividad económica, podría estar obligando a los empleadores de estas manzanas a ofrecer mayores salarios con el fin de captar a trabajadores residentes de zonas más alejadas. La forma en que las firmas logran afrontar estos mayores costos laborales es a partir de la reducción de costos generada por las economías de aglomeración.

En cuanto a las variables vinculadas al transporte público, en dos de las tres especificaciones estimadas se obtiene un coeficiente negativo y estadísticamente significativo para la variable de distancia al subte, lo que sugiere que un empleo con peor acceso a este medio de transporte ofrece un menor salario, ceteris paribus. Por otro lado, solo se encuentra una relación estadísticamente significativa entre la remuneración pagada en una manzana y su frecuencia de colectivos cuando se utiliza MCP. En ese caso, la relación entre las variables es inicialmente positiva, lo cual se explica por la predominancia del efecto positivo derivado de la mayor facilidad para movilizarse dentro de la ciudad. Sin embargo, la relación se invierte para valores altos de frecuencia, lo que indica que, en este tramo del soporte, los costos de la mayor frecuencia de colectivos (mayor congestión del tránsito) sobrepasan sus beneficios.

Respecto de la densidad poblacional, en dos de las tres especificaciones se encuentra un efecto negativo y estadísticamente significativo sobre la remuneración pagada por los empleadores. Este resultado podría estar reflejando los mayores costos relativos a los que se enfrentan las firmas en las zonas más densas poblacionalmente, posiblemente debido a los costos adicionales asociados a la congestión.

Por su parte, se encuentra que los empleos ubicados en distritos económicos reciben salarios mayores al promedio en una magnitud de entre 5 y 15% dependiendo del método de estimación seleccionado. Esta ganancia puede ser atribuida a las economías de aglomeración particularmente fuertes en esas zonas (donde la concentración de unidades económicas pertenecientes a una misma rama productiva factiblemente implica considerables derrames tecnológicos) o bien a otros beneficios tales como créditos subsidiados y exenciones impositivas.

A continuación, se complejizan los modelos incluyendo rezagos especiales en base a las versiones GNS y SAC discutidas previamente. Los resultados se presentan en los cuadros 6 y 7.

Cuadro 6. Modelos espaciales: Modelo 1

Versión del modelo		**GNS**		**SAC**	
Método de estimación		**MC2E**	**MV**	**MC2E**	**MV**
Panel 1: Coeficientes estimados					
θ Aglomeración		-1,991*** (0,445)	0,394** (0,174)	-	-
δ		5,128*** (0,253)	3,574*** (0,026)	2,735*** (0,187)	4,701*** (0,046)
λ		2,154*** (0,117)	2,531*** (0,068)	2,749*** (0,299)	2,361*** (0,046)
N		9452	9452	9452	9452
Pseudo R2		0,107	0,043	-0,003	0,088
Panel 2: Efectos estimados					
Commuting (ln)	Efecto directo	0,088*** (0,029)	0,092*** (0,013)	0,086*** (0,027)	0,090*** (0,013)

Continúa en la página siguiente.

Cuadro 6. Modelos espaciales: Modelo 1 (continuación)					
Versión del modelo		**GNS**		**SAC**	
Método de estimación		**MC2E**	**MV**	**MC2E**	**MV**
Panel 2: Efectos estimados					
Commuting (ln)	Efecto indirecto	-0,109*** (0,035)	-0,128*** (0,018)	-0,136*** (0,044)	-0,114*** (0,016)
	Total	-0,021*** (0,007)	-0,036*** (0,005)	-0,050*** (0,017)	-0,024*** (0,003)
Aglomeración (ln)	Efecto directo	0,327*** (0,105)	0,010 (0,060)	0,186*** (0,032)	0,022 (0,034)
	Efecto indirecto	0,075 (0,055)	-0,167*** (0,048)	-0,293*** (0,053)	-0,028 (0,044)
	Efecto total	0,403*** (0,080)	-0,157*** (0,051)	-0,107*** (0,023)	-0,006 (0,009)
Distancia al subte (ln)	Efecto directo	-0,001 (0,007)	-0,005 (0,007)	0,003 (0,008)	-0,005 (0,007)
	Efecto indirecto	0,002 (0,008)	0,007 (0,010)	-0,005 (0,012)	0,007 (0,009)
	Efecto total	0,000 (0,002)	0,002 (0,003)	-0,002 (0,004)	0,001 (0,002)
Frecuencia de colectivos (ln)	Efecto directo	0,589*** (0,134)	0,752*** (0,118)	0,642*** (0,131)	0,692*** (0,114)
	Efecto indirecto	-0,729*** (0,159)	-1,040*** (0,163)	-1,011*** (0,209)	-0,878*** (0,145)
	Efecto total	-0,140*** (0,031)	-0,289*** (0,045)	-0,370*** (0,084)	-0,186*** (0,031)
Frecuencia de colectivos al cuadrado (ln)	Efecto directo	-0,046*** (0,010)	-0,057*** (0,008)	-0,048*** (0,009)	-0,052*** (0,008)
	Efecto indirecto	0,056*** (0,011)	0,079*** (0,012)	0,076*** (0,015)	0,067*** (0,010)
	Efecto total	0,011*** (0,002)	0,022*** (0,003)	0,028*** (0,006)	0,014*** (0,002)
Densidad (ln)	Efecto directo	0,019*** (0,007)	0,023*** (0,004)	0,022*** (0,007)	0,021*** (0,004)

Continúa en la página siguiente.

Cuadro 6. Modelos espaciales: Modelo 1 (continuación)					
Versión del modelo		**GNS**		**SAC**	
Método de estimación		**MC2E**	**MV**	**MC2E**	**MV**
Panel 2: Efectos estimados					
Densidad (ln)	Efecto indirecto	-0,024*** (0,009)	-0,032*** (0,006)	-0,034*** (0,012)	-0,027*** (0,006)
	Efecto total	-0,005*** (0,002)	-0,009*** (0,002)	-0,012*** (0,005)	-0,006*** (0,001)

***Estadísticamente significativo al 99%, **al 95%, *al 90%. Variable dependiente expresada en logaritmo natural.
Fuente: elaboración propia con base en datos provenientes de SIPA y AFIP.

De acuerdo a la estimación de δ reportada en el panel 1 del cuadro 6, es posible afirmar que existen *spillovers* significativos en la variable dependiente: la remuneración promedio de los habitantes de una manzana se relaciona positivamente con la remuneración promedio de los habitantes de las manzanas cercanas. Distintas hipótesis pueden racionalizar estos resultados. En primer lugar, es esperable que personas que viven en manzanas cercanas entre sí se enfrenten a similares costos de vida y que, por ende, reciban salarios similares para poder afrontar estos costos. De esta forma, si aumentan sistemáticamente los salarios en una manzana y sus residentes modifican sus hábitos de consumo de manera acorde, entonces es esperable que las manzanas aledañas también se encarezcan, lo que induciría a los trabajadores de menores ingresos a abandonar la zona (mecanismo que puede ser reforzado por cambios en el precio de los inmuebles). Por otro lado, los *spillovers* en la remuneración también pueden explicarse a partir de *spillovers* en el consumo en caso de que los comercios de la zona se encuentren en manos de sus residentes. En este caso, un aumento en la remuneración promedio de una manzana aumenta el consumo tanto en esa manzana como en las manzanas cercanas, lo que aumentaría la remuneración de aquellos trabajadores que son dueños de un negocio en la zona.

Corresponde recordar que las estimaciones de δ, λ y θ no son en sí mismas medidas de tamaños de efectos indirectos sino insumos para un cálculo recursivo que da forma a los efectos directos, indirectos y totales. La forma correcta de interpretar los resultados de estas estimaciones es a través de los valores presentados en el panel 2, en el que se reportan los efectos que surgen de incorporar simultáneamente todos los *spillovers* espaciales. Los resultados de este panel reflejan la idea de que, cuando una variable explicativa cambia en una manzana particular, la variable dependiente no solo cambia en esa manzana sino que también cambia en las manzanas cercanas, lo cual en un efecto de segunda ronda vuelve a impactar sobre la manzana analizada inicialmente.

Al controlar por los derrames geográficos, el coeficiente que acompaña la variable de *commuting* es menor al hallado por MCO, incluso negativo. Esto sugiere que, cuando se incorporan los efectos espaciales, se encuentra una predominancia del efecto asociado al hecho de que las personas de mayores ingresos deciden endógenamente residir cerca de sus trabajos de modo tal de ahorrar en costos de viaje (tanto en términos de costos monetarios como de oportunidad). Este resultado es robusto a las distintas especificaciones.

En cuanto a la variable de aglomeración, los resultados son poco robustos tanto a la especificación del modelo como al método de estimación implementado. En las estimaciones de la especificación preferida, el modelo GNS, el efecto total es positivo y estadísticamente significativo (de acuerdo con el resultado

teórico esperado) cuando se estima por MC2E y no significativo cuando se estima por MV. En cambio, en el modelo que no incluye *spillovers* en esta variable el resultado es negativo o no significativo.

Los coeficientes asociados a la frecuencia del servicio de colectivos muestran un patrón de parábola convexa, un resultado llamativo ya que es opuesto al obtenido en la estimación no espacial ponderada. En estos modelos espaciales, el salario se minimiza en niveles intermedios de este regresor, lo que sugiere que los salarios más altos se ubican en zonas con frecuencias muy bajas o muy altas en el servicio de colectivos. Este resultado es producto de efectos indirectos grandes y significativos, que operan en sentido opuesto a los efectos directos y los compensan.

Finalmente, controlando por los derrames geográficos, se observa una asociación negativa entre la densidad poblacional de una manzana y los salarios de los habitantes de la misma. Presumiblemente, una de las razones detrás de estos resultados sea que el efecto predominante en la relación es el negativo derivado de la mayor congestión espacial y contaminación. Estas externalidades negativas aumentan con la densidad; si la menor demanda de viviendas en estas zonas se traduce en menores precios inmobiliarios, entonces es posible que quienes allí residan sean trabajadores de menores ingresos, dispuestos a afrontar estos costos de congestión con el fin de reducir los costos monetarios de la vivienda.

Cuadro 7. Modelos espaciales: Modelo 2

Versión del modelo		**GNS**		**SAC**	
Método de estimación		**MC2E**	**MV**	**MC2E**	**MV**
Panel 1: Coeficientes estimados					
θ Aglomeración		-1,340*** (0,162)	-0,357 (0,231)	-	-
δ		3,139*** (0,217)	0,704*** (0,184)	2,831*** (0,315)	1,770*** (0,218)
λ		0,677 (0,449)	2,900*** (0,140)	1,730*** (0,109)	1,941*** (0,164)
N		9299	9299	9299	9299
Pseudo R2		0,072	0,277	0,070	0,180
Panel 2: Efectos estimados					
Commuting (ln)	Efecto directo	0,049*** (0,010)	0,051*** (0,008)	0,050*** 0,010	0,052*** (0,008)
	Efecto indirecto	-0,072*** (0,016)	0,122 (0,109)	-0,077*** (0,017)	-0,120*** (0,026)
	Efecto total	-0,023*** (0,006)	0,174 (0,111)	-0,027*** (0,008)	-0,068*** (0,022)

Continúa en la página siguiente.

Cuadro 7. Modelos espaciales: Modelo 2 (continuación)					
Versión del modelo		**GNS**		**SAC**	
Método de estimación		**MC2E**	**MV**	**MC2E**	**MV**
Panel 2: Efectos estimados					
Aglomeración (ln)	Efecto directo	0,328*** (0,059)	0,302*** (0,096)	-0,014 (0,052)	0,143*** (0,054)
	Efecto indirecto	0,145*** (0,038)	-0,487 (0,649)	0,021 (0,080)	-0,329** (0,157)
	Efecto total	0,473*** (0,041)	-0,185 (0,588)	0,007 (0,028)	-0,186* (0,106)
Distancia al subte (ln)	Efecto directo	-0,011 (0,008)	-0,015 (0,012)	-0,016 (0,010)	-0,011 (0,010)
	Efecto indirecto	0,017 (0,012)	-0,035 (0,041)	0,024 (0,015)	0,026 (0,024)
	Efecto total	0,005 (0,004)	-0,050 (0,050)	0,009 (0,005)	0,015 (0,014)
Frecuencia de colectivos (ln)	Efecto directo	-0,221 (0,153)	-0,246 (0,186)	-0,138 (0,158)	-0,205 (0,175)
	Efecto indirecto	0,324 (0,224)	-0,585 (0,681)	0,214 (0,246)	0,470 (0,414)
	Efecto total	0,103 (0,072)	-0,831 (0,814)	0,075 (0,089)	0,265 (0,243)
Frecuencia de colectivos al cuadrado (ln)	Efecto directo	0,014 (0,011)	0,015 (0,013)	0,008 (0,011)	0,013 (0,013)
	Efecto indirecto	-0,020 (0,016)	0,035 (0,045)	-0,013 (0,017)	-0,030 (0,029)
	Efecto total	-0,006 (0,005)	0,050 (0,055)	-0,004 (0,006)	-0,017 (0,017)
Densidad (ln)	Efecto directo	-0,018*** (0,007)	-0,021*** (0,007)	-0,015** (0,007)	-0,020*** (0,007)
	Efecto indirecto	0,026*** (0,010)	-0,049 (0,046)	0,023** (0,011)	0,046** (0,018)
	Efecto total	0,008** (0,003)	-0,069 (0,049)	0,008* (0,004)	0,026** (0,012)

Continúa en la página siguiente.

Cuadro 7. Modelos espaciales: Modelo 2 (continuación)					
Versión del modelo		**GNS**		**SAC**	
Método de estimación		**MC2E**	**MV**	**MC2E**	**MV**
Panel 2: Efectos estimados					
Dummy de distrito económico	Efecto directo	0,031* (0,016)	0,055*** (0,020)	0,028 (0,018)	0,046** (0,018)
	Efecto indirecto	-0,045* (0,024)	0,130 (0,121)	-0,044 (0,029)	-0,105** (0,047)
	Efecto total	-0,014* (0,008)	0,184 (0,129)	-0,015 (0,011)	-0,059* (0,031)

***Estadísticamente significativo al 99%, **al 95%, *al 90%. Variable dependiente expresada en logaritmo natural.
Fuente: elaboración propia con base en datos provenientes de SIPA y AFIP.

Al promediar las variables en función de las personas que trabajan en cada manzana, vuelve a encontrarse una correlación espacial positiva en la variable dependiente, lo que indica que la remuneración pagada por los empleadores de una manzana está asociada a lo que sucede en las manzanas cercanas. La explicación más sencilla es que los empleadores de una determinada manzana tienen costos muy similares a los costos de otros empleadores cercanos (por ejemplo, similares costos de alquiler y de transporte), lo que podría traducirse en similares retribuciones pagadas a sus trabajadores. Por otro lado, esta autocorrelación espacial positiva en la remuneración también podría estar reflejando cierta competencia entre las firmas en el mercado laboral: los empleadores de una determinada manzana, ante un aumento en la remuneración pagada por los empleadores vecinos, deberían aumentar sus salarios con el fin de evitar volverse relativamente menos atractivos a la vista de los trabajadores.

Por otro lado, comparando los resultados del panel 2 con los resultados del cuadro 5, se observa que el efecto total de la variable de distancia de *commuting* se vuelve negativo. Este resultado indica que el efecto que predomina es el efecto negativo de la relación; si bien los trabajos mejor pagos son más atractivos (lo que haría que los trabajadores estén dispuestos a viajar distancias más largas para llegar a ellos), también es cierto que los trabajadores de ramas de alta remuneración podrían priorizar trabajar cerca del hogar, de forma tal de reducir la desutilidad generada por el traslado. De predominar este último caso, se obtendría un efecto negativo en la relación.

Por su parte, en el caso de la variable de aglomeración, el efecto total estimado es menor al obtenido por MCO, incluso negativo en algunos casos. En la especificación preferida (modelo GNS), el efecto total es estadísticamente significativo únicamente al utilizar MC2E, con signo positivo. La remuneración de una manzana no solo está influida (positivamente) por la propia aglomeración, sino también por la aglomeración de las manzanas aledañas. Este es un resultado razonable si se considera que las economías de escala posiblemente traspasen los límites de las manzanas (por ejemplo, a través de derrames tecnológicos entre manzanas). Por otro lado, de no considerarse este tipo de derrame geográfico (modelo SAC), el efecto total de la variable de aglomeración se vuelve negativo y estadísticamente significativo al estimar por medio de MV.

En cuanto a la densidad poblacional, se encuentra un efecto positivo y estadísticamente significativo sobre la remuneración pagada por los empleadores. Este es un resultado robusto a las distintas

especificaciones, aunque económicamente poco relevante. Si bien el efecto directo es negativo, posiblemente debido a los mayores costos derivados de la mayor congestión, el efecto indirecto es positivo; esto último podría explicarse por el hecho de que tener manzanas de mayor densidad poblacional cerca puede significar una mayor demanda y, por ende, ciertas ganancias en términos de economías de escala.

Por último, en la *dummy* de distrito económico se encuentran efectos derrame negativos que compensan en magnitud los positivos, resultado posiblemente determinado por el hecho de que los distritos económicos están situados en torno a zonas generalmente poco favorecidas de la ciudad en términos de desarrollo económico y social local, de modo que ubicarse cerca de un distrito es en general un indicador de valores bajos en variables inobservables del problema.

6. Comentarios finales

Este trabajo provee evidencia novedosa sobre la distribución geográfica del empleo formal en CABA. Los resultados obtenidos capturan distintas características espaciales de la economía de la ciudad, en particular, una distribución salarial que ubica las remuneraciones más elevadas en el centro y en el corredor norte de la ciudad, aunque con presencia de clusters de ingresos altos en otras zonas. Estos atributos son consistentes con las predicciones teóricas para una ciudad monocéntrica en la que existen diversos medios de transporte y también considerable variabilidad en la distribución del ingreso.

La estimación de ecuaciones de salarios tradicionales indica correlación positiva entre remuneraciones, aglomeración económica y distancia de *commuting*, lo que sugiere que los efectos positivos de la concentración de la actividad productiva superan sus efectos adversos y que los trabajadores están dispuestos a recorrer mayores distancias en sus traslados laborales a cambio de un salario mayor. Sin embargo, algunos de estos resultados desaparecen o se invierten cuando se incorporan efectos espaciales. Es preciso ahondar en estos hallazgos en futuros trabajos, en la medida en que la baja robustez de estas últimas estimaciones a la especificación del modelo y el método de estimación elegidos puede explicar al menos parte de estas divergencias.

También se encuentra que la distribución sectorial del empleo exhibe patrones espaciales. Esto, sumado a la existencia de distritos económicos en los que se otorga beneficios fiscales a determinadas actividades productivas (con su correspondiente efecto en salarios) sugiere también la posibilidad de aportes futuros en la investigación, particularmente a través de la aplicación de métodos de inferencia causal. Una consideración similar aplica a la oferta de transporte público, variables donde nuevamente los resultados varían de acuerdo al modelo que se elija, y donde el estudio de efectos causales puede resultar particularmente interesante.

CABA presenta una división norte-sur que puede rastrearse a las epidemias de fiebre amarilla que tuvieron lugar en el siglo XIX. Que esa distribución histórica se asemeje a la actual es un resultado que puede ser interpretado de diversas maneras, entre las que destaca la *path-dependence*, en la medida en que la distribución de la población en el espacio geográfico factiblemente supone importantes costos hundidos. Aquí también es posible contemplar la aplicación de diversas herramientas econométricas que permitan inferir efectos temporales de interés.

Referencias bibliográficas

Arnott, R. J. y Stiglitz, J. E. (1979). Aggregate land rents, expenditure on public goods, and optimal city size. *The Quarterly Journal of Economics, 93*(4), pp. 471-500.

Alonso, W. (1964). Location and land use, publications of the Joint Center for Urban Studies of the Massachusetts Institute of Technology and Harvard University.

Becker, G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal, 75*(299), pp. 493-517.

Carmona, R. (2017). The economic districts in the autonomous city of Buenos Aires as a new form of urban intervention. *Direito da Cidade, 9*(4), pp. 1862-1884.

Chauvin, J. P., Glaeser, E., Ma, Y., y Tobio, K. (2017). What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics, 98*, pp. 17-49.

Ferreyra, M. M., y Roberts, M. (eds.) (2018). Raising the bar for productive cities in Latin America and the Caribbean. World Bank Publications.

Glaeser, E. L., Kahn, M. E. y Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics, 63*(1), pp. 1-24.

González, F. y Anapolsky, S. (2022) Identificando la desigualdad en los patrones de movilidad en transporte público. Inter-American Development Bank.

Krugman, P. (1992). A Dynamic Spatial Model. *NBER Working Papers 4219, National Bureau of Economic Research, Inc*.

Krugman, P. (1991a). Geography and Trade. MIT Press.

Krugman, P. (1991b). Increasing returns and economic geography. *Journal of Political Economy*, *99*(3), pp. 483-499.

LeRoy, S. F., y Sonstelie, J. (1983). Paradise lost and regained: Transportation innovation, income, and residential location. *Journal of Urban Economics, 13*(1), pp. 67-89.

Mendiola, L., y González, P. (2021). Urban Development and Sustainable Mobility: A Spatial Analysis in the Buenos Aires Metropolitan Area. *Land*, *10*(2), 157.

Mills, E. S. (1967). An aggregative model of resource allocation in a metropolitan area. *The American Economic Review, 57*(2), pp. 197-210.

Mills, E. y Hamilton, B. (1995). Urban Economics. Prentice Hall.

Muth, R. (1968). Cities and Housing. Univ. of Chicago Press.

Quintero, L. y Roberts, M. (2018). Explaining spatial variations in productivity: evidence from Latin America and the Caribbean. World Bank Policy Research Working Paper, (8560).

Roberts, M., Blankespoor, B., Deuskar, C., y Stewart, B. (2017). Urbanization and development: is Latin America and the Caribbean different from the rest of the world?. World Bank Policy Research Working Paper, (8019).

Scobie, J. R. (1986). Buenos Aires del centro a los barrios 1870-1910. Ediciones Solar.

Vargas, J., Brassiolo, P., Sanguinetti, P., Daude, C., Goytia, C., Álvarez, F., … Fajardo, G. (2017). RED 2017. Urban growth and access to opportunities: a challenge for Latin America. Bogotá: CAF. Recuperado de http://scioteca.caf.com/handle/123456789/1091.