



**Transporte público y heterogeneidad de la informalidad laboral: evidencia para
Metrocable Línea J de Medellín**

Walter José Marín Barrios

Tesis de maestría presentada para optar al título de Magíster en Economía

Tutor

Héctor Mauricio Posada Duque, Doctor (PhD) en Economía

Universidad de Antioquia
Facultad de Ciencias Económicas
Maestría en Economía
Medellín, Antioquia, Colombia
2022

Cita	(Marín Barrios, 2022)
Referencia	Marín Barrios, W. (2018). <i>Transporte público y heterogeneidad de la informalidad laboral: evidencia para Metrocable Línea J de Medellín</i> [Tesis de maestría].
Estilo APA 7 (2020)	Universidad de Antioquia, Medellín, Colombia.



Maestría en Economía, Cohorte XVII.

Grupo de Investigación Microeconomía Aplicada.

Seleccione centro de investigación UdeA (A-Z).



Centro de Documentación Economía

Repositorio Institucional: <http://bibliotecadigital.udea.edu.co>

Universidad de Antioquia - www.udea.edu.co

Rector: John Jairo Arboleda Céspedes.

Decano/Director: Sergio Iván Restrepo Ochoa.

Jefe departamento: Wilman Arturo Gómez Muñoz.

El contenido de esta obra corresponde al derecho de expresión de los autores y no compromete el pensamiento institucional de la Universidad de Antioquia ni desata su responsabilidad frente a terceros. Los autores asumen la responsabilidad por los derechos de autor y conexos.

Nota del autor

Este trabajo fue realizado con apoyo financiero del Programa de Investigación Colombia Científica-Alianza EFI, con código 60185 y número de contrato FP44842-220-2018, financiado por El Banco Mundial a través de la convocatoria Ecosistemas Científicos, gestionada por el Ministerio de Ciencia, Tecnología e Innovación de Colombia. Los errores, opiniones y omisiones son nuestros y no comprometen a las instituciones.

Agradezco a Colombia Científica-Alianza Programa de Investigación Economía Formal e Inclusiva, por la oportunidad de hacerme parte de su grupo de trabajo como joven investigador y a mi tutor Héctor Mauricio Posada Duque por sus valiosos aportes y comentarios a la elaboración de esta tesis de maestría en economía de la Universidad de Antioquía.

Resumen

La literatura ha evidenciado que el transporte público es una herramienta esencial para reducir el desajuste espacial y con ello, la informalidad laboral, porque puede mejorar la asignación de trabajadores de la economía informal cerca del hogar, a la formal en el centro de la ciudad. Sin embargo, es posible que su impacto puede darse de manera diferencial dada la heterogeneidad ocupacional. Por ello, este estudio investiga el efecto causal del Metrocable Línea J (MCJ) sobre la informalidad laboral en Medellín, con especial interés en verificar si existe un efecto diferencial por tipo de ocupación (cuenta propia y asalariados) para captar quienes aprovechan mejor la infraestructura de tránsito. Usando información de la Encuesta de Calidad de Vida de Medellín y una estrategia de identificación que combina especificación de triple diferencia con Control Sintético, se estima para toda la zona de tratamiento que la informalidad se reduce 5,9 p.p aproximadamente en asalariados, mientras que en los cuenta propia aumenta 5,8 p.p. El diferencial entre ambos efectos es amplio y significativo, lo que sugiere que MCJ impacta de manera diferente según la posición ocupacional.

Palabras clave: informalidad laboral, heterogeneidad ocupacional, Metrocable, desajuste espacial, inferencia causal, Control Sintético, triple diferencia, Medellín.

Clasificación JEL: C31, R12, R41, R42, J46

Abstract

The literature has shown that public transportation is an essential tool to reduce spatial mismatch and thus labor informality, because it can improve the allocation of workers from the informal economy close to home to the formal economy in the city center. However, it is possible that its impact may be differential given occupational heterogeneity. Therefore, this study investigates the causal effect of the Metrocable Line J (MCJ) on labor informality in Medellín, with special interest in verifying whether there is a differential effect by type of occupation (self-employed and salaried) to capture those who take better advantage of the transit infrastructure. Using information from the Encuesta de Calidad de Vida (Survey of Living Conditions, ECV for its acronym in Spanish) and an identification strategy that combines triple difference specification with Synthetic Control, it is estimated for the entire treatment area that informality is reduced by approximately 5.9 p.p. for salaried workers, while it increases by 5.8 p.p. The differential between both effects is large and significant, suggesting that MCJ impacts differently according to occupational position.

Keywords: labor informality, occupational heterogeneity, Metrocable, spatial mismatch, causal inference, Synthetic Control, triple difference, Medellín.

Clasificación JEL: C31, R12, R41, R42, J46

Introducción

Es usual que la informalidad denote un aspecto negativo relacionado con el incumplimiento total o parcial de la normatividad formal del mercado laboral, definida comúnmente por la falta de acceso a seguridad social (Bonet et al., 2016). Los países en desarrollo de África y América Latina son más vulnerables a estos casos de informalidad. Según la Organización Internacional del Trabajo (OIT), en países como Etiopía, Madagascar, Rwanda, Uganda y Zambia, más del 70% de la población ocupada es informal, mientras que las economías latinoamericanas presentan en promedio entre el 20% y el 50% de informalidad laboral, cifra que alcanzó un 46% en Colombia (2019), según (DANE, 2019).

Estudios de empleo en zonas de pobreza sugieren un mercado laboral característico de redes sociales informales (Damm, 2014), bajos niveles de educación y altas tasas de desempleo, inactividad e informalidad (Amin, 2016; Mitra, 1990). Estas características desfavorables se han explicado por condiciones estructurales de demanda u oferta laboral o por la falta de servicios urbanos como escuelas, centros de salud y transporte público (Cobbinah et al., 2015) pero otra línea de literatura se ha centrado en las condiciones espaciales de los individuos. La relación entre la desconexión espacial y los resultados laborales se denomina Hipótesis de Desajuste Espacial (SMH) (Bradbury & Chalmers, 2003; Da Piedade et al., 2015; Gobillon et al., 2007; Kain, 1968; Offner, 1972). En esta se explica la asociación geográfica entre centros de actividad económica o de empleo y lugar de residencia de los individuos. Heilmann (2014) sugiere que las oportunidades de empleo dependen en gran medida de la ubicación de los individuos, si residen en un lugar segregado espacialmente, se reduce la movilidad laboral porque no pueden acercarse a sus trabajos tanto como quisieran, lo cual tiene implicaciones negativas en términos de reducción de los ingresos laborales, una baja tasa de empleo o puestos de trabajo de baja calidad.

Esto es más frecuente en ciudades de economías en desarrollo y emergentes que presentan altos niveles de segregación socioeconómica. Los empleos formales se concentran en zonas centrales con buena accesibilidad, mientras las zonas periféricas presentan accesibilidad limitada y alta incidencia de empleo de baja calidad porque carecen de información sobre oportunidades laborales. Esta división se ve reforzada por la prestación subóptima y sesgada del transporte público urbano con grandes distancias por recorrer y mayores tiempos de desplazamiento para esa misma distancia (Moreno-Monroy, 2016). Sumado a eso, el crecimiento de barrios ubicados en zonas

periféricas de las ciudades y la escasa planificación de estos conlleva a un desarrollo vial deficiente, con calles estrechas, sin pavimentación e incluso sin espacio para la infraestructura vial¹. De hecho, según Organización de Naciones Unidas (2018), en 2018, el 53% de los residentes urbanos en el mundo tenía acceso conveniente al transporte público, mientras que esta cifra alcanza el 50% en Latino América, 24 puntos porcentuales menos que en economías desarrolladas de Europa y Norte América. Esto ha causado que, en regiones con baja accesibilidad, se incentiven los modos de transporte informales para cerrar las brechas de accesibilidad. Datos para ciudades colombianas muestran que, en Medellín, según Encuesta Origen Destino 2017, viajar desde la zona de bajos ingresos occidental hasta el centro de la ciudad tarda al menos 50 minutos, mientras que, viajar de la zona de mayores ingresos suroriental tarda 36 minutos en promedio. Caso similar sucede en Bogotá, trabajadores de bajos ingresos en la zona pobre del sur de la ciudad tardan en promedio 60 minutos en recorrer 2 a 3 kilómetros, mientras que, viajar la misma distancia en la zona norte de mayores ingresos toma solo 35 minutos (Bocarejo & Oviedo, 2012).

De esta manera, el transporte público es una herramienta esencial para la inclusión social porque puede reducir el desajuste espacial entre la periferia y las economías de aglomeración a partir de sus mecanismos para mejorar resultados laborales (Gobillon et al., 2007). Estos mecanismos pueden contribuir a la reducción del empleo informal, pues una mejora en la infraestructura de transporte permite un mejor acceso a información sobre empleos formales, haciendo que los individuos valoren menos los empleos informales cercanos (Moreno-Monroy, 2016). La literatura ha abarcado el impacto de la infraestructura de transporte en diferentes variables como el valor de la propiedad (Martínez & Viegas, 2009), dispersión de la ciudad (Iaçu et al., 2011), pobreza (Glaeser et al., 2008), crimen (Khanna et al., 2020), calidad de la vivienda (Posada & García-Suaza, 2021) y en mercado laboral, en participación y subempleo (Jaramillo & Rengifo, 2018; Seki & Yamada, 2020). La informalidad suele explicarse por condiciones de oferta o estructura de la demanda y apenas se ha examinado el papel de la accesibilidad al transporte, con aproximaciones por parte de Moreno & Roman, (2019) y Zárate, (2020). En Medellín, la literatura se ha centrado en impactos de Metroplús sobre resultados del mercado laboral (Jaramillo & Rengifo, 2018) y la calidad de vida (Gómez & Semeshenko, 2018), impactos de líneas del

¹ Este es el caso de Jalousie, Haití (The Guardian, 2018).

Metrocable sobre el crimen (Bea, 2016) y calidad de la vivienda (Posada & García-Suaza, 2021), pero aún no se ha explorado su efecto heterogéneo sobre la informalidad laboral.

Respecto a esta heterogeneidad, Amin (2016) clasifica los trabajadores informales en comerciantes ambulantes, conductores de transporte informal, obreros de construcción, recolectores de basura y criadas. Por otro, Banerjee & Goswami (2019)² comentan que las categorías ocupacionales más comunes, en especial en barrios marginales, son trabajadores por cuenta propia, trabajo familiar, empleados asalariados en empresas no constituidas legalmente y empleadores de empresas privadas no constituidas en sociedad. Este último refleja en parte la motivación de este trabajo, ya que, según este autor, la literatura en el contexto de barrios marginales rara vez reconoce la heterogeneidad ocupacional del trabajo informal y que existe una escasez de estudios que especifiquen como las políticas y desarrollos de infraestructura pública pueden enfocarse y diseñarse de acuerdo con los determinados tipos de informalidad laboral.

Banerjee & Goswami (2019) contribuyen a esta literatura y comentan en favor de Meagher (2016) acerca de hacer esfuerzos por considerar qué tipos de informalidad suelen ser más favorables para compensar la pobreza de los habitantes de tugurios, en países en desarrollo. Si bien es cierto que la informalidad como fuente de compensación de la pobreza es debatible, porque usualmente es visto como último recurso para los más desfavorecidos, existe otra literatura que define trabajadores cuenta propia como impulsores del crecimiento económico (Schumpeter, 1943), la base del desarrollo (De Soto, 1990) o cómo lo establece (Akerlof, 1970) los comerciantes pueden convertirse en los principales empresarios. De esta manera, las políticas públicas y las intervenciones de infraestructura deben generar las condiciones para que empleos informales en distintos tipos puedan mejorar sus medios de vida o les permita transitar hacia empleos de mejor calidad. De ahí la importancia de comprender cuáles de estas ocupaciones, cuenta propia o asalariados, ofrecen mejores recursos para una mejor calidad de vida y quienes aprovechan mejor la dotación de infraestructura como el transporte público. Este estudio considera dos tipos de

² Barrios marginales de cuatro distritos de Bengala Occidental, India.

ocupaciones, asalariados³ y cuenta propia⁴. En Colombia, según datos de la Gran Encuesta Integrada de Hogares (GEIH), en 2019 los cuenta propia representaron el 40% del total de ocupados y el 65 % de los informales por seguridad social, mientras que los asalariados representaron un 54% del total y el 33 % de informales.

Lo anterior brinda una oportunidad de investigación en la literatura sobre el transporte público e informalidad laboral, haciendo énfasis en su heterogeneidad. Por ello, se investiga el efecto causal del sistema de transporte Metrocable Línea J (MCJ) sobre la informalidad laboral en Medellín, en particular, verificar si existe efecto heterogéneo por tipo de empleo cuenta propia o asalariados para captar quienes aprovechan mejor el transporte público. Para este estudio se usa información de la Encuesta de Calidad de Vida de Medellín para los años 2005, 2007, 2008, 2010 y 2013 con una estrategia de identificación que combina Control Sintético (CS) para elegir un contrafactual adecuado con comunas no tratadas para equilibrar covariables pretratamiento y posteriormente una estimación de triple diferencia (Olden & Moen, 2020) para captar la heterogeneidad ocupacional. También se apoyan las estimaciones con un estudio de eventos y regresiones placebo para ofrecer evidencia sobre los supuestos de tendencias paralelas y comportamiento anticipado (Sun & Abraham, 2020)⁵.

Respecto al comportamiento anticipado, este está relacionado con la migración de personas de altos ingresos hacia la nueva estación antes del tratamiento, lo cual implicaría que un efecto de reducción de la informalidad pueda deberse a la llegada de estos residentes y no a un beneficio de la infraestructura. No obstante, la evidencia de estos sucesos es debatible, por un lado, Warnes (2020) muestra que la infraestructura de tránsito podría aumentar la segregación espacial entre trabajadores cualificados y poco cualificados dentro de la ciudad, por otro, Padeiro et al., (2019) argumentan que grupos de bajos ingresos podrían continuar dominando los vecindarios con nuevo servicio de transporte público, ya que las personas de altos ingresos podrían evitar mudarse por la

³ **Asalariados:** trabajadores de una unidad institucional que reciben en compensación una remuneración fundamentada en el tiempo de trabajo o en algún otro indicador de la cantidad de trabajo realizado. Los conforman los trabajadores privados, del gobierno y domésticos (DANE).

⁴ **Cuenta propia:** son trabajadores que dirigen su propia empresa o ejercen su profesión sin contratar a ningún trabajador a su cargo, hacen parte del grupo de trabajadores independientes junto con patronos y trabajadores familiares sin remuneración. (DANE).

⁵ Para el estudio de eventos se agregará el año 2006 y 2018 para reforzar la verificación de los supuestos y mostrar si el sistema MCJ tiene efectos de largo plazo.

congestión, la escasez de apartamentos grandes y cómodos, escasas posibilidades de estacionamiento y crimen. Incluso, comentan que aún con un aumento de residentes más acomodados, los grupos vulnerables podrían permanecer en la zona y aprovechar la mejor accesibilidad a trabajos y otros recursos. Por lo tanto, los beneficios resultantes podrían no ser el resultado de desplazamientos ex ante, sino de mejoras generales en las condiciones económicas como resultado de un mecanismo de transmisión de política.

Los resultados muestran que MCJ tiene efectos diferenciales en la informalidad según tipos de ocupación. MCJ logra disminuir la informalidad en los asalariados y la aumenta para los cuenta propia. En particular, para toda la zona de tratamiento, se observa que la informalidad en asalariados se redujo significativamente 5,9 puntos porcentuales (p.p), mientras que en zonas de influencia la reducción va desde 6,4 p.p (500 metros) hasta 6,2 p.p (1500 metros). En cambio, la informalidad en cuenta propia aumenta 5,8 p.p en toda la comuna tratada, mientras que en zonas de influencia el aumenta es cada vez menor, pasando de 8,5 p.p (500 metros) hasta 7,8 p.p (1500 metros). Además, las diferencias entre ambos efectos resultan amplias y significativas en todas las estimaciones. Este trabajo contribuye a la extensa literatura sobre impactos del transporte público y a los pocos estudios sobre efectos en la informalidad con un enfoque metodológico diferente. Aporta nueva evidencia de que la infraestructura de transporte tiene efectos diferenciados por posición ocupacional. Por último, en términos de política pública, contribuye a entender cómo se reflejan los beneficios de las intervenciones de transporte en el mercado laboral en términos de calidad del empleo, dependiendo de la ocupación del trabajador y su forma de valorar la accesibilidad. Así, se busca responder los siguientes interrogantes: ¿puede la provisión de transporte ser una forma de reducir la informalidad laboral? y ¿qué tipo de ocupación, cuenta propia o asalariados, tiende a formalizarse más dadas la presencia de infraestructura de transporte?

El documento está organizado de la siguiente manera: primero la introducción, luego la revisión de literatura sobre heterogeneidad laboral de los trabajadores informales y el impacto causal de la infraestructura de transporte. Los datos y la metodología se describen posterior a este y, por último, muestro resultados, discusión y conclusiones.

1 Revisión de literatura

Esta revisión dará cuenta de lo que se ha hecho en materia de mercado laboral en asentamientos, en particular, los estudios de informalidad laboral dentro de estas zonas, como se relaciona con la hipótesis de desajuste espacial, su heterogeneidad ocupacional y las implicaciones del transporte público. Esto para ubicar al lector en lo que se sabe y no se sabe entorno a estos fenómenos, ilustrar el vacío que se pretende llenar aquí e incentivarle sobre nuevos rumbos de investigación.

Antes de guiar esta sección por los estudios relacionados al objeto de investigación, es necesario abordar los enfoques que se le ha dado históricamente a existencia y persistencia de la informalidad laboral, donde se destaca el enfoque estructuralista e institucionalista. Por un lado, Bourguignon (1979) argumenta desde el enfoque estructuralista que la informalidad laboral es un sector residual que se debe a fallos del mercado laboral, por su escaso desarrollo moderno, que limita la productividad y no alcanza a absorber toda la fuerza laboral existente, por tanto, los individuos optan por trabajar en el sector informal antes de quedarse desempleados. Entre las variables más comunes para caracterizar este primer enfoque se tienen: el nivel de desempleo, el nivel de capital físico, el grado de desarrollo tecnológico, entre otros (Uribe et al., 2006). Por otro, el pensamiento institucionalista asocia este fenómeno a factores institucionales que provocan la evasión de las normas legales, ya que las cargas fiscales e ineficiencias gubernamentales (como los impuestos, costos laborales, regulaciones ambientales, pago diferencial de servicios públicos, entre otros) incentivan el incumplimiento de las reglas institucionales y, por tanto, la informalidad. Las variables más utilizadas para medir este enfoque son: medidas de recaudo fiscal, medidas de restricciones del mercado laboral y el grado de calidad de las instituciones gubernamentales (Loayza, 1996).

Entre los estudios que asocian informalidad laboral con asentamientos informales se encuentran Alvarez-Rivadulla (2009), Elgin & Oyvat (2013) y Bonet et al., (2016), cuyos objetivos son determinar el efecto de vivir en un asentamiento informal sobre probabilidad de obtener un trabajo informal. Han utilizado encuestas de hogares como Encuesta Nacional de Hogares Ampliada 2006 (ENHA) y la Gran Encuesta Integrada de Hogares (GEIH) que ofrecen la posibilidad de identificar a personas que viven en tierras ocupadas ilegalmente o sin tierras (asentamientos informales) y que no están afiliadas a la seguridad social (trabajo de informalidad).

Autores encuentran que la informalidad laboral no es única de las zonas informales y que vivir en asentamientos informales aumenta posibilidades de que un joven tenga un empleo informal. Empero, otros estudios comentan importancia de la relación causal entre ambos tipos de informalidad. Para el caso, Abramo (2003 y 2009) reconoce que el hecho de ser trabajador informal aumenta posibilidad que estas personas vivan en un asentamiento irregular. Esta evidencia sugiere tener cuidado al estudiar los efectos de los asentamientos sobre la informalidad laboral o viceversa, pues debe considerarse la relación bidireccional entre ambos tipos de informalidad. Por su parte, Banerjee & Goswami (2019), sobre la base de encuestas a 240 trabajadores de barrios informales de Bengala Occidental (País), concluyen que empleos como vendedores y conductores entre los trabajadores cuenta propia son más favorables para compensar la pobreza, mientras que entre los asalariados son más beneficiosos los que trabajan en sectores eléctrico y electrónico.

Esta literatura va de la mano con la hipótesis de desajuste espacial (Spatial Mismatch). Existe una amplia línea de investigación dedicada al estudio de la correspondencia espacial entre los centros de empleo y la residencia de los trabajadores. Uno de los primeros trabajos en abordar esta literatura fue Kain (1968) quien encontró una relación negativa entre nivel de segregación del mercado de vivienda y niveles de empleo para afroamericanos en Detroit (Estados Unidos). Además, evidenció que desplazamiento de la demanda laboral del centro de las ciudades a suburbios (fenómeno de suburbanización del empleo), incentivó a trabajadores a cambiar su residencia para estar cerca de lugares de trabajo. No obstante, los afroamericanos siguieron habitando los barrios centrales y pobres, lo que generó una especie de desconexión espacial con los lugares de empleo y finalmente resultados laborales desfavorables como desempleo, alta duración de este y baja remuneración salarial. Siguiendo el planteamiento original de la hipótesis de desajuste espacial, otros trabajos analizan resultados laborales como el desempleo, participación laboral, las horas trabajadas y la segmentación del mercado laboral a través de la segregación en el mercado de vivienda (Brueckner & Zenou, 2003; Gonzales, 2000), las características sociales del vecindario y proximidad del trabajo a economías de aglomeración (Weinberg et al., 2004) y la organización espacial de la ciudad (Bradbury & Chalmers, 2003; Dujardin et al., 2008). La evidencia de estos estudios apunta a que segregación espacial excluye personas de oportunidades laborales, redes de información acerca de ellas y aumenta el costo de participar por la lejanía de sus hogares a sitios de trabajo (Gobillon et al., 2007). Lo anterior sugiere además que los resultados

laborales no solo dependen de las características individuales del trabajador como la educación, edad, experiencia, entre otras, sino también de la ubicación geográfica de su lugar de residencia.

En este sentido, la mejora en accesibilidad juega un papel importante en las políticas para reducir las desigualdades espaciales y sociales en poblaciones vulnerables, sobre todo en países en desarrollo. Las estrategias van desde mejorar la conectividad del transporte hasta reducir los costos mediante subsidios al transporte. Para estimar el impacto causal de la expansión del transporte o los subsidios a este, en resultados laborales, la evidencia empírica recurre usualmente a diseños cuasiexperimentales como la regresión discontinua (Hernández, 2019), métodos de diferencias en diferencias (DD) (Holzer et al., 2003; Moreno & Roman, 2019; Rodríguez et al., 2016; Rotger & Nielsen, 2015), diseños experimentales como el método de Intención de Tratar (ITT) (Franklin, 2018; Phillips, 2014) y no experimentales como el método de variables instrumentales (VI) (Bastiaanssen et al., 2021; Rodríguez et al., 2016).

Por el lado de las economías desarrolladas y la expansión de los sistemas de transporte, Holzer et al., (2003) examinan, con un enfoque de diferencias en diferencias, el impacto de ampliar el sistema ferroviario Bay Area Rapid Transit (BART) de San Francisco, EE. UU, sobre contratación de minorías alrededor de dos nuevas estaciones. Muestran que accesibilidad sólo importa ligeramente para trabajadores latinos, pero no para afroamericanos. Por su parte, Rotger & Nielsen (2015) usa el mismo enfoque metodológico con panel individual para evaluar el efecto de la accesibilidad a puestos de trabajo urbanos en ingresos individuales a través del sistema de transporte público Metro en Copenhague, Dinamarca. Evidencian un impacto positivo entre accesibilidad e ingreso. Estudios más recientes como el de Bastiaanssen et al., (2021) examina en Gran Bretaña la accesibilidad al empleo en el transporte público sobre las probabilidades de empleo. Combinan la estrategia VI, para controlar la endogeneidad de la accesibilidad, con un modelo probit binomial usando como instrumento la densidad poblacional por hectárea. Encuentran que una mejor accesibilidad al empleo en el transporte público aumenta las probabilidades de empleo individual, en particular en áreas con tasas más bajas de propiedad de automóviles y en vecindarios de bajos ingresos.

En este mismo contexto, pero analizando subsidios al transporte, algunos trabajos analizan si costos de transporte limitan la búsqueda de empleo en mercados laborales urbanos de bajos salarios en Washington, DC (Phillips, 2014) y Etiopía (Franklin, 2018), a través de diseños experimentales. Phillips (2014) proporciona subsidio de transporte a clientes seleccionados al azar

de una agencia de empleo sin ánimo de lucro y estima un modelo ITT con efectos fijos de días desde una contratación. Sus resultados evidencian que individuos previamente desempleados que recibieron un subsidio de transporte público aumentaron su intensidad de búsqueda de empleo en comparación con individuos del grupo de control. Franklin (2018) usa un enfoque ITT con VI para estimar el impacto causal de los subsidios ofrecidos a jóvenes desempleados sobre su intensidad de búsqueda, incluyendo el rezago de la variable dependiente para controlar las diferencias en los resultados de línea base que instrumenta con la asignación aleatoria al grupo de tratamiento. Sus resultados muestran que tratados aumentan la intensidad de búsqueda de empleo, lo que implica más probabilidades de encontrar buenos trabajos permanentes e inducen una reducción a corto plazo del trabajo temporal.

Sobre la estrategia de subsidios al transporte en el contexto de países en desarrollo, hay algunos estudios para Bogotá, Colombia. Rodríguez et al., (2016) aprovechan la información del SISBÉN Bogotá 2015 y un plan de subvenciones al transporte público en favor de individuos de bajos ingresos para evaluar vínculo causal entre subsidio y resultados del mercado laboral como empleo (participación laboral, formalidad, desempleo y horas trabajadas) e ingresos por hora (hogares y trabajadores informales), mediante una estrategia de identificación DD con efectos fijos a nivel manzana. Sus hallazgos solo evidencian impactos significativos y positivos sobre los ingresos de los informales. Mientras tanto, Hernández (2019) busca evaluar el efecto de una política de subsidios al uso del Sistema Integrado de Transporte Público (SITP) sobre la probabilidad de participar laboralmente. Para esto, emplea datos de la Encuesta de Movilidad para el año 2015 y la Encuesta Multipropósito de 2017 para implementar un diseño de regresión discontinua borrosa usando el puntaje del SISBEN como variable de asignación y un enfoque de variable instrumental local para controlar distintas fuentes de endogeneidad (autoselección y doble causalidad). Los resultados indican que los subsidios al transporte público no incentivan a las personas en hogares que pertenecen a los estratos 1 y 2 a participar en la fuerza laboral, tampoco incrementa su tasa de ocupación o reducen su probabilidad de estar desempleado.

Por otra parte, las aproximaciones hacia el impacto del transporte público sobre la informalidad han sido estudiadas por Boisjoly et al., (2017), Moreno-Monroy & Posada (2018) y Moreno & Roman (2019). El primero Explora la relación entre informalidad laboral y la accesibilidad al empleo mediante el transporte público en la Región Metropolitana de São Paulo (RMSP), Brasil. También utilizan como medida de accesibilidad, el número de puestos de trabajo

a los que se puede llegar desde un área espacial ponderada utilizando el tránsito. Estima un modelo logístico de efectos mixtos para trabajadores que ganan más del salario mínimo comparado con aquellos que ganan menos del mismo, con el fin de captar la heterogeneidad del sector informal. Encuentran que mayor accesibilidad al transporte reduce probabilidad de formar parte del sector informal, sobretudo en segmento que gana menos del salario mínimo. Sin embargo, por limitaciones de datos, no controlan la endogeneidad presente. El segundo, desde un enfoque teórico, realizan un modelo basado en agentes para investigar el impacto de subsidios al transporte en tasas de informalidad urbana. En este trabajo, autores suponen que trabajadores formales viajan todos los días para trabajar al Distrito Central de Negocios (CBD), mientras que trabajadores informales tienen la opción de trabajar en su lugar de residencia. El estudio encuentra que un subsidio al transporte dirigido exclusivamente a trabajadores informales puede no tener impacto en tasas de informalidad, mientras que un subsidio dirigido a trabajadores formales puede conducir a una reducción de las tasas de informalidad. Por último, Moreno & Roman (2019) usa un enfoque combinado de DD y VI teniendo como instrumento un plan histórico de transporte. Evidencia que la expansión del transporte reduce tasas de informalidad en grupos que se favorecieron de mejoras en transporte público entre 2000 y 2010.

Estos estudios son de particular importancia para las pretensiones de este trabajo, ya que son primera aproximaciones en calcular efectos de las expansiones del transporte público sobre informalidad laboral. Moreno & Roman (2019) comentan que información detallada de opciones individuales de trabajo permitiría una interpretación más significativa del impacto estimado. Atendiendo a esta sugerencia, este trabajo evaluará el impacto del transporte público teniendo en cuenta la heterogeneidad laboral con diferentes tipos de ocupación de los trabajadores (cuenta propia y asalariados).

De acuerdo con la revisión, se puede concluir tres aspectos: primero, existe un consenso en la literatura que apunta a que la infraestructura de transporte logra reducir el desajuste espacial para los trabajadores de zonas periféricas de las ciudades. Segundo, los enfoques metodológicos usados generan grupos de control ad-hoc basándose usualmente en la similaridad entre zonas tratadas y zonas con probables intervenciones a futuro. Tercero, los trabajos tanto en economías desarrolladas como en desarrollo se han centrado en desempleo, participación laboral, horas trabajadas e ingresos con poca evidencia sobre la informalidad y ninguna sobre la heterogeneidad ocupacional de los

trabajadores. Este último aspecto muestra un importante vacío tanto metodológico como de evidencia más detallada en la literatura, el cual se pretende llenar con este trabajo.

2 Metodología y datos

El sistema Metrocable ya existía en otras ciudades del mundo como en Nueva York, con el teleférico hasta la isla de Roosevelt desde 1976 o en Singapur, con el teleférico para cruzar el puerto de Keppel desde 1974. Sin embargo, en 2004 Medellín construyó el primer Metrocable que se integró al sistema de transporte de una ciudad. La línea K se construyó en el noreste de la ciudad, conectando Santo Domingo, uno de los barrios más pobres y violentos, con la línea A del metro, que conecta el norte y el sur. En la actualidad, los teleféricos son una tendencia urbana novedosa, ya que se encuentran en muchas ciudades de América Latina, como Caracas, Río y La Paz, aprovechando su capacidad para conectar espacialmente zonas de extrema segregación y generando mayores oportunidades socioeconómicas gracias a la mejora de la movilidad.

Este estudio se centra en el Metrocable Línea J, que sirve directamente a la comuna de San Javier con dos estaciones (San Javier y Juan XXIII), a la comuna de Robledo con una estación (Vallejuelos) y al corregimiento de San Cristóbal con una estación (La Aurora). Sin embargo, se excluyen las estaciones La Aurora y San Javier, la primera por ser de la zona rural y la segunda porque la dinámica alrededor por actividad turística en barrio Las Independencias y el acceso al Metro puede generar factores de confusión que sobre (o sub) estimen los efectos. Este proyecto fue inaugurado en marzo de 2008 y cuenta con una extensión de 2,7 kilómetros y 119 cabinas para beneficiar directamente alrededor de 120 mil personas de los barrios San Javier, Juan XXIII, La Divisa, La Quiebra, Olaya entre otros.

En este contexto, se explorará el impacto del Metrocable línea J (MCJ) en Medellín sobre la informalidad laboral definida normalmente en la literatura como aquel trabajador que no cotiza al sistema de seguridad social del país. Esto con el fin de probar si esta infraestructura de tránsito mejora la asignación de trabajadores de economía informal a la formal, bajo premisa que trabajadores en lugares segregados prefieren trabajos informales mal remunerados cerca de su residencia en lugar de incurrir en altos costos de transporte para acceder a empleo formal (Zárate, 2020). Sin embargo, más allá del impacto sobre la mera informalidad, es de particular interés el impacto heterogéneo por cuenta propia y asalariados, para tener una mejor información sobre qué

tipo de ocupación, se reasignan mejor dada la mejora en el transporte. Los resultados parten de la Encuesta de Calidad de Vida de Medellín (ECV), la cual es representativa a nivel de estrato y comuna, y permite tener un escenario ideal con datos de experimento natural para evaluar el impacto del transporte. Mientras tanto, la información espacial se obtuvo de la plataforma Geomedellin de la alcaldía local.

El efecto causal de este sistema de transporte requiere entender el problema de inferencia causal subyacente. En primer lugar, en el efecto del tratamiento $\tau_1 = Y_i(1) - Y_i(0)$ para un individuo, uno de los dos resultados no es observable, sea resultado cuando es tratado o cuando no (Bernal & Peña, 2011), por lo que el análisis del tratamiento se debe hacer con el efecto promedio sobre las personas tratadas (ATT⁶). El análisis ATT clásico se presenta de la siguiente manera:

$$\delta_{ATT} = E[Y_i(1)|D_i = 1] - E[Y_i(0)|D_i = 1] \quad (1)$$

La ecuación (1) muestra el análisis de la diferencia entre media de la variable de interés (informalidad) de las comunas beneficiadas de la mejora $E[Y_i(1)|D_i = 1]$ y la media que hubiesen obtenido sin no hubiesen tenido dicha mejora $E[Y_i(0)|D_i = 1]$. No obstante, ambos resultados no se pueden observar al mismo tiempo, ya que el segundo resultado es hipotético (inobservable), lo cual se denomina en la literatura como contrafactual. Por ello, se debe hallar un grupo contrafactual (control) que se asemeje al grupo tratamiento antes de la intervención.

Ahora bien, como el interés es evidenciar efectos heterogéneos de la Línea J para trabajadores por cuenta propia y asalariados, el análisis ATT se estima a partir de una triple diferencia (DDD). Siguiendo a Olden & Moen (2020), el estimador de interés que acompaña la triple interacción entre indicador de tiempo (*Post*), de tratamiento (*T*) y variable de heterogeneidad (*B*) está dado por:

$$\begin{aligned} \delta_{ATT} = & [(E[Y|T = 1, B = 1, Post = 1] - E[Y|T = 1, B = 1, Post = 0]) \\ & - (E[Y|T = 1, B = 0, Post = 1] - E[Y|T = 1, B = 0, Post = 0])] \\ & - [(E[Y|T = 0, B = 1, Post = 1] - E[Y|T = 0, B = 1, Post = 0]) \\ & - (E[Y|T = 0, B = 0, Post = 1] - E[Y|T = 0, B = 0, Post = 0])] \end{aligned} \quad (2)$$

⁶ Average Treatment on the Treated, por sus siglas en inglés (ATT).

Intuitivamente, la ecuación (2) establece el análisis contrafactual entre grupos de trabajadores cuenta propia ($B = 1$) y asalariados ($B = 0$). La idea es comparar trabajadores cuenta propia en comuna de tratamiento con los cuenta propia en comuna de control. Según Olden & Moen (2020), esto no sería válido si diferentes comunas tienen diferentes condiciones económicas, de modo que los cuenta propia en la comuna de tratamiento habrían evolucionado de forma diferente a los cuenta propia en las comuna de control, independientemente de la Línea J. Sin embargo, se puede suponer razonablemente que diferencias económicas generales no afectarán resultados relativos de ambos grupos. En ese caso, se puede utilizar la diferencia relativa para estimar lo que habría ocurrido con los resultados relativos de cada grupo en comunas de tratamiento en ausencia de éste (en la sección 3.3 se amplía esta discusión a los supuestos de tendencias paralelas y comportamiento anticipado).

Sustituyendo los valores esperados por sus contrapartes muestrales, se obtiene:

$$\hat{\delta} = [(\bar{Y}_{trat,B,post} - \bar{Y}_{trat,B,pre}) - (\bar{Y}_{control,B,post} - \bar{Y}_{control,B,pre})] - [(\bar{Y}_{trat,A,post} - \bar{Y}_{trat,A,pre}) - (\bar{Y}_{cont,A,post} - \bar{Y}_{cont,A,pre})] \quad (3)$$

La ecuación (3) es el estimador de triple diferencia que mostrará el efecto diferencial del Metrocable Línea J entre trabajadores cuenta propia y asalariados. Así, \bar{Y} es la media del resultado de mi interés, una variable binaria que toma valor de uno si la persona es trabajador informal y cero si es formal, por tanto, se mide la probabilidad de ser trabajador informal. B es el grupo de cuenta propia, A el grupo de asalariados y, pre y $post$ son los periodos previos y posteriores a la entrada del proyecto. La primera doble diferencia sería el efecto tratamiento para trabajadores cuenta propia, mientras que la segunda sería el efecto para asalariados.

Según lo anterior, la especificación vía regresión del estimador de triple diferencia está dada, en forma general, por:

$$Y_{it} = c + \alpha_0 T + \alpha_1 B + \alpha_2 Post + \alpha_3 T * B + \alpha_4 T * Post + \alpha_5 B * Post + \delta T * B * Post + \varepsilon_{it} \quad (4)$$

Estimar el impacto causal requiere de una zona que sirva de contrafactual (una zona de tratamiento si no hubiese existido la intervención), para zona de intervención. Comúnmente, la literatura sobre impactos del sistema de transporte usa diferencias en diferencias que suelen identificar zonas de control aquellas donde se planearon proyectos y no se ejecutaron o aquellas que en un futuro cercano recibirán proyectos de transporte basados en la similitud aparente antes de la intervención (Moreno & Roman, 2019). Sin embargo, en este trabajo se recurre a la estrategia de Control Sintético (CS) que ofrece un método formalizado y riguroso para identificar casos de comparación.

El CS fue propuesto por Abadie & Gardeazabal (2003) y Abadie et al., (2010) como una generalización del enfoque de diferencias en diferencias. Básicamente sirve para identificar casos de comparación mediante la construcción de zona de control “sintética” que representa combinación ponderada de varios casos no tratados llamados “grupo de donantes”. Los pesos se calculan para maximizar la similitud entre el control sintético y el tratamiento en términos de variables de “coincidencia” especificadas. Este enfoque tiene varias ventajas: primero, maximiza la similitud observable de los casos de control y tratamiento, fortaleciendo los supuestos como tendencias paralelas inherentes al enfoque DD. Segundo, el método es factible incluso cuando no existe un solo caso no tratado suficientemente similar al caso de tratamiento. Por último, los investigadores pueden apuntar a un enfoque formal y objetivo para la selección de zonas de control, en vez de justificar elecciones ad hoc (Robbins & Davenport, 2021).

En esta sección se explicará formalmente el análisis usual de CS. Este enfoque de CS se realiza a nivel agregado suponiendo $J+1$ zonas con $j = 1, 2, \dots, J+1$. La primera ($j=1$) es zona impactada por intervención, que para este caso son dos, comunas San Javier y Robledo. El conjunto de comparaciones potenciales llamado “conjunto de donantes” $j = 2, \dots, J + 1$ es grupo de zonas no afectadas por intervención, que serían el resto de comunas de la ciudad. Se suponen T períodos (2007, 2008, 2013 y 2018), donde T_0 son períodos anteriores a la intervención (en este caso solo uno, 2007). Para cada zona j y tiempo t , se observa resultado de interés Y_{jt} y conjunto de k predictores del resultado, X_{1j}, \dots, X_{kj} . Asimismo, para cada zona j y t se define Y_{jt}^{NI} como respuesta potencial sin intervención (contrafactual) y para zona intervenida $j=1$ post-tratamiento $t > T_0$ se define Y_{jt}^I como respuesta potencial con intervención. De este modo, efecto del tratamiento para zona de intervención vendría dado por $\tau_{1t} = Y_{1t}^I - Y_{1t}^{NI}$. Es importante notar que efecto cambia con

el tiempo, porque los efectos de la intervención pueden no ser instantáneos y pueden acumularse o disiparse con el tiempo después de la intervención (Abadie, 2020).

El interés de este método es estimar el contrafactual para la zona de tratamiento (Y_{1t}^{NI}) o el valor de la variable de resultado que se habría observado para la zona intervenida en ausencia de la intervención, para lo cual, se genera un promedio ponderado del conjunto de donantes mediante un vector $J \times 1$ de pesos no negativos, $\mathbf{W} = (w_2, \dots, w_{j+1})'$. Dado \mathbf{W} , los estimadores del contrafactual Y_{1t}^{NI} y el efecto tratamiento τ_{1t} son, respectivamente:

$$\hat{Y}_{1t}^{NI} = \sum_{j=2}^{j+1} w_j Y_{jt} \quad (5)$$

$$\hat{\tau}_{1t} = Y_{1t} - \hat{Y}_{1t}^N \quad (6)$$

Ahora bien, Abadie et al., (2010) proponen elegir w_2, \dots, w_{j+1} de manera que la variable de resultado del control sintético se asemeje lo más posible a valores de variable de resultado antes de la intervención para zona tratada. Esto es, dado un conjunto de constantes no negativas, v_1, \dots, v_k que reflejan importancia relativa del control sintético, los pesos óptimos, $\mathbf{W}^* = (w_2^*, \dots, w_{j+1}^*)'$ deberían minimizar $|\mathbf{X} - \mathbf{X}_0 \mathbf{W}|$ (ver ecuación (9) en apéndice A) sujeto a la restricción de que $w_2, \dots, w_{j+1} \geq 0$ y $\sum_{j=2}^{j+1} w = 1$. De manera intuitiva, esta restricción garantiza que un número limitado de zonas dentro del grupo de donantes contribuyan a estimar contrafactual y contribución de cada zona es representada por ponderación de control sintético. La ecuación (9) es una medida de distancia que indica reducción de distancia entre valores de una determinada covariable de zona tratada y resto de valores que toma esa covariable en el resto de zonas y se suma a través de las unidades no tratadas.

No obstante, este enfoque se ha aplicado en contexto de un solo caso de tratamiento con un número limitado de casos no tratados para la comparación, lo cual ocurre porque usualmente se usan datos agregados. Este procedimiento se realiza a través del paquete *Synth* en los software estadísticos, el cual está limitado a la escasez de casos de tratamiento y que a su vez limitan: desarrollar un control sintético que coincida exactamente con el caso de tratamiento, estimar con

precisión el efecto del tratamiento, medir la importancia de ese efecto e incorporar conjuntamente múltiples variables de resultado. Para el propósito de este estudio, se utiliza la metodología propuesta por Robbins & Davenport (2021) en el paquete *microsynth* que abarca las limitaciones antes descritas en la implementación con *Synth*. Este enfoque también es utilizado para datos de panel agregados a nivel granular (como manzanas), pero admite datos individuales de sección transversal repetido⁷, sección transversal ($T = 1$)⁸ y más de una zona de tratamiento.

Esta alternativa de control sintético supone que hay datos sobre J casos totales de manera que los primeros J_0 no son tratados o “grupo de donantes” y los últimos $J - J_0$ son tratados. Asimismo, como en *Synth*, hay T períodos de tiempo totales (2007, 2008, 2013 y 2018), con T_0 periodos previos a la intervención (2007) y $T - T_0$ (2013 y 2018) posteriores a la intervención⁹. Hay L variables de resultado que varían en el tiempo, siendo Y_{ljt} el valor de la variable de resultado l para cada zona j en el momento t . También, supone k covariables invariantes en el tiempo (es decir, de línea de base) X_{1j}, \dots, X_{kj} para cada zona j . Se calcula un grupo de control sintético asignando un peso a cada zona no tratada. Estas ponderaciones se denominan w_j para $j \in (1, \dots, J_0)$. El algoritmo *microsynth* hace coincidir tratamiento y control sintético calculando pesos que satisfacen tres restricciones. En primer lugar, la suma de pesos es igual a número de casos en el área de tratamiento. Esta restricción tiene la misma intuición que la restricción de pesos de *Synth*. En segundo lugar, el control sintético agregado (ponderado) coincide con el área de tratamiento agregada a través de las covariables. Por último, el control sintético y el tratamiento también coinciden en todos los puntos temporales previos a la intervención de cada variable de resultado (ver ecuación (10), (11) y (12) en apéndice A para formalización de estas restricciones).

De esta manera, se puede configurar una estrategia de identificación que combine el método de triple diferencia (DDD) con CS. La idea no es estimar el efecto de tratamiento con CS, es ponderar las observaciones de estas zonas de control sintético por sus respectivos pesos y equilibrar las características de individuos en estas zonas con las de tratamiento. Esta estrategia ha sido

⁷ En cohorte transversal repetido, *microsynth*, agrega los datos individuales a través una variable que identifique múltiples observaciones de un mismo caso, para esta investigación es el identificador de comuna.

⁸ *microsynth* también se puede usar para calcular pesos de tipo puntaje de propensión con datos en sección transversal ($T=1$).

⁹ *microsynth* proporciona parámetros para ayudar a encontrar pesos factibles cuando hay escasez de datos, con lo cual no es inconveniente tomar solo un año pre-tratamiento.

utilizada por Bauhoff (2014) para evaluar efectos a corto plazo de la política de nutrición escolar de 2004 para Los Ángeles con un enfoque de diferencias en diferencias con control sintético usando datos de corte transversal repetido. El autor, luego de aplicar el control sintético para el periodo antes de la intervención, toma pesos y pondera observaciones de dichos controles. Una vez ponderadas las observaciones, estima un modelo DD. El estudio fue guiado por Alberto Abadie, por lo que es una alternativa robusta al método de control sintético tradicional cuando no se tienen disponibles datos de panel.

Por ello, la especificación para el efecto heterogéneo de la Línea J está dada por:

$$Y_{ijt}^b = c^b + \alpha_0^b T_i + \alpha_1^b I_t + \alpha_2^b D_i^b + \alpha_3^b T_i^b D_i^b + \alpha_4^b T_i^b I_t + \alpha_5^b D_i^b I_t + \delta^b I_t T_i^b D_i^b + \sum_{j=1}^p \beta^{b'} w_j \mathbf{X}_i^b + \varepsilon_{it}^b \quad (7)$$

Donde Y_{ijt} es la variable de resultado: dummy igual a uno si el individuo i es informal (no cotiza pensión). T_i representa la variable dicótoma de tratamiento, que toma valor de uno si trabajador reside en las comunas de tratamiento (San Javier y Robledo). I_t es uno si observación corresponde a periodo posterior a la intervención. D_i es efecto diferencial entre tipos de empleo y toma valor de uno si individuo i es cuenta propia. El término $I_t T_i D_i$ es la triple interacción entre variable de tratamiento, dummy del periodo y variable de heterogeneidad ocupacional. b es indicador que sugiere que la estimación se hace para diferentes zonas de influencia de la estación. Este indicador se define usando círculos concéntricos alrededor de las estaciones de la línea J, tomando zonas de influencia circulares de 0-500, 0-1000 y 0-1500 metros. UN-Habitat (2018) define como área de servicio del transporte público zonas a menos de 500 metros a pie de un sistema de transporte de baja capacidad¹⁰ y/o a 1000 metros de un sistema de mediana-alta capacidad¹¹. Este último se justifica en el hecho de que, es más probable que los pasajeros recorran a pie distancias más largas para acceder a los sistemas de transporte público de mediana-alta capacidad que las que recorrerían para acceder a los de baja capacidad. El área de influencia de 1500 metros se usa solo para verificar si más allá de los 1000 metros estándar se logra capturar un

¹⁰ Incluye sistemas como autobuses, tranvías y tránsito rápido de autobuses (BRT).

¹¹ Incluye sistemas como trenes, metros y teleféricos.

efecto, sobretodo en ciudades con topografía escarpada, donde se ha evidenciado que las pendientes elevadas reducen la actividad a pie y se puede optar por otras formas de transporte (Meeder et al., 2017). Los coeficientes de interés son α_4 y δ . La estimación DD usual del efecto del tratamiento para los asalariados es α_4 y la estimación DD del efecto del tratamiento para los cuenta propia sería $(\alpha_4 + \delta)^{12}$. La diferencia entre los dos efectos permite calcular los efectos diferenciales de la Línea J según el tipo de ocupación. De hecho, el efecto del tratamiento para los cuenta propia y asalariados difiere en $(\alpha_4 + \delta) - \alpha_4 = \delta$ que se supone que es el efecto causal heterogéneo.

El vector de características socioeconómicas X_i de trabajadores que ayudan a controlar los efectos del tratamiento estarán ponderadas con w_j para las comunas de control sintético, lo cual garantiza la similitud entre zonas y controlar los factores de confusión no medidos. Debido a que la variable de respuesta es uno si el individuo es informal y cero si es formal, se utiliza un Modelo Lineal de Probabilidad (MLP) para estimar los parámetros en vez de modelos no lineales como logit o probit. El uso de MCO en resultado dicotómico presenta ventajas respecto a modelos no lineales: es fácil de estimar, la interpretación del coeficiente de interacción es fácil y hay posibilidad de corregir correlación de errores. Estas ventajas se pierden en los modelos no lineales, sobre todo porque estos son más complejos de estimar y no están dentro del alcance de este estudio (Ai & Norton, 2003). Los inconvenientes por superar son: heterocedasticidad y predicción por fuera del intervalo 0 y 1. El primero es fácil de superar, mientras que el segundo no, pero el interés es el impacto causal y no la valores pronosticados individuales. Ahora bien, para la estimación principal se tienen 5 periodos de observación, dos años antes de la puesta en marcha del MCJ que sería el periodo de pre-tratamiento (2005 y 2007) y 2 años posteriores a la puesta en marcha (2010 y 2013), teniendo como año de tratamiento 2008 cuando se inauguró la línea. Se estima la ecuación (7) a nivel general para las comunas tratadas y control sintético y para las diferentes áreas de influencia definidas antes. Cabe aclarar que el control sintético se hace a nivel comuna, mientras que la estimación DDD se hace a nivel individual.

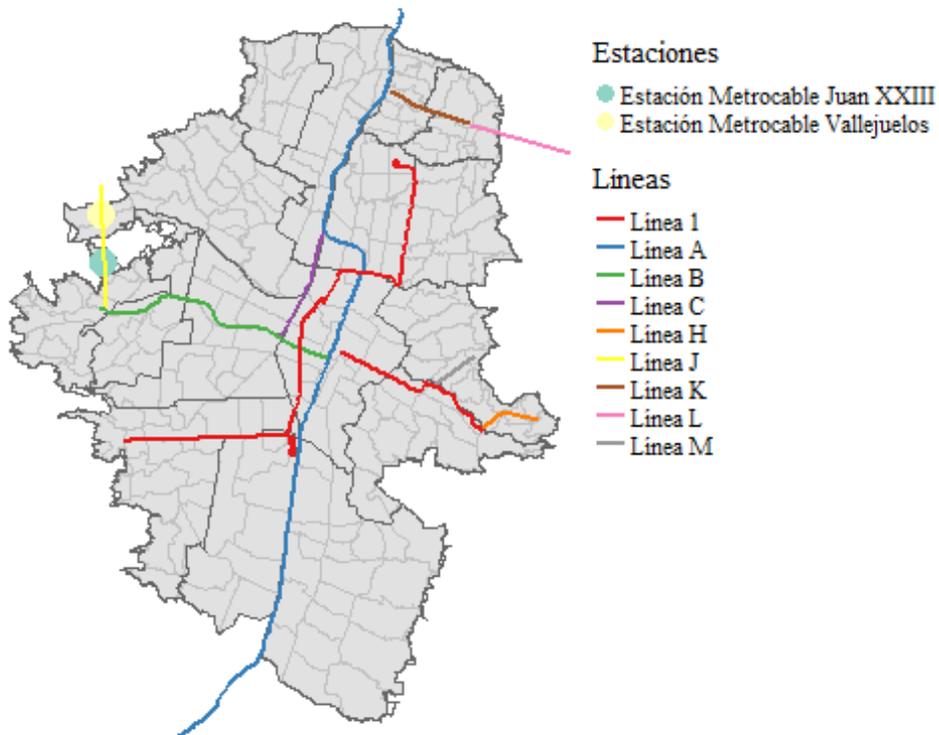
A continuación, la **Figura 1** presenta la distribución del transporte público en Medellín según el Plan de Ordenamiento Territorial (POT) 2014. Las estaciones de la Línea J de Metrocable conectan con el resto de la ciudad a través de la Línea B (San Javier - San Antonio). Otras Líneas

¹² Para obtener este estimador con su error estándar se debe cambiar la categoría base de la dummy tipo de empleo.

de Metrocable están distribuidas hacia el norte de la ciudad con la Línea K y en la zona oriental con la Línea H y M que también conectan con el centro a través del Metro.

Figura 1.

Distribución de infraestructura de transporte en Medellín



Fuente: Elaboración propia con datos Plan de Movilidad de Medellín 2006-2030 GeoMedellín

3 Resultados

3.1 Control sintético

La **Tabla 1** resume resultados de aplicar metodología de control sintético de (Robbins & Davenport, 2021). Las primeras tres columnas muestran correspondencia del tratamiento con el control sintético en covariables invariantes en el tiempo (sexo, estado civil, parentesco y tipo de empleo)¹³ y variables de resultado que varían en el tiempo (informalidad y gasto del hogar en logaritmo), aprovechando la ventaja que ofrece *microsynth* de incluir más de una variable de resultado para que haya mayor precisión en el ajuste. Se confirma que cumple todas las restricciones descritas en 2, ya que entradas de la columna de tratamiento son aproximadamente iguales a las de la columna Control Sintético (ver ecuación (10), (11) y (12) en apéndice A). Obsérvese que primera fila (Intercepto) indica que hay 2 comunas de tratamiento y que ponderaciones del control sintético suman ese valor y los resultados de las filas siguientes son indicativos y no sugieren aquí una interpretación económica. En la sección 3.2 se mostrará estas correspondencias a través de un ejercicio de diferencia de medias.

Del lado derecho están los cálculos de pesos para cada comunas de control potenciales, mostrando que comunas Doce de Octubre (0.9716770), Poblado (0.0552449) y Belén (0.9730780) conforman unidades de control sintético que se usaran como contrafactual para la estimaciones de la triple diferencia. Empero, se excluye la comuna Poblado, pues es una comuna con bajas tasas de precariedad laboral en la ciudad (lo cual puede confundir el efecto del tratamiento) y su peso es muy bajo. La **Figura 2** muestra la identificación de zonas de tratamiento, con un área general donde se incluyen todos los barrios de las comunas San Javier y Robledo y zonas de influencia circulares de 500, 1000 y 1500 metros. Se observa que las zonas de influencia de 1000 y 1500 metros toman barrios de La América (comuna 12) y los barrios alrededor de la estación San Javier (San Javier N°1 y San Javier N°2), pero estos se omiten en las estimaciones para evitar factores de confusión. Otra preocupación es que la ruta de Metroplús en la comuna 16 también puede generar factores de

¹³ Se usaron estas covariables ya que el uso de *microsynth* requiere que las covariables sean invariantes en el tiempo y las variables de resultados varíen con el tiempo. Haciendo la verificación de la variabilidad, se encuentra que estas son las que menos varían en el tiempo.

confusión, sin embargo, para no omitir esta comuna como zona de control, se deja que las ponderaciones de CS ayuden a controlar este factor de confusión no medido.

Tabla 1.

Control Sintético

Variable	Tratados	Control Sintético	Comunas	Pesos
Intercepto	2	2	Popular (1)	0.0000000
Sexo	0	1.45232e-08	Santa Cruz (2)	0.0000000
Estado civil	0	0.05524492	Manrique (3)	0.0000000
Tipo empleo	1	0.9730781	Aranjuez (4)	0.0000000
Parentesco	2	2	Castilla (5)	0.0000000
Informales (2007)	1	0.973078	Doce octubre (6)	0.9716770
Gasto (2007)	26.76352	27.6822	Villa Hermosa (8)	0.0000000
			Buenos Aires (9)	0.0000000
			La Candelaria (10)	0.0000000
			Laureles-Estadio (11)	0.0000000
			La América (12)	0.0000000
			Poblado (14)	0.0552449
			Guayabal (15)	0.0000000
			Belén (16)	0.9730780
			Total	1.9999999

Nota: Número de la comuna en paréntesis.

Fuente. Cálculos propios con ECV Medellín.

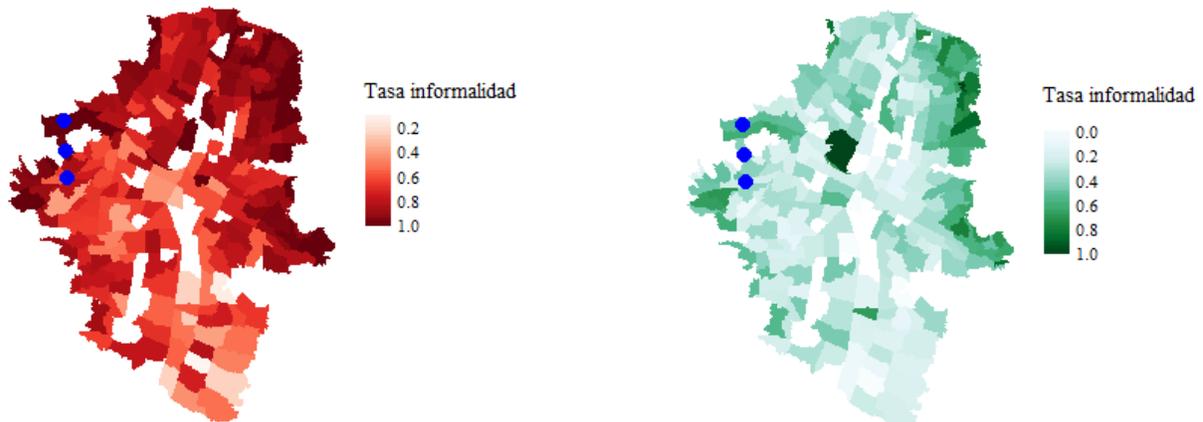
Figura 2.*Definición de zonas de tratamiento y Control Sintético*

Fuente. Elaboración propia con Plan de Movilidad de Medellín 2006-2030 GeoMedellín.

3.2 Estadísticas descriptivas

La **Figura 3** muestra la distribución espacial de la tasa de informalidad (TI) de cuenta propia **3(a)** y asalariados **3(b)** a nivel de barrio. A pesar de que informalidad es mayor para los cuenta propia, que, para asalariados, se observa un patrón espacial en ambos grupos; sus tasas de informalidad más altas se concentran en zonas periféricas de la ciudad. En particular, se observa mayor incidencia de empleo informal por cuenta propia que asalariado en comunas donde se ubica la Línea J.

La **Tabla 2** presenta un conjunto de características socioeconómicas de trabajadores para el año pre-tratamiento 2007 en un buffer de 1000 metros. Se compara a través de una prueba de diferencia de medias el grupo de tratamiento y control sintético, indicando las medias para las observaciones ponderadas por pesos de las comunas que arrojó el control sintético (w) y las no ponderadas en estas mismas comunas. Incluyo características como el sexo, estado civil, parentesco, tipo de empleo (cuenta propia o asalariado), edad y años de educación. Los datos revelan que las diferencias entre comunas tratadas y control sintético en pretratamiento se reducen con la ponderación, lo que sugiere que la restricción dos tiende a cumplirse (ver ecuación (11) en apéndice A). No obstante, hay diferencias en educación promedio y porcentaje de jefes que se pueden controlar en la estimación. Este mismo ejercicio se realizó por cuenta propia y asalariados con resultados similares (ver **Tabla 6** y **Tabla 7** en anexo 1).

Figura 3.*Distribución espacial de informalidad cuenta propia y asalariados**(a) Distribución informalidad cuenta propia**(b) Distribución informalidad asalariados*

Fuente: Elaboración propia con datos de ECV 2008 y GeoMedellín.

Tabla 2.*Resumen estadísticas a nivel individuo 2007 con zona de influencia 1000 metros*

	Control Sintético	Tratamiento	Valor P	Sig.
% Hombre	0.4157 (0.4929)	0.3998 (0.4901)	0.361	
% Hombre (w)	0.4043 (0.4794)	0.3998 (0.4901)	0.796	
% Casados	0.5082 (0.5000)	0.4796 (0.4998)	0.106	
% Casados (w)	0.4943 (0.4863)	0.4796 (0.4998)	0.404	
% Jefes	0.4487 (0.4974)	0.5020 (0.5003)	0.003	**
% Jefes (w)	0.4363 (0.4837)	0.5020 (0.5003)	<0.001	**
% Cuenta propia	0.2805 (0.4493)	0.2444 (0.4299)	0.019	**
% Cuenta propia (w)	0.2728 (0.4369)	0.2444 (0.4299)	0.063	
Edad promedio	38.9593 (12.2051)	38.3763 (12.3726)	0.183	
Edad promedio (w)	37.8883 (11.8715)	38.3763 (12.3726)	0.263	
Educación promedio	9.4195 (4.6106)	8.0798 (4.2004)	<0.001	**
Educación promedio (w)	9.1613 (4.4860)	8.0798 (4.2004)	<0.001	**

Nota: **w** es el peso de las comunas de CS, Sig.: Significancia, desviación estándar entre paréntesis. Se omiten de la muestra los individuos residentes en los barrios San Javier N°1, San Javier N°2 y Las Independencias para evitar factores de confusión por el efecto spillover que genera la actividad turística en Las Independencias.

Fuente: Elaboración propia con datos ECV 2007.

3.3 Estudio de eventos y regresiones placebo

Dos supuestos de identificación del efecto causal en la estrategia DD estándar y que se pretenden evidenciar aquí son: tendencias paralelas y comportamiento no anticipado pre-tratamiento. El primero sugiere que ambos grupos deberían tener la misma tendencia de la variable de resultado en pre-tratamiento. En el caso del estimador DDD, aunque sea la diferencia entre dos diferencias en diferencias, no necesita dos supuestos de tendencia paralela, más bien, requiere que el diferencial del resultado del grupo *B* (cuenta propia) y del grupo *A* (asalariados) en la comuna de tratamiento tenga la misma tendencia que el diferencial del resultado del grupo *B* y del grupo *A* en el control sintético, en ausencia de tratamiento (Olden & Moen, 2020; Sun & Abraham, 2020). El segundo busca que se garantice que los individuos de la zona de tratamiento no hayan cambiado su comportamiento antes de la entrada de la Línea J (Sun & Abraham, 2020).

Si no se cumplen estos supuestos, no será válido utilizar los resultados observados del control sintético después del tratamiento como contrafactual para los resultados después del tratamiento para los tratados. Hay varias formas comunes para evaluar tendencias paralelas, las cuales, en todo caso, se consideran evidencias sugestivas, pero ofrecen una formalización a estos supuestos (Huntington-Klein, 2014). En este estudio se usarán dos: la primera, un diseño de triple diferencia de estudio de eventos con variables binarias para cada periodo antes (2005, 2006 y 2007) y después (2010, 2013 y 2018) usando como categoría base el año de apertura de MCJ, 2008¹⁴. Aprovechando la especificación dinámica de estudio de eventos, se eligen estos años como post-tratamiento para verificar si existen efectos de corto (luego de dos años, 2010), mediano (después de 5 años, 2013) y largo plazo (10 años después, 2018). La segunda, una regresiones placebo¹⁵ con solo observaciones pre-tratamiento (2005, 2006 y 2007). Se usa 2006 y 2007 como periodos placebo (falso) post-tratamiento. Lo anterior puede ofrecer cierta evidencia para la identificación y validez

¹⁴ Aprovechando que MCJ entró en operación el 3 marzo de 2008 y la fecha de recolección de ECV 2008 empezó en noviembre de ese año y terminó en enero 2009, hay dos ventajas al tomar 2008 como base: primero, permite observar todos los efectos antes de la apertura de MCJ y segundo, todos los efectos se interpretan respecto a un año donde se supone razonablemente un efecto, ya que ECV 2008 pudo haber capturado cambios generados por MCJ respecto a ECV de años anteriores. Lo anterior porque la literatura suele usar como base el año justo antes de la intervención, lo que supone un efecto nulo en este año, pero si no fuese nulo, los resultados estarían sesgados, dado que la interpretación es relativa al periodo de referencia (Huntington-Klein, 2014).

¹⁵ Martínez et al., (2018) usa la misma prueba para su triple diferencia.

interna de las estimaciones de interés¹⁶. Explotando los datos previos al tratamiento, la especificación de regresión placebo es igual a (7) pero con $t \in \{2005, 2006, 2007\}$ y se espera δ^b sea no significativo, mientras que estudio de eventos para una estrategia DDD está dada por:

$$Y_{ijt} = c^b + \sum_{\tau=2005}^{2018} \gamma_{\tau}^b I_{i\tau} + \alpha^b T_i + \pi^b D_i + \sum_{\tau=2005}^{2018} \theta_{\tau}^b I_{i\tau} T_i + \sum_{\tau=2005}^{2018} \alpha_{\tau}^b I_{i\tau} D_i + \lambda^b T_i D_i + \sum_{\tau=2005}^{2018} \delta_{\tau}^b I_{i\tau} T_i D_i + \sum_{i=1}^p \beta^{i'b} w_j X_i + \varepsilon_{it}^b \quad (8)$$

Todo se define igual que (7). Para la ecuación (8) interesa los coeficientes δ_{τ} con $\tau \in \{2005, 2006, 2007, 2008, 2013, 2018\}$. Si δ_{τ} para $\tau \leq 2007$ resultan estadísticamente iguales a cero, sugiere que la hipótesis de tendencia paralela se mantiene para la triple diferencia¹⁷. Los coeficientes para $\tau \geq 2010$ muestra una primera estimación del impacto causal de interés. Además, si tendencias paralelas se cumple, es sensato pensar que no hubo comportamiento anticipado que pudiera confundir los resultados.

La **Tabla 3** muestra los resultados de ambas pruebas, el Panel A muestra la regresión de estudio de eventos, mientras que Panel B muestra regresión placebo. La columna 1 indica una estimación para todos los individuos de comunas intervenidas y las columnas 2 a 4 indican estimaciones con individuos en las zonas de influencia. El Panel A se divide en tres partes, las tres primera filas informan los efectos post-tratamiento para asalariados, las siguientes tres los efectos post-tratamiento para trabajadores cuenta propia y las seis últimas, el efecto diferencial para cada año. Estas últimas son las que ofrecen evidencia de tendencias paralelas en DDD. Los resultados sugieren, en general, que la hipótesis de tendencias paralelas se mantiene, con algún efecto en 2007 (columnas 2) y 2005 (columna 3). Lo anterior indica que, en un área de 500 metros, las tendencias

¹⁶ Otras pruebas pueden ser el método gráfico y regresiones placebo con grupos no tratados. Este último no se usa porque no aborda el problema de tendencias paralelas directamente. Huntington-Klein (2014) sugiere que no es realmente un problema si las tendencias paralelas fallan entre los grupos no tratados.

¹⁷ Estudios que han probado tendencias paralelas usando DDD de estudio de eventos, regresiones placebo o ambos, solo muestran los resultados de triple interacción para verificar la hipótesis de tendencias paralelas. (Basant & Sen, 2019; Martinez et al., 2018; Muralidharan & Prakash, 2017).

eran consistentes entre sí, pero ocurrió una desviación en 2007, mientras que en área de 1000 metros hubo una brecha en 2005 y luego se mantuvieron constantes en los siguientes años pre-tratamiento. Una ventaja del estimar DDD con estudio de eventos es que muestra efectos para cada año post-tratamiento. Se evidencia que luego de 2 años (2010) de la entrada de MCJ, no se observan efectos significativos ni en asalariados ni en cuenta propia. Después de 5 años (2013), la informalidad en asalariados se reduce de manera significativa, mientras que en cuenta propia el efecto es nulo. Por último, el efecto se pierde en el largo plazo (2018). Asimismo, Panel B indica efectos placebo de tratamiento estadísticamente no significativos. A pesar de que las brechas encontradas en estudio de eventos son significativas y amplias, no se logran mantener en todos los años pre-tratamiento, por tanto, la validez de las estimaciones adelante no está comprometida del todo. La **Tabla 8** del anexo 2 muestra el estudio de eventos y regresión placebo de diferencias en diferencias estándar con evidencia consistente de esta hipótesis.

Tabla 3.*Pruebas de tendencias paralelas para DDD*

	E(Informal = 1 X)			
	General (1)	500 (2)	1000 (3)	1500 (4)
Panel A: Estudio de eventos				
I ₂₀₁₀ *MCJ (Asalariados)	-0.0237 (0.031)	-0.011 (0.0396)	-0.011 -0.0361	-0.0252 -0.0345
I₂₀₁₃*MCJ (Asalariados)	-0.0979 *** (0.0347)	-0.0977 (0.071)	-0.1022 *** -0.0431	-0.1156 *** -0.0388
I ₂₀₁₈ *MCJ (Asalariados)	-0.0477 (0.0354)	0.0183 (0.037)	0.0081 (0.0377)	-0.0428 (0.0413)
I ₂₀₁₀ *MCJ (Cuenta propia)	0.0385 (0.0368)	0.0037 (0.0567)	-0.0071 (0.0428)	0.0292 (0.0414)
I₂₀₁₃*MCJ (Cuenta propia)	0.027 (0.036)	0.0224 (0.047)	-0.0065 (0.042)	0.013 (0.0379)
I ₂₀₁₈ *MCJ (Cuenta propia)	-0.0637 (0.0558)	-0.1261 (0.1359)	-0.1359 (0.0845)	-0.0852 (0.0707)
I ₂₀₀₅ *MCJ*Cuenta propia	-0.0709 (0.0432)	-0.0449 (0.0413)	-0.1458 ** (0.0571)	-0.0760 (0.0569)
I ₂₀₀₆ *MCJ*Cuenta propia	0.0272 (0.0404)	0.0400 (0.0548)	-0.0399 (0.0538)	0.0067 (0.0479)
I ₂₀₀₇ *MCJ*Cuenta propia	-0.0368 (0.0529)	-0.1643 *** (0.0543)	-0.1034 (0.0646)	-0.0878 (0.0627)
I ₂₀₁₀ *MCJ*Cuenta propia	0.0622 (0.0428)	0.0147 (0.0527)	0.0038 (0.0489)	0.0544 (0.0466)
I₂₀₁₃*MCJ*Cuenta propia	0.1249 ** (0.0493)	0.1202 (0.0818)	0.0957 (0.0626)	0.1286 ** (0.0571)
I ₂₀₁₈ *MCJ*Cuenta propia	-0.0160 (0.0664)	-0.1443 (0.1212)	-0.1439 (0.0912)	-0.0424 (0.0829)
Controles	Si	Si	Si	Si
N	42387	26316	29694	32471
Adj. R-squared	0.2629	0.2648	0.2654	0.2638
Panel B: Regresión placebo 2006-2007				
I₀₆₋₀₇*MCJ*cuenta propia	0.0656 * (0.0372)	-0.0097 (0.0676)	0.0744 (0.0452)	-0.0370 (0.0421)
Controles	Si	Si	Si	Si
N	23081	14157	15999	17575
Adj. R-squared	0.2096	0.2284	0.2265	0.2198

Nota: ***p < .01; **p < .05; *p < .10. Errores estándar entre paréntesis agrupados a nivel barrio, MCJ refiere Metro Cable Línea J. Se omiten de la muestra los individuos residentes en los barrios San Javier N°1, San Javier N°2 y Las Independencias para evitar factores de confusión por el efecto spillover que genera la actividad turística en Las Independencias.

Fuente. Cálculos propios con datos ECV Medellín.

3.4 Efecto causal heterogéneo del Metrocable línea J

Antes de pasar al efecto heterogéneo, muestro una estimación DD estándar (sin diferenciar por tipo de empleo) en la **Tabla 9** del anexo 2, siguiendo la misma intuición de columnas en las estimaciones de sección 3.3. En esta se busca verificar si MCJ reduce las tasas de informalidad en zonas intervenidas. Se observa un efecto de reasignación entre los empleados en zonas de tratamiento, aunque con reducciones significativas en la estimación general y a 1500 metros. Esto va en línea con evidencias anteriores sobre el impacto del transporte público en la informalidad laboral a nivel general para la región metropolitana de Sao Paulo (Boisjoly et al., 2017; Moreno & Roman, 2019; Posada & García-Suaza, 2021).

Respecto al efecto heterogéneo, la **Tabla 4** muestra la estimación de la ecuación (7) para toda la zona de tratamiento (columna 1) y para diferentes buffer alrededor de la Línea J (columnas 2:4) usando los controles expuestos en la Tabla 2. Esta estimación refleja el impacto causal heterogéneo por tipo de ocupación. Por un lado, indica el efecto de la Línea J para los asalariados ($\hat{\alpha}_4$) y cuenta propia ($\hat{\alpha}_4 + \hat{\delta}$) y por otro, el efecto diferencial o brecha entre los efectos de cuenta propia y asalariado ($\hat{\delta}$).

Todas las columnas informan que la infraestructura de transporte aumentó significativamente la informalidad para los cuenta propia y disminuyó de manera significativa para los asalariados. En particular, la columna 1 muestra que luego de la entrada de MCJ, la informalidad por cuenta propia aumentó 5,8 p.p más en zonas de tratamiento que en zonas de CS, mientras que la informalidad en asalariados se redujo 5,9 p.p más en comunas de intervención que en CS. Con esto, se evidencia que ambos efectos difieren significativamente en 11,7 p.p. Por su parte, las columnas 2 a 4 muestran que el efecto se reduce tanto para asalariados como para cuenta propia a medida que aumenta el área de influencia. En particular, para trabajadores asalariados en una zona de influencia de 500 metros se reduce la informalidad en 6,4 p.p, mientras que en un área de influencia de 1500 metros se redujo 6,2 p.p. En cambio, para los cuenta propia en una zona de influencia de 500 metros, aumentó la informalidad en 8,5 p.p, mientras que en zona de influencia de 1500 metros aumentó 7,8 p.p. A pesar de que el efecto en una zona de influencia mayor a la estándar se reduce en mayor proporción, se encuentra un efecto significativo, lo que demuestra que las personas pueden estar dispuestas a caminar hasta 15 minutos para acceder a transportes de

mediana-alta capacidad. Esto va en línea con estudios que evidencian efectos heterogéneos del transporte público, como el caso del BRT y la línea 1 de tren ligero elevado en Lima sobre la participación laboral (Martinez et al., 2018) y el Metro de Delhi en India sobre el empleo (Seki & Yamada, 2020), ambos diferenciados por género.

Lo anterior sugiere evidencia en favor de que la infraestructura de transporte si logra tener efectos heterogéneos sobre la informalidad, mostrando que los asalariados en este contexto aprovechan la conexión que ofrece el transporte público con las zonas centrales de empleo para transitar hacia empleos formales, es decir, MCJ tiene efectos de reasignación en estos trabajadores. Mientras tanto, es posible que el resultado de los cuenta propia este explicado por dos posibles razones relacionadas: según datos de la ECV para 2008¹⁸, si bien tanto trabajadores cuenta propia (71,5%) como asalariados (94%) laboran lejos del barrio donde residen, la proporción de cuenta propia que trabaja en el mismo barrio (28,6%) es significativamente mayor a la de asalariados (6%), lo que sugiere que los cuenta propia tienden a valorar los empleos informales cercanos. Debido a esto, es sensato pensar que MCJ atrajo aglomeraciones de puestos de trabajo cuenta propia y que estos aprovechan de otra forma la infraestructura de transporte público. Por ejemplo, aquellos trabajadores cuenta propia que tienen negocios pequeños pueden aprovechar el flujo de pasajeros para aumentar sus ingresos. Esto coincide con datos de UN-Habitat (2018b), que reportan que se ha incrementado en un 40% la actividad comercial con pequeñas empresas familiares y restaurantes alrededor de las estaciones de Metrocable.

¹⁸ La variable sobre si trabajan lejos o cerca del barrio donde residen solo está planteada en para el año 2008, por ello no se consideró dentro de la estimación.

Tabla 4.*Impacto heterogeneo del Metrocable Linea J sobre la informalidad laboral*

	E(Informal = 1 X)			
	General (1)	500 (2)	1000 (3)	1500 (4)
MCJ	0.069 ** (0.029)	0.083 *** (0.029)	0.067 ** (0.033)	0.086 *** (0.030)
Tiempo	-0.122 *** (0.018)	-0.122 *** (0.018)	-0.121 *** (0.018)	-0.121 *** (0.018)
Cuenta propia	0.370 *** (0.015)	0.367 *** (0.015)	0.368 *** (0.015)	0.368 *** (0.015)
Tiempo*MCJ (asalariados)	-0.059 ** (0.026)	-0.064 ** (0.026)	-0.06308 ** (0.029)	-0.06219 ** (0.027)
MCJ*Cuenta propia	-0.061 *** (0.022)	-0.055 *** (0.020)	-0.052 ** (0.025)	-0.071 *** (0.025)
Tiempo*Cuenta propia	0.042 ** (0.019)	0.042 ** (0.019)	0.042 ** (0.019)	0.042 ** (0.019)
Tiempo*MCJ*Cuenta propia	0.117 *** (0.030)	0.149 *** (0.026)	0.153 *** (0.035)	0.141 *** (0.038)
Constante	0.835 *** (0.029)	0.843 *** (0.030)	0.845 *** (0.030)	0.845 *** (0.029)
Tiempo*MCJ (Cuenta propia)	0.05803 *** (0.02171)	0.085 *** (0.03374)	0.08978 *** (0.02146)	0.0788 *** (0.02622)
Controles	Si	Si	Si	Si
N	31345	19658	22096	24007
Adj. R-squared	0.245	0.245	0.246	0.245

Nota: ***p < .01; **p < .05; *p < .10. Errores estándar entre paréntesis agrupados a nivel barrio, MCJ refiere Metro Cable Línea J. Se omiten de la muestra los individuos residentes en los barrios San Javier N°1, San Javier N°2 y Las Independencias para evitar factores de confusión por el efecto spillover que genera la actividad turística en Las Independencias.

Fuente. Cálculos propios con datos ECV Medellín.

3.5 Chequeo de robustez

Para el chequeo de robustez se realizan estimaciones eligiendo un contrafactual de forma ad hoc. En este caso, aprovechando los proyectos del plan de expansión Metrocable, se escoge como grupo de control la comuna 6 (12 de octubre), dado que aquí se tenía previsto la entrada de Metrocable Línea P, beneficiando 350 mil habitantes aproximadamente. Esta estrategia ha sido ampliamente utilizada en la literatura (Moreno & Roman, 2019; Posada & García-Suaza, 2021; Zárate, 2020) y se basa en el supuesto de similitud aparente respecto a características relevantes para la asignación de infraestructura de transporte (Moreno & Roman, 2019). Como el dominio temporal de este trabajo es hasta 2013, tiene sentido considerar la comuna 6 como contrafactual porque la Línea P recién se inauguró en 2021.

La **Tabla 5** muestra estimaciones consistentes respecto a la estimación con control sintético, la informalidad se reduce para los asalariados y aumenta para los cuenta propia en las comunas tratadas respecto a la comuna de control. Los efectos en cada estimación son superiores a los 2 p.p en asalariados y alrededor de 5 p.p en cuenta propia, aunque no son significativos estadísticamente, a excepción del efecto calculado en una zona de influencia a 1000 metros para trabajadores cuenta propia. Esto demuestra que el CS parece ser robusta frente a una alternativa ad hoc, que permite controlar factures de confusión invariantes en el tiempo y las diferencias preexistentes entre grupos.

Tabla 5.

Impacto heterogéneo del Metrocable Línea J sobre la informalidad laboral (comuna 6 como contrafactual)

	E(Informal = 1 X)			
	General (1)	500 (2)	1000 (3)	1500 (4)
MCJ	0.02167 (0.03084)	0.02582 (0.03713)	0.01509 (0.03547)	0.03950 (0.03443)
Tiempo	-0.16108 *** (0.02899)	-0.16161 *** (0.02974)	-0.16062 *** (0.02941)	-0.16030 *** (0.02931)
Cuenta propia	0.36133 *** (0.02223)	0.36395 *** (0.02281)	0.36260 *** (0.02259)	0.36089 *** (0.02251)
Tiempo*MCJ (asalariados)	-0.02340 (0.03423)	-0.02539 (0.03458)	-0.02744 (0.03703)	-0.02815 (0.03525)
MCJ*Cuenta propia	-0.04635 * (0.02759)	-0.04248 * (0.02562)	-0.04200 (0.03023)	-0.05971 ** (0.02971)
Tiempo*Cuenta propia	0.10848 *** (0.02591)	0.10929 *** (0.02670)	0.10857 *** (0.02628)	0.10794 *** (0.02613)
Tiempo*MCJ*Cuenta propia	0.05144 (0.03453)	0.07806 ** (0.03216)	0.08729 ** (0.03907)	0.07664 * (0.04274)
Constante	0.98114 *** (0.03491)	0.93366 *** (0.03708)	0.97290 *** (0.03785)	0.97988 *** (0.03638)
Tiempo*MCJ (Cuenta propia)	0.028 (0.029)	0.053 (0.039)	0.060 ** (0.029)	0.048 (0.033)
Controles	Si	Si	Si	Si
N	21294	9607	12045	13956
Adj. R-squared	0.24535	0.24555	0.24774	0.24247

Nota: ***p < .01; **p < .05; *p < .10. Errores estándar entre paréntesis agrupados a nivel barrio, MCJ refiere Metro Cable Línea J. Se omiten de la muestra los individuos residentes en los barrios San Javier N°1, San Javier N°2 y Las Independencias para evitar factores de confusión por el efecto spillover que genera la actividad turística en Las Independencias. *Fuente:* Cálculos propios

Conclusiones

La literatura ha revelado tres hechos empíricos relevantes respecto a la relación entre accesibilidad e informalidad hasta el momento: los trabajadores informales son más sensibles a costos de desplazamiento que sus contrapartes formales, los altos costos de desplazamiento inducen a trabajadores de la periferia a elegir trabajar en la informalidad cerca de su residencia y, mientras que la estrategia de subsidios al transporte parecen no funcionar bien los países en desarrollo, la expansión de infraestructura de tránsito reduce tasas de informalidad en áreas tratadas. Este trabajo investiga el efecto causal del acceso al transporte público sobre la informalidad laboral, con especial interés en verificar si existe un efecto diferencial por tipo de ocupación, cuenta propia y asalariados. Se realiza el análisis para la ciudad de Medellín, que se ha mostrado al mundo como una de las más innovadoras en transporte público, promoviendo sistemas de transporte tipo cable para reducir la segregación espacial de las zonas periféricas.

Se confirma, por un lado, (al menos en la estimación DD de estudio de eventos) que MCJ tiene efectos de reasignación, reduciendo significativamente la informalidad en zonas de influencia de 1000 y 1500 metros. Por otro, mediante una estrategia de identificación que combina control sintético con una triple diferencia, se evidencia que la infraestructura de transporte tiene efectos diferenciales por tipo de ocupación, la informalidad se reduce para trabajadores asalariados mientras que aumenta para los cuenta propia en comunas donde opera MCJ respecto a las comunas CS. En todas las estimaciones se observaron que ambos efectos difieren significativamente. Los resultados sugieren que la infraestructura de transporte puede ser aprovechada de manera distinta entre tipos de ocupación. Mientras que asalariados aprovechan la conexión que ofrece con el centro de la ciudad para transitar hacia empleos formales, los cuenta propia, al parecer, tienden a valorar más el empleo informal cercano aprovechando los flujos de pasajeros para aumentar sus ingresos en sus negocios. Asimismo, estos efectos van disminuyendo entre más grande sea la zona de influencia de la estación. Por último, se muestra que la estrategia metodológica con el enfoque de CS para construir contrafactuales objetivos es robusta ante las alternativas ad hoc.

En términos de política pública, este estudio puede mostrar dos importantes contribuciones. Primero, los resultados muestran evidencia en favor de que la expansión de transporte hacia zonas segregadas espacialmente puede ofrecer un mejor acceso a información sobre empleos formales, haciendo que los individuos valoren menos los empleos informales cercanos. Segundo, el hecho

de que se reduzca la informalidad laboral en asalariados y aumente en los cuenta propia es un resultado interesante porque plantean la discusión sobre tener en cuenta la heterogeneidad ocupacional de los trabajadores para entender mejor los beneficios del transporte público en el mercado laboral a través de la forma en como los trabajadores lo valoran.

No obstante, este estudio posee dos limitaciones relevantes: primero, las zonas de influencia circulares son una forma limitada de reflejar una área de servicio: primero, porque el área creada puede tomar un pedazo muy pequeño del barrio, con lo cual, al no tener la georreferenciación de los individuos obliga a incluir a todos los individuos de ese barrio en la muestra. Segundo, las personas generalmente no pueden caminar en línea recta en todas las direcciones desde la estación de transporte público. Si el barrio tiene poca integración en sus calles, con características geográficas o artificiales en el camino, como un río o una carretera sin paso de peatones, este puede tardar más tiempo en recorrer 500, 1000 o 1500 metros hasta la estación. Segundo, a pesar de que la metodología de CS propuesta por Robbins & Davenport (2021) con *microsynth* aborda algunos inconvenientes de CS con *Synth*, existen métodos como el de Ben-Michael et al. (2021) quienes realizan una crítica reciente a la restricción $w_j \geq 0$, argumentando que CS en ocasiones no ofrece un ajuste perfecto en pretratamiento, por lo que ofrecen un método llamado “Ridge Augmented Synthetic Control (Ridge ASC)”, que admite ponderaciones negativas, utilizando extrapolación para mejorar el ajuste del pretratamiento. El parámetro de regularización en Ridge ASC penaliza la distancia de los pesos del CS. Este método se puede implementar en R a través del paquete “*augsynth*”.

De acuerdo con lo anterior, se deben considerar las siguientes recomendaciones: primero, los nuevos rumbos de investigación deberían optar por herramientas más precisas para construir mejor zonas de influencia, tales como la herramienta “*Area de Servicio*” disponible ArcGIS. Las Áreas de Servicio, también conocidas como isócronas o polígonos de tiempo de conducción (o polígonos de distancia a pie, en este caso), muestran el área a la que se puede llegar desde una ubicación determinada dentro de un límite de tiempo o distancia. Si se calculan Áreas de Servicio se obtendrá una estimación mucho más precisa del área atendida por su sistema de tránsito. Segundo, hay varios enfoques de CS en la literatura además de los expuesto en este trabajo. Sería interesante emplear métodos estos métodos recientes que combinan simultáneamente CS con DD, tales como: Estimador Sintético de Diferencias en Diferencias (SDID en inglés) propuesto por Arkhangelsky et al., (2020) o Método de Control Sintético Generalizado: Inferencia Causal con

Modelos de Efectos Fijos Interactivos propuesto por Xu (2017). Por último, dado los resultados en este estudio, se podría analizar las ganancias en ingreso que obtienen los asalariados por aprovechar la conexión con empleos formales y los cuenta propia por aprovechar el flujo de pasajeros para sus negocios.

Referencias

- Abadie, A. (2020). Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*.
- Abadie, A., Diamond, A., & Hainmueller, A. J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's Tobacco control program. *Journal of the American Statistical Association*, 105(490), 493–505. <https://doi.org/10.1198/jasa.2009.ap08746>
- Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque country. *American Economic Review*, 93(1), 113–132. <https://doi.org/10.1257/000282803321455188>
- Abramo, P. (2009). Social innovation, reciprocity and the monetarization of territory in informal settlements in Latin American cities. *Social Innovation and Territorial Development*, 115–130.
- Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, 80(1), 123–129. [https://doi.org/10.1016/S0165-1765\(03\)00032-6](https://doi.org/10.1016/S0165-1765(03)00032-6)
- Akerlof, G. A. (1970). The Market for “Lemons”: Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3), 488–500.
- Alvarez-Rivadulla, M. J. (2009). Informalidades: efectos de la informalidad urbana sobre el empleo de los jóvenes. *Política y Gestión*, 11, 1–37.
- Amin, N. (2016). The Informal Sector and Urban Poor. *ResearchGate*.
- Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2020). *Synthetic Difference in Differences*. <http://arxiv.org/abs/1812.09970>
- Banerjee, S., & Goswami, K. (2019). Self-employed or Paid Employed: Who can Earn more among the Slum Dwellers and Why? *Progress in Development Studies*, 20(1), 7–25. <https://doi.org/10.1177/1464993419870961>
- Basant, R., & Sen, G. (2019). Impact of Affirmative Action in Higher Education for the Other Backward Classes in India. *Journal of Development Studies*, January. <https://doi.org/10.2139/ssrn.2814128>
- Bastiaanssen, J., Johnson, D., & Lucas, K. (2021). Does better job accessibility help people gain employment? The role of public transport in Great Britain. *Urban Studies*, 34–40. <https://doi.org/10.1177/00420980211012635>
- Bauhoff, S. (2014). The effect of school district nutrition policies on dietary intake and overweight: A synthetic control approach. *Economics and Human Biology*, 12(1), 45–55. <https://doi.org/10.1016/j.ehb.2013.06.001>
- Bea, D. (2016). Transport Engineering and Reduction in Crime: The Medellín Case. *Transportation Research Procedia*, 18(June), 88–92. <https://doi.org/10.1016/j.trpro.2016.12.012>
- Ben-Michael, E., Feller, A., & Rothstein, J. (2021). The Augmented Synthetic Control Method. *Journal of the American Statistical Association*, 116(536), 1789–1803. <https://doi.org/10.1080/01621459.2021.1929245>
- Bernal, R. S., & Peña, X. (2011). Guía práctica para la evaluación de impacto. In *Universidad de los Andes*.
- Bocarejo, J. P., & Oviedo, D. R. (2012). Transport accessibility and social inequities: a tool for identification of mobility needs and evaluation of transport investments. *Journal of Transport Geography*, 24, 142–154. <https://doi.org/10.1016/j.jtrangeo.2011.12.004>
- Boisjoly, G., Moreno-Monroy, A., & El-Geneidy, A. (2017). Informality and accessibility to jobs

- by public transit: Evidence from the São Paulo Metropolitan Region. *Journal of Transport Geography*, 64(August), 89–96. <https://doi.org/10.1016/j.jtrangeo.2017.08.005>
- Bonet, J., Perez, J., & Chiriví, E. (2016). *Labor Informality and Informal Settlements: A First Approach for Colombian Labor informality and informal settlements : A first approach for Colombian cities.*
- Bourguignon, F. (1979). Pobreza y dualismo en el sector urbano de las economías en desarrollo: El caso de Colombia. In *Revista Desarrollo y Sociedad* (Issue 1, pp. 39–72). <https://doi.org/10.13043/dys.1.2>
- Bradbury, B., & Chalmers, J. (2003). Housing , location and e mployment. *Australian Housing and Urban Research Institute*, 44.
- Brueckner, J. K., & Zenou, Y. (2003). Space and Unemployment: The Labor-Market Effects of Spatial Mismatch. *Journal of Labor Economics*, 21(1), 242–266. <https://doi.org/10.1086/344129>
- Cobbinah, P., Gaisie, E., & Owusu-Amponsah, L. (2015). Peri-urban morphology and indigenous livelihoods in Ghana. *Habitat International*, 50, 120–129. <https://doi.org/10.1016/j.habitatint.2015.08.002>
- Da Piedade, M., De Oliveira, B., & De Albuquerque, C. (2015). Residential segregation and social exclusion in brasilian housing market. *Institute for Applied Economic Reasearch*, 1813.
- Damm, A. (2014). Neighborhood quality and labor market outcomes: Evidence from quasi-random neighborhood assignment of immigrants. *Journal of Urban Economics*, 79, 139–166. <https://doi.org/10.1016/j.jue.2013.08.004>
- DANE. (2019). *Gran Encuesta Integrada de Hogares-GEIH.*
- De Soto, H. (1990). *The other path the invisible revolution in the third world* (Harper and, Vol. 22).
- Dujardin, C., Selod, H., & Thomas, I. (2008). Residential segregation and unemployment: The case of Brussels. *Urban Studies*, 45(1), 89–113. <https://doi.org/10.1177/0042098007085103>
- Elgin, C., & Oyvatt, C. (2013). Lurking in the cities: Urbanization and the informal economy. *Structural Change and Economic Dynamics*, 27, 36–47. <https://doi.org/10.1016/j.strueco.2013.06.003>
- Franklin, S. (2018). Location, Search Costs and Youth Unemployment: Experimental Evidence from Transport Subsidies. *Economic Journal*, 128(614), 2353–2379. <https://doi.org/10.1111/eoj.12509>
- Glaeser, E. L., Kahn, M. E., & Rappaport, J. (2008). Why do the poor live in cities? The role of public transportation. *Journal of Urban Economics*, 63(1), 1–24. <https://doi.org/10.1016/j.jue.2006.12.004>
- Gobillon, L., Selod, H., & Zenou, Y. (2007). Mechanism of spatial mismatch. *Urban Studies*, 44(12), 2401–2427.
- Gómez, Y., & Semeshenko, V. (2018). Transporte y calidad de vida urbana. Estudio de caso sobre el Metroplús de Medellín, Colombia. *Lecturas de Economía*, 89, 103–131. <https://doi.org/https://doi.org/g63z>
- Gonzales, G. (2000). Segregación socioespacial, integración los grupos de ingreso medio en la zona al mercado de trabajo y deterioro de conurbada Zacatecas-Guadalupe, México, 2000 Guadalupe. *Papeles de Población*, 46.
- Heilmann, K. (2014). *Can public transport investment relieve spatial mismatch?*
- Hernández, J. (2019). *El impacto de subsidios al uso del transporte público sobre el desempeño de los individuos en el mercado laboral : el caso de Bogotá.*

- Holzer, H. J., Quigley, J. M., & Raphael, S. (2003). Public Transit and the Spatial Distribution of Minority Employment: Evidence from a Natural Experiment. *Journal of Policy Analysis and Management*, 22(3), 415–441. <https://doi.org/10.1002/pam.10139>
- Huntington-Klein, N. (2014). The Effect: An Introduction to Research Design and Causality. In *Paper Knowledge . Toward a Media History of Documents* (CRC Press). Taylor & Francis Group.
- Iațu, C., Munteanu, A., Boghinciu, M., Cernescu, R., & Ibănescu, B. (2011). The effects of transportation system on the urban sprawl process for the city of Iasi, Romania. *WIT Transactions on the Built Environment*, 116, 291–301. <https://doi.org/10.2495/UT110251>
- Jaramillo, A. M., & Rengifo, J. C. (2018). Impacto del sistema Metroplús sobre el Mercado laboral de las comunas Manrique y Aranjuez de Medellín, Colombia. *Lecturas de Economía*, 89, 133–161. <https://doi.org/10.17533/udea.le.n89a05>
- Kain, J. (1968). Housing Segregation, Negro Employment, and Metropolitan Decentralization. *The Quarterly Journal of Economics*, 82(4827), 175–197.
- Khanna, G., Medina, C., Nyshadham, A., Ramos, D., Tamayo, J., & Tiew, A. (2020). *Spatial Mobility, Economic Opportunity, and Crime* (Vol. 02163).
- Loayza, N. V. (1996). The economics of the informal sector: a simple model and some empirical evidence from Latin America. *Carnegie-Rochester Conference Series on Public Policy*, 45, 129–162. [https://doi.org/10.1016/s0167-2231\(96\)00021-8](https://doi.org/10.1016/s0167-2231(96)00021-8)
- Martinez, D., Mitnik, O., Salgado, E., Scholl, L., & Yanez, P. (2018). Connecting to economic opportunity: The role of public transport in promoting women’s employment in Lima. *Inter-American Development Bank*.
- Martínez, L. M., & Viegas, J. M. (2009). Effects of transportation accessibility on residential property values: Hedonic price model in the Lisbon, Portugal, metropolitan area. *Transportation Research Record*, 2115, 127–137. <https://doi.org/10.3141/2115-16>
- Meagher, K. (2016). The Scramble for Africans: Demography, Globalisation and Africa’s Informal Labour Markets. *Journal of Development Studies*, 52(4), 483–497. <https://doi.org/10.1080/00220388.2015.1126253>
- Meeder, M., Aebi, T., & Weidmann, U. (2017). The influence of slope on walking activity and the pedestrian modal share. *Transportation Research Procedia*, 27, 141–147. <https://doi.org/10.1016/j.trpro.2017.12.095>
- Mitra, A. (1990). Employment Structure and Poverty Incidence : The Slum Perspective. *Indian Economic Review*, 25(1), 57–73.
- Moreno-Monroy, A. (2016). Access to public transport and labor informality. In *IZA World of Labor* (Issue July). <https://doi.org/10.15185/izawol.274>
- Moreno-Monroy, A., & Posada, H. (2018). The effect of commuting costs and transport subsidies on informality rates. *Journal of Development Economics*, 130(April 2015), 99–112. <https://doi.org/10.1016/j.jdeveco.2017.09.004>
- Moreno, A., & Roman, F. (2019). The impact of public transport expansions on informality: The case of the São Paulo Metropolitan Region. *Research in Transportation Economics*. <https://doi.org/10.1016/j.retrec.2020.100928>
- Muralidharan, K., & Prakash, N. (2017). Cycling to school: Increasing secondary school enrollment for girls in India. *American Economic Journal: Applied Economics*, 9(3), 321–350. <https://doi.org/10.1257/app.20160004>
- Offner, P. (1972). Labor Force Participation in the Ghetto. *The Journal of Human Resources*, 7(4), 460–481.

- Olden, A., & Moen, J. (2020). The Triple Difference Estimator. In *Department of Business and Management Science* (Issue April). <https://doi.org/10.2139/ssrn.3582447>
- Organizacion de Naciones Unidas. (2018). *Make cities and human settlements inclusive, safe, resilient and sustainable*.
- Padeiro, M., Louro, A., & Marques, N. (2019). Desarrollo y gentrificación orientados al tránsito: una revisión sistemática. *Transport Reviews*, 733–754.
- Phillips, D. C. (2014). Getting to work: Experimental evidence on job search and transportation costs. *Labour Economics*, 29, 72–82. <https://doi.org/10.1016/j.labeco.2014.07.005>
- Posada, H., & García-Suaza, A. (2021). *Transit Infrastructure and Informal Housing : Assessing an Expansion of the Medellin ' s Metrocable System* (Issue 267).
- Robbins, M. W., & Davenport, S. (2021). Microsynth: Synthetic control methods for disaggregated and micro-level data in R. *Journal of Statistical Software*, 97(2), 1–31. <https://doi.org/10.18637/jss.v097.i02>
- Rodríguez, C., Gallego, J. M., Martínez, D., Montoya, S., & Peralta-Quiros, T. (2016). Examining implementation and labor market outcomes of targeted transit subsidies: Subsidy by sistema nacional de selección de beneficiarios for urban poor in Bogotá, Colombia. *Transportation Research Record*, 2581, 9–17. <https://doi.org/10.3141/2581-02>
- Rotger, G. P., & Nielsen, T. S. (2015). Effects of job accessibility improved by public transport system: Natural experimental evidence from the copenhagen metro. *European Journal of Transport and Infrastructure Research*, 15(4), 419–441. <https://doi.org/10.18757/ejtir.2015.15.4.3090>
- Schumpeter, J. A. (1943). *Capitalism, Socialism & Democracy* (U. U. Books. (ed.)).
- Seki, M., & Yamada, E. (2020). *Heterogeneous Effects of Urban Public Transportation on Employment by Gender : Evidence from the Delhi Metro* (Issue 207).
- Sun, L., & Abraham, S. (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*. <https://doi.org/10.1016/j.jeconom.2020.09.006>
- The Guardian. (2018). *Dignity, community and hope in the Haitian slum of Jalousie – in pictures*.
- UN-Habitat. (2018a). *Metadata on SDGs Indicator 11.1.1 Indicator category: Tier I. March*, 1–12.
- UN-Habitat. (2018b). *Metrocable Medellín: Estudio de Caso*.
- Uribe, J., Ortiz, C., & Correa, J. (2006). ¿Cómo deciden los individuos en el mercado laboral Modelos y estimaciones para Colombia. *Lecturas de Economía, unknown*(64), 61–89.
- Warnes, P. (2020). *Columbia - Transport Infrastructure Improvements and Spatial Sorting : Evidence from Buenos Aires - Done*.
- Weinberg, B. A., Reagan, P. B., & Yankow, J. J. (2004). Do neighborhoods affect hours worked? Evidence from longitudinal data. *Journal of Labor Economics*, 22(4), 891–924. <https://doi.org/10.1086/423158>
- Xu, Y. (2017). Generalized synthetic control method: Causal inference with interactive fixed effects models. *Political Analysis*, 25(1), 57–76. <https://doi.org/10.1017/pan.2016.2>
- Zárate, R. (2020). *Spatial Misallocation, Informality, and Transit Improvements: Evidence from Mexico City*.

Apéndice

A. Aspectos formales del Control Sintético

$$\| \mathbf{X} - \mathbf{X}_0 \mathbf{W} \| = \left(\sum_{h=1}^k v_h (X_{h1} - w_2 X_{h2} - \dots - w_{j+1} X_{hj+1})^2 \right)^{\frac{1}{2}} \quad (9)$$

Las restricciones del Control Sintético de Robbins & Davenport (2021) son:

$$\sum_{j=1}^{J_0} w_j = J - J_0 \quad (10)$$

$$\sum_{j=1}^{J_0} w_j X_{kj} = \sum_{j=J_0+1}^J X_{kj} \quad \forall k \in (1, \dots, K) \quad (11)$$

$$\sum_{j=1}^{J_0} w_j Y_{lj} = \sum_{j=J_0+1}^J Y_{lj} \quad \forall l \in (1, \dots, L) \quad (12)$$

Anexos

1. Diferencia de media por cuenta propia y asalariados

Tabla 6.

Resumen estadísticas cuenta propia 2007 con zona de influencia 1000 metros

	Control Sintético	Tratamiento	P-Valor	Sig.
% Hombre	0.3762 (0.4846)	0.4047 (0.4920)	0.436	
% Hombre (w)	0.3681 (0.4742)	0.4047 (0.4920)	0.315	
% Casados	0.4261 (0.4947)	0.4605 (0.4996)	0.354	
% Casados (w)	0.4168 (0.4840)	0.4605 (0.4996)	0.238	
% Jefes	0.5757 (0.4945)	0.6140 (0.4880)	0.293	
% Jefes (w)	0.5642 (0.4847)	0.6140 (0.4880)	0.171	
Edad promedio	43.4786 (13.0547)	43.5674 (12.6326)	0.925	
Edad promedio (w)	42.5757 (12.7877)	43.5674 (12.6326)	0.293	
Educación promedio	8.2581 (4.7319)	7.1395 (4.2962)	0.001	**
Educación promedio (w)	8.0851 (4.6297)	7.1395 (4.2962)	0.004	**

Nota: w es el peso de las comunas de CS, Sig.: Significancia, desviación estándar entre paréntesis. Se omiten de la muestra los individuos residentes en los barrios San Javier N°1, San Javier N°2 y Las Independencias para evitar factores de confusión por el efecto spillover que genera la actividad turística en Las Independencias

Fuente: Cálculos propia con datos de ECV 2007 Medellín.

Tabla 7.

Resumen estadísticas asalariados 2007 con zona de influencia 1000 metros

	Control Sintético	Tratamiento	P-Valor	Sig.
% Hombre	0.4240 (0.4943)	0.3813 (0.4861)	0.050	**
% Hombre (w)	0.4157 (0.4847)	0.3813 (0.4861)	0.112	
% Casados	0.5258 (0.4994)	0.4795 (0.5000)	0.038	**
% Casados (w)	0.5155 (0.4897)	0.4795 (0.5000)	0.105	
% Jefes	0.4326 (0.4955)	0.5074 (0.5004)	0.001	**
% Jefes (w)	0.4244 (0.4862)	0.5074 (0.5004)	<0.001	**
Edad promedio	36.6519 (11.2791)	36.4975 (12.0693)	0.772	
Edad promedio (w)	35.9338 (11.0502)	36.4975 (12.0693)	0.288	
Educación promedio	9.5338 (4.6150)	8.0753 (4.3688)	<0.001	**

	Control Sintético	Tratamiento	P-Valor	Sig.
Educación promedio (w)	9.3470 (4.5225)	8.0753 (4.3688)	<0.001	**

Nota: **w** es el peso de las comunas de CS, Sig.: Significancia, desviación estándar entre paréntesis. Se omiten de la muestra los individuos residentes en los barrios San Javier N°1, San Javier N°2 y Las Independencias para evitar factores de confusión por el efecto spillover que genera la actividad turística en Las Independencias

Fuente: Cálculos propia con datos de ECV 2007 Medellín.

2. Estudio de eventos, regresión placebo y efecto causal con DD estandar del Metrocable línea J

Tabla 8.

Pruebas de tendencias paralelas para DD

	E(Informal = 1 X)			
	General (1)	500 (2)	1000 (3)	1500 (4)
Panel A: Estudio de eventos				
$I_{2005} * MCJ$	-0.0016 (0.0257)	-0.0660 ** (0.0322)	-0.0015 (0.0303)	-0.0159 (0.0290)
$I_{2006} * MCJ$	-0.0177 (0.0311)	-0.0750 * (0.0440)	-0.0531 (0.0335)	-0.0432 (0.0345)
$I_{2007} * MCJ$	0.0501 (0.0388)	0.0562 (0.0700)	0.0235 (0.0418)	0.0317 (0.0401)
$I_{2010} * MCJ$	-0.0134 (0.0267)	-0.0027 (0.0265)	-0.0025 (0.0274)	-0.0220 (0.0296)
$I_{2013} * MCJ$	-0.0420 (0.0291)	-0.0673 (0.0420)	-0.0620 ** (0.0284)	-0.0860 *** (0.0302)
$I_{2018} * MCJ$	-0.0429 (0.0306)	-0.0193 (0.0527)	-0.0196 (0.0371)	-0.0533 (0.0378)
Controles	Si	Si	Si	Si
N	42387	26316	29694	32471
Adj. R-squared	0.1499	0.1511	0.1510	0.1527
Panel B: Regresión placebo 2006-2007				
$I_{06-07} * MCJ$	0.0235 (0.0280)	0.0582 (0.0675)	-0.0089 (0.0374)	0.0162 (0.0363)
Controles	Si	Si	Si	Si
N	23081	14157	15999	17575
Adj. R-squared	0.1128	0.1237	0.1251	0.1216

Nota: ***p < .01; **p < .05; *p < .10. Errores estándar entre paréntesis agrupados a nivel barrio, MCJ refiere Metro Cable Línea J. Se omiten de la muestra los individuos residentes en los barrios San Javier N°1, San Javier N°2 y Las Independencias para evitar factores de confusión por el efecto spillover que genera la actividad turística en Las Independencias.

Fuente: Cálculos propia con datos de ECV Medellín.

Tabla 9.*Impacto causal estandar del Metrocable Linea J sobre la informalidad laboral*

	E(Informal = 1 X)			
	General (1)	500 (2)	1000 (3)	1500 (4)
Tratamiento	0.06176 ** (0.02568)	0.06628 ** (0.03261)	0.04732 * (0.02773)	0.06699 ** (0.02682)
Tiempo	-0.08426 *** (0.01486)	-0.08387 *** (0.01502)	-0.08376 *** (0.01496)	-0.08388 *** (0.01494)
Tiempo*MCJ	-0.04577 ** (0.02115)	-0.02114 (0.02564)	-0.0309 (0.01988)	-0.04035 ** (0.01911)
Constante	0.80903 *** (0.02524)	0.80748 *** (0.02595)	0.80865 *** (0.02569)	0.81308 *** (0.02553)
Controles	Si	Si	Si	Si
N	31345	19658	22096	24007
Adj. R-squared	0.12582	0.12317	0.12241	0.12534

Nota: ***p < .01; **p < .05; *p < .10. Errores estándar entre paréntesis agrupados a nivel barrio, MCJ refiere Metro Cable Línea J. Se omiten de la muestra los individuos residentes en los barrios San Javier N°1, San Javier N°2 y Las Independencias para evitar factores de confusión por el efecto spillover que genera la actividad turística en Las Independencias.

Fuente: Cálculos propia a partir de ECV