

Marzo 2022

WP4-2022-001

N° de serie

DOCUMENTO DE TRABAJO

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

Autores

Walter Jose Marin

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

Marín-Barrios, W. J¹

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 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 Control Sintético con una estimación de triple diferencia, se estima para toda la zona de tratamiento que la informalidad se reduce 5 p.p aproximadamente en asalariados, mientras que en los cuenta propia el efecto es nulo. El diferencial sugiere que el efecto tratamiento entre trabajadores cuenta propia y asalariados difiere significativamente en 5,3 p.p.

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

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

¹ Universidad de Antioquia, walter.marin1@udea.edu.co

Agradezco a Colombia Científica-Alianza Programa de Investigación EFI, 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. Los errores, opiniones y omisiones en este documento son míos y no comprometen a la alianza.

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

Marín-Barrios, W. J

–Introducción. –I. Revisión de la literatura. –II. Metodología y datos. –III. Resultados – Conclusiones. –Anexos. –Referencias

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.

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 vial². En Medellín, por ejemplo, según datos de la Encuesta origen destino 2017, viajar desde la zona pobre occidental hasta el centro de la ciudad tarda al menos 50 minutos, mientras que, viajar de la zona más rica 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 recorrer de 2 a 3 kilómetros, mientras que, viajar la misma distancia en la zona norte más rica toma solo 35 minutos (Bocarejo & Oviedo, 2012).

De esta manera, el transporte público es una herramienta esencial para reducir el desajuste espacial entre la periferia y las economías de aglomeración porque ofrece mecanismos para mejorar los 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 (Iașu et al., 2011), la pobreza (Glaeser et al., 2008), el crimen (Khanna et al., 2020), la calidad de la vivienda (Posada & García-Suaza, 2021) y en el mercado

² Este es el caso de Jalousie, Haití.

laboral sobre la 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 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 cuya proporción está creciendo debido a la escases de transporte público en zonas marginales, obreros de construcción, recolectores de basura y criadas. Por otro, Banerjee & Goswami (2019)³ comenta que las categorías ocupacionales más comunes, en especial en los 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 públicas pueden enfocarse y diseñarse de acuerdo con los determinados tipos de informalidad laboral.

Banerjee & Goswami (2019) contribuye a esta literatura y comenta en favor de Meagher (2016) que, dada la persistencia de la informalidad laboral en países en desarrollo, se deben hacer esfuerzos por considerar qué tipos de informalidad suelen ser más favorables para compensar la pobreza de los habitantes de tugurios. 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 a los trabajadores cuenta propia como impulsores del crecimiento económico (Schumpeter, 1943), la base del desarrollo (De Soto, 1990) o que los comerciantes pueden convertirse en los principales empresarios (Akerlof, 1970). De esta manera, las políticas públicas deben considerar, dada la persistencia de este fenómeno, generar las condiciones para que los empleos informales en distintos tipos puedan mejorar sus medios de vida

³ Barrios marginales de cuatro distritos de Bengala Occidental, India.

o que les permita transitar hacia empleos de mejor calidad. De ahí la importancia de comprender cuáles de estos empleos, 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 ocupaciones, asalariados⁴ y cuenta propia⁵. 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, pretendo investigar el efecto causal del sistema de transporte Metrocable Línea J sobre la informalidad laboral en Medellín, en particular, verificar si existe efectos heterogéneos 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 2007, 2008, 2013 y 2018 y una estrategia de identificación que combina control sintético 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 apoyo mi estimación con un estudio de eventos para ofrecer evidencia sobre los supuestos de tendencias paralelas y comportamiento anticipado (Sun & Abraham, 2020).

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) muestran 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) argumenta que los 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

⁴ **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).

⁵ **Independientes:** Los independientes trabajan para sí mismos y son dueños de sus propias empresas no constituidas en sociedad (excepto las cuasi sociedades). Lo conforman los cuenta propia, los patronos y los trabajadores familiares sin remuneración. Los cuenta propia son ocupados que dirigen su propia empresa o ejercen su profesión sin contratar a ningún trabajador asalariado (DANE).

por la 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 el Metrocable tiene efectos diferenciales en la informalidad según tipos de ocupación. En particular, para toda la zona de tratamiento, la informalidad se reduce 5 p.p aproximadamente en asalariados, mientras que en los cuenta propia el efecto es nulo. El diferencial sugiere que el efecto tratamiento entre trabajadores cuenta propia y asalariados difiere significativamente en 5,3 p.p. Este resultado se mantiene para áreas de influencia de 500, 1000 y 1500 metros con efectos decrecientes con la distancia. Este trabajo contribuye de varias maneras: primero, apoya la extensa literatura sobre impactos del transporte público y contribuye a la ampliación de los pocos estudios sobre sus efectos en la informalidad. Segundo, nuestro evidencia antigua del impacto en la informalidad en general con un enfoque metodológico diferente y al mismo tiempo, nueva evidencia analizando la heterogeneidad ocupacional. Así pues, el lector puede identificar que pretendo 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 presento esta introducción, luego la revisión de literatura sobre la 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, nuestro resultados, discusión y conclusiones.

I. 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 & Oyvatt (2013) y Bonet et al., (2016), cuyos objetivos son determinar el efecto de vivir en un asentamiento informal sobre la 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 las personas que viven en tierras ocupadas ilegalmente o sin tierras (asentamientos informales) y que no están afiliadas a la seguridad social (trabajo de informalidad). Encuentran por un lado que la informalidad laboral no es única de las zonas informales y que el hecho de vivir en asentamientos informales aumenta las posibilidades de que un joven tenga un empleo informal. Empero, otros estudios comentan la importancia de la relación causal entre ambos tipos de informalidad. Para el caso, Abramo (2009) y Abramo (2003) reconoce que el hecho de ser un trabajador informal aumenta la posibilidad de que estas personas vivan en un asentamiento

irregular. Estas evidencias sugieren que se debe tener cuidado al estudiar los efectos de los asentamientos sobre la informalidad laboral o viceversa, pues hay que tener en cuenta 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 los 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 el nivel de segregación del mercado de vivienda y los niveles de empleo para los afroamericanos en Detroit (Estados Unidos). Evidenció que el desplazamiento de la demanda laboral del centro de las ciudades a los suburbios (fenómeno de suburbanización del empleo), incentivó a los trabajadores a cambiar su residencia para estar cerca de los 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 las 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 la segregación espacial excluye las personas de oportunidades laborales, redes de información acerca de ellas y aumentan el costo de participar por la lejanía de sus hogares a los 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, sobretodo 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 los resultados laborales, la evidencia empírica ha recurrido usualmente a diseños cuasiexperimentales como la regresión discontinua (Hernández, 2019) y 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 la ampliación del sistema ferroviario Bay Area Rapid Transit (BART) de San Francisco, EE. UU sobre la contratación de minorías alrededor de dos nuevas estaciones. Muestran que la accesibilidad sólo importa ligeramente para los trabajadores latinos, pero no para los afroamericanos. Por su parte, Rotger & Nielsen (2015) usa el mismo enfoque metodológico con panel individual para evaluar el efecto de la accesibilidad a los puestos de trabajo urbanos en los ingresos individuales a través del sistema de transporte público Metro en Copenhague, Dinamarca. Evidencian un impacto positivo de una mejor accesibilidad en el 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 mejora 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 los subsidios al transporte, algunos trabajos analizan si los costos de transporte limitan la búsqueda de empleo en los mercados laborales urbanos de bajos salarios en Washington, DC (Phillips, 2014) y Etiopía (Franklin, 2018) a través de diseños experimentales. Phillips (2014) proporcionan subsidios 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 apoyan la literatura sobre búsqueda de empleo, evidenciando que los individuos previamente desempleados que recibieron un subsidio de transporte público aumentaron su intensidad de búsqueda de empleo en comparación con los individuos del grupo de control. Usando un diseño similar, Franklin (2018) también 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 los tratados aumentan la intensidad de la 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 los pobres para evaluar el vínculo causal entre el subsidio y los resultados del mercado laboral en cuanto a 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 la 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 y aquellos que ganan menos del mismo con el fin de captar la heterogeneidad del sector informal. Encuentran que la mayor accesibilidad al transporte reduce la probabilidad de formar parte del sector informal, sobretudo en el 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 los subsidios al transporte en las tasas de informalidad urbana. En este trabajo, los autores suponen que los trabajadores formales viajan todos los días para trabajar al Distrito Central de Negocios (CBD), mientras que los trabajadores informales tienen la opción de trabajar en su lugar de residencia. El estudio encuentra que un subsidio al transporte dirigido exclusivamente a los trabajadores informales puede no tener impacto en las tasas de informalidad, mientras que un subsidio dirigido a los 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 las primera aproximaciones a los efectos de las expansiones del transporte público sobre la informalidad laboral. Moreno & Roman (2019) comentan que una información detallada de las 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 la 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. Por otro lado, los nuevos rumbos de investigación deberían considerar analizar la relación abordada aquí con datos de panel y metodologías más novedosas como el Estimador Sintético de Diferencias en Diferencias (SDID en inglés) propuesto por (Arkhangelsky et al., 2020).

II. 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), este último se excluye por ser de la zona rural. 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 Quebra, Olaya entre otros.

En este contexto, pretendo explorar el impacto del Metrocable línea J (MCJ) en Medellín sobre la informalidad laboral definida normalmente en la literatura como aquel trabajador 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 la economía informal a la formal, bajo la premisa de que los trabajadores en lugares segregados prefieren trabajar en trabajos informales mal remunerados cerca de su residencia en lugar de incurrir en altos costos de transporte para acceder al 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 adelante 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 de 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, ya sea el 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 contrafactual de la diferencia entre la 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 $E[Y_i(0)|D_i = 1]$ (contrafactual). No obstante, este contrafactual es un resultado hipotético, por tanto, la aproximación ideal es definir una muestra dividida en un grupo de tratamiento definido por trabajadores residentes en las comunas que recibieron la nueva infraestructura y grupo de control, que no fueron tratados pero que guarda cierta similitud con el grupo de tratamiento.

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

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

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

Intuitivamente, la ecuación (2) establece el análisis contrafactual entre grupos de trabajadores cuenta propia ($B = 1$) y asalariados ($B = 0$). La idea es comparar los cuenta propia en la comuna de tratamiento con los cuenta propia en la comuna de control. Según Olden & Moen (2020), esto no sería válido si las 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 sintético, independientemente de la Línea J. Sin embargo, se puede suponer razonablemente que las diferencias económicas generales no afectarían a los 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 las comunas de tratamiento en ausencia de éste (En la sección C se amplía esta discusión a los supuestos de tendencias paralelas y comportamiento anticipado).

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

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

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 los cuenta propia, mientras que la segunda sería el efecto para los asalariados.

Según lo anterior, la especificación vía regresión del estimador de triple diferencia está dada 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 para la zona de intervención. Esta se define como el resultado que hubiese obtenido la zona de tratamiento si no hubiese existido la intervención. Comúnmente, la literatura sobre los impactos del sistema de transporte utiliza diferencias en diferencias que suelen utilizar zonas de control ad hoc 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, recurro 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 una zona de control “sintética” que representa una combinación ponderada de muchos 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 controles, 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 la zona impactada por la intervención o política de interés que para este caso son las comunas San Javier y Robledo. El conjunto de comparaciones potenciales llamados “conjunto de donantes,” $j = 2, \dots, J + 1$ es un conjunto de zonas no afectadas por la intervención, que serían el resto de comunas de la ciudad. Se suponen T períodos (2007, 2008, 2013 y 2018), donde los primeros T_0 períodos son anteriores a la intervención (2007). Para cada zona j y tiempo t , se observa el resultado de interés Y_{jt} y un conjunto de k predictores del resultado, X_{1j}, \dots, X_{kj} . Asimismo, para cada zona j y t se define Y_{jt}^{NI} como la respuesta potencial sin intervención (contrafactual) y para la zona intervenida $j=1$ post-tratamiento $t > T_0$ se define Y_{1t}^I como la respuesta potencial con intervención. De este modo, el efecto del tratamiento para la zona de intervención vendría dado por $\tau_{1t} = Y_{1t}^I - Y_{1t}^{NI}$. Es importante notar que este 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, $\mathbf{W} = (w_2, \dots, w_{j+1})'$. 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 los valores de la variable de resultado antes de la intervención para la zona tratada. Esto es, dado un conjunto de constantes no negativas,

v_1, \dots, v_k que reflejan la 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 el contrafactual y la contribución de cada zona es representada por ponderación de control sintético. La ecuación (9) es una medida de distancia que indica la reducción de la distancia entre los valores de una determinada covariable de la zona tratada y el resto de valores que toma esa covariable en el resto de las zonas y se suma a través de las unidades no tratadas.

No obstante, este enfoque se ha aplicado en el 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. La escasez de casos de tratamiento y comparación 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. Además, la implementación de esta metodología a través del paquete *Synth* en los software estadísticos está limitada a datos de panel balanceados y la ECV usada aquí es una encuesta anual de corte transversal repetido, es decir, encuestan los mismos temas a diferentes personas y hogares en cada año. Para el propósito de este estudio, se utiliza la metodología propuesta por Robbins & Davenport (2021) en el paquete *mycosynth* que abarca las limitaciones del enfoque agregado y la implementación con *Synth*. Este enfoque, en principio, es utilizado para datos de panel con un nivel de observación granular (manzanas) pero admite datos de sección transversal repetido, sección transversal ($T = 1$) 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 y los últimos $J - J_0$ son tratados. Además, hay L resultados variables en el tiempo medidos a lo largo de T períodos de tiempo totales con T_0 periodos previos a la intervención y los últimos $T - T_0$ son posteriores a la intervención. Y_{ijt} se utiliza para representar el valor del resultado del individuo i en el momento t para la zona j . También, hay k covariables totales invariantes en el tiempo (es decir, en la línea de base) $X_{1ijt}, \dots, X_{kijt}$ para cada i en el momento t y la zona j . Se calcula un grupo de control sintético asignando un peso a cada caso no tratado. Estas ponderaciones se denominan w_j para $j \in (1, \dots, J_0)$. El algoritmo *mycosynth* hace coincidir el tratamiento y el control sintético calculando pesos que satisfacen tres clases de

restricciones. En primer lugar, la suma de los pesos es igual al número de casos en el área de tratamiento. Esta restricción tiene la misma intuición que la restricción de pesos de Synth. En segundo lugar, la ponderación del control sintético 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, puedo 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, más bien, una vez se generen los pesos para las comunas que sumen las unidades tratadas, se ponderan las observaciones de estas zonas de control sintético por sus respectivos pesos y así equilibrar las características de los individuos en estas zonas con las de tratamiento. Esta estrategia ha sido utilizada por Bauhoff (2014) para evaluar los efectos a corto plazo de la política de nutrición escolar de 2004 para Los Ángeles con un enfoque de control sintético usando datos de corte transversal repetido. El autor, luego de aplicar el control sintético para el periodo antes de la intervención, toma los pesos y pondera las 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:

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

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, la cual toma el valor de 1 si el trabajador pertenece al grupo de tratamiento. I_t es 1 si la observación corresponde al periodo posterior a la intervención. D_i es la dummy que captura el efecto diferencial entre los tipos de empleo, que es igual a uno si el individuo i es cuenta propia. El término $I_t T_i D_i$ es la triple

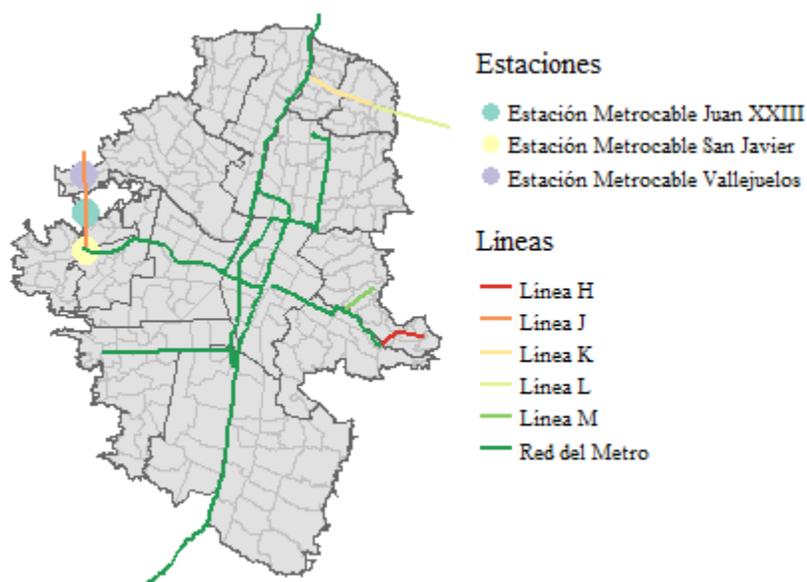
interacción entre la variable de tratamiento, la dummy del periodo en que se encuentra la observación y la dummy de heterogeneidad ocupacional. b es indicador de distancia a la estación que se usa en la estimación. Los coeficientes de interés son α_4 y δ . La estimación DD usual del efecto del tratamiento para los asalariados es α_4 y la estimación DD del efecto del tratamiento para los cuenta propia sería $(\alpha_4 + \delta)$. 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 los trabajadores que ayudan a controlar los efectos del tratamiento y estarán ponderadas con w_j para las comunas de control sintético. Para estimar los parámetros uso estimaciones MCO para cada variable de resultado. En particular, para la variable de respuesta que indica si el individuo es formal o informal utilizo un Modelo Lineal de Probabilidad (MLP) corrigiendo por heterocedasticidad. Debido a que mi interés es el impacto causal y no la predicción, no es necesario usar modelos probit o logit.

Ahora bien, cuento con 4 periodos de observación, un año antes de la puesta en marcha del MCJ que sería el periodo de pre-tratamiento (2007) y 2 años posteriores a la puesta en marcha (2013 y 2018), teniendo como año de tratamiento 2008 cuando se inauguró la línea. Estimo la ecuación (7) a nivel general para las comunas tratadas y control sintético y para diferentes buffer. Defino estas áreas usando círculos concéntricos alrededor de las estaciones de la línea J donde el área de influencia son los círculos más interiores de 500, 1000 y 1500 metros. Fijo un área de influencia de hasta 1500 metros porque es la distancia máxima usual que una persona camina para llegar a una estación de transporte público (Martinez et al., 2018).

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 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 a partir de POT 2014

III. Resultados

A. Control sintético

La Tabla 1 resume los resultados de haber aplicado la metodología de control sintético de (Robbins & Davenport, 2021) . Las primeras tres columnas muestran la correspondencia del tratamiento con el control sintético en las covariables (sexo, estado civil, parentesco y tipo de empleo) y variables de resultado (informalidad y gasto del hogar en logaritmo), aprovechando la ventaja que ofrece este enfoque de incluir más de una variable de resultado para que haya mayor precisión en el ajuste. Se confirma que se cumplen todas las restricciones descrita en II, ya que las 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 la primera fila (Intercepto) indica que hay 2 comunas de tratamiento y que las 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 B se mostrará estas correspondencias a través de un ejercicio de diferencia de medias.

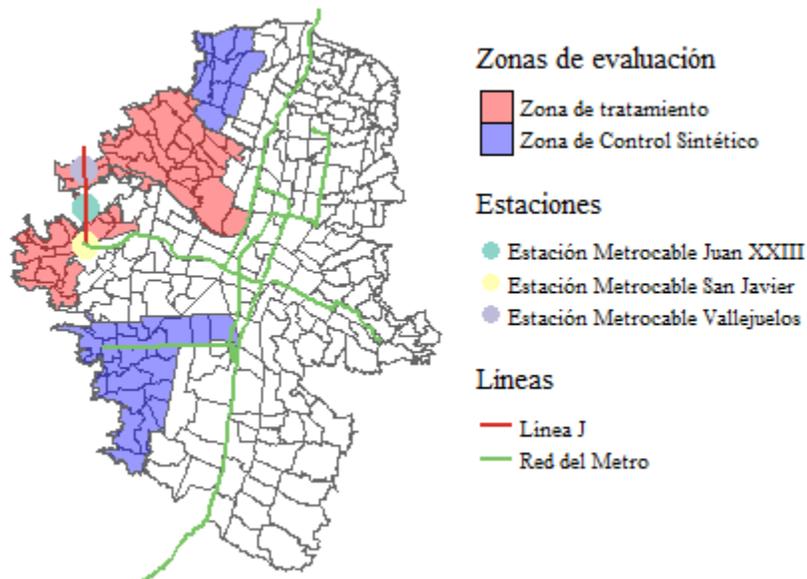
En el panel derecho están los cálculos de los pesos para cada una de las comunas de control potenciales, mostrando que las comunas Doce de Octubre (0.9716770), Poblado (0.0552449) y Belén (0.9730780) conforman las unidades de control sintético que se usaran como contrafactual para la estimaciones de triple diferencia. Empero, se excluye la comuna Poblado, pues es una comuna con bajas tasas de precariedad laboral en la ciudad y su peso es muy bajo. La Figura 2 muestra las zonas de tratamiento y control sintético.

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

Fuente: Elaboración propia

Figura 2. Definición de zonas de tratamiento y Control Sintético



Fuente: Elaboración propia con POT 2014

B. Estadísticas descriptivas

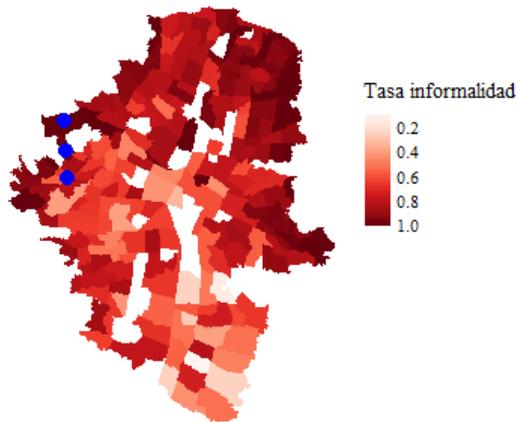
La Figura 3 muestra la distribución espacial de la tasa de informalidad (TI) por cuenta propia 3(a) y asalariados 3(b) a nivel de barrio. A pesar de la informalidad es mayor en los cuenta propia que, en los asalariados, se observa un patrón espacial en ambos grupos, sus tasas de informalidad se concentran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

La Tabla 2 presenta un conjunto de características socioeconómicas de los trabajadores para el año pre-tratamiento 2007 en un buffer de 1500 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 los 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.

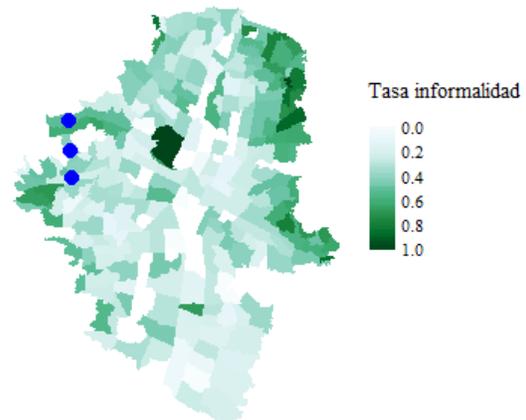
Los datos revelan que las diferencias entre las comunas tratadas y control sintético en pretratamiento se reducen más con la ponderación, lo que sugiere que la restricción dos tiende a cumplirse (ver ecuación (11) en apéndice A). No obstante, hay diferencias en la educación promedio y porcentaje de jefes que se pueden controlar en la estimaciones. Este mismo ejercicio se realizó por cuenta propia y asalariados con resultados similares (ver Tabla 5 y Tabla 6 en anexo 1).

Figura 3. *Distribución de informalidad cuenta propia y asalariados*

(a) *Distribución informalidad cuenta propia*



(b) *Distribución informalidad*



Fuente: Elaboración propia a partir de ECV 2008

Tabla 2. Resumen estadísticas a nivel individuo 2007 con buffer 1500

	Control Sintético	Tratamiento	Valor P	Sig.
% Hombre	0.3874 (0.4875)	0.4104 (0.4920)	0.217	
% Hombre (w)	0.3874 (0.4875)	0.4022 (0.4822)	0.426	
% Casados	0.4746 (0.4997)	0.4975 (0.5001)	0.230	
% Casados (w)	0.4746 (0.4997)	0.4875 (0.4901)	0.497	
% Jefes	0.5351 (0.4991)	0.4732 (0.4993)	0.001	**
% Jefes (w)	0.5351 (0.4991)	0.4641 (0.4898)	<0.001	**
% Cuenta propia	0.2603 (0.4391)	0.2836 (0.4508)	0.166	
% Cuenta propia (w)	0.2603 (0.4391)	0.2777 (0.4415)	0.298	
Edad promedio	38.3378 (12.5990)	38.5881 (12.2026)	0.601	
Edad promedio (w)	38.3378 (12.5990)	37.8176 (11.9491)	0.276	
Educación promedio	7.8317 (4.3668)	9.1720 (4.6833)	<0.001	**
Educación promedio (w)	7.8317 (4.3668)	8.9891 (4.5880)	<0.001	**

Nota: **w** es el peso de las comunas de CS, Sig.: Significancia

Fuente: Elaboración propia ECV 2007.

C. Estudio de eventos

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 sugiere que no haya conocimiento de la futura intervención para que no puedan cambiar su comportamiento antes del tratamiento y, por tanto, el resultado potencial antes del tratamiento pueda representar los resultados de referencia (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. El diseño de estudio de eventos puede ofrecer cierta evidencia para la identificación de estos supuestos y apoya la validez interna de las estimaciones de interés. Explotando los datos previos al tratamiento, la especificación de estudio de eventos para una estrategia DDD sería:

$$\begin{aligned}
Y_{ijt} = & c + \sum_{\tau=2005}^{2018} \gamma_{\tau} I_{i\tau} + \alpha T_i + \pi D_i + \sum_{\tau=2005}^{2018} \theta_{\tau} I_{i\tau} * T_i + \sum_{\tau=2005}^{2018} \alpha_{\tau} I_{i\tau} * D_i + \lambda T_i * D_i \\
& + \sum_{\tau=2005}^{2018} \delta_{\tau} I_{i\tau} T_i D_i + \sum_{i=1}^p \beta' w_j X_i + \varepsilon_{it}
\end{aligned} \tag{8}$$

Todo se define igual que (7). Para la ecuación (8) interesa los coeficientes θ_{τ} y δ_{τ} con $\tau \in \{2005, 2008, 2013, 2018\}$ teniendo como categoría base el año inmediatamente anterior a la intervención, 2007. Si θ_{2005} , θ_{2008} , $(\theta_{2005} + \delta_{2005})$ y $(\theta_{2008} + \delta_{2008})$ son estadísticamente iguales a cero, sugiere que la hipótesis de tendencia paralela se mantiene para asalariados y cuenta propia respectivamente. Por su parte, si δ_{2005} y δ_{2008} resultan no significativos, sugiere que la hipótesis se mantiene para el diferencial de ambos grupos. Por último, los coeficientes para $\tau \geq 2013$ muestra una estimación del impacto causal de interés. Además, si tendencias paralelas se cumple, puede suponerse que no hubo comportamiento anticipatorio. De esta manera, puedo ofrecer evidencia del cumplimiento de los supuestos antes mencionados.

La Tabla 3 muestra los resultados del estudio de eventos con y sin controles (Tabla 2). En aras del espacio, las primeras cuatro filas muestran la estimación del efecto tratamiento para los asalariados $\hat{\theta}_{\tau}$, mientras que las siguientes cuatro filas proveen el efecto diferencial de la Línea J para cuenta propia comparado con el efecto para asalariados $\hat{\delta}_{\tau}$ que es la triple interacción en (8). Los resultados sugieren que la hipótesis de tendencias paralelas se mantiene tanto para trabajadores cuenta propia como para asalariados. Asimismo, se mantiene la hipótesis para el diferencial entre esos dos efectos. Lo anterior apoya la validez interna de las estimaciones en D. Además, los

estimadores post-tratamiento resultaron significativos lo cual es una primera evidencia del impacto heterogéneo de la Línea J sobre la informalidad, mostrando también que el efecto se reduce en el tiempo. La Tabla 7 del anexo 2 muestra el estudio de eventos de diferencias en diferencias estándar con evidencia consistente de esta hipótesis.

Tabla 3. Estudio de eventos sobre la informalidad laboral

	E(Informal = 1 X)	
	DDD informal	DDD informal
$I_{2005} * T$ (asalariados)	-0.029 (0.020)	-0.026 (0.021)
$I_{2008} * T$ (asalariados)	-0.008 (0.021)	-0.003 (0.022)
$I_{2013} * T$ (asalariados)	-0.119 *** (0.025)	-0.108 *** (0.026)
$I_{2018} * T$ (asalariados)	-0.066 *** (0.022)	-0.082 *** (0.023)
$I_{2005} * T * D$	-0.044 (0.032)	-0.036 (0.033)
$I_{2008} * T * D$	0.015 (0.034)	0.046 (0.035)
$I_{2013} * T * D$	0.151 *** (0.044)	0.184 *** (0.045)
$I_{2018} * T * D$	0.025 (0.050)	0.048 (0.051)
$I_{2005} * T$ (cuenta propia)	-0.001 (0.024)	0.005 (0.025)
$I_{2007} * T$ (cuenta propia)	0.007 (0.027)	0.043 (0.027)
$I_{2013} * T$ (cuenta propia)	0.032 (0.036)	0.076 ** (0.037)
$I_{2018} * T$ (cuenta propia)	-0.041 (0.044)	-0.034 (0.046)
Controles	Si	No
N	24124	24124
R-squared	0.281	0.205
Adj. R-squared	0.280	0.204
Residual Std. Error	0.424	0.446
F Statistic	392.206 ***	327.037 ***

Nota: ***p < .01; **p < .05; *p < .1. DDD se refiere a triple diferencia
Fuente: Elaboración propia.

D. Efectos causal heterogéneo del Metrocable línea J

Antes de pasar al efecto heterogéneo, muestro evidencia del enfoque DD estándar (sin diferenciar por tipo de empleo) en la Tabla 8 del anexo 2, siendo la columna 1 la estimación para toda la zona de tratamiento y las columnas siguientes (2 a 4) la estimación para los diferentes buffers. La columna 1 no verifica que la provisión del Metrocable Línea J reduzca la informalidad laboral en las zonas de tratamiento respecto al grupo de control sintético. Sin embargo, las columnas 2 a 4 muestran que el efecto no solo se reduce con la distancia a la estación, también se pierde, revelando que el impacto depende de la intensidad del tratamiento, cuanto más cerca se está de la estación, mayor es la reducción en la informalidad laboral. En particular, el resultado indica que, para los trabajadores en un radio de influencia de 500 metros, se reduce la probabilidad de trabajar en un empleo informal en 10 p.p, para aquellos a 1000 metros el efecto se reduce a casi la mitad (4,4 p.p) y a 1500 metros el efecto es nulo.

Dado que la proporción de informales en 2007 fue de 52 % aproximadamente, implica que en una zona a 500 metros de la estación se reduciría la informalidad laboral a 42 % luego de la entrada de la Línea, para la zona al doble de distancia, se reduciría a 45,6 % postratamiento y en una zona al triple de distancia no habría cambio. Lo anterior 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). Además, sugiere que la estrategia de CS para conseguir contrafactuales objetivos es robusta ante las alternativas ad hoc.

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 evidencia el impacto causal heterogéneo por tipo de empleo. 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 entre los efectos de cuenta propia y asalariado ($\hat{\delta}$).

La estimación informa que, mientras para los cuenta propia no hay resultados estadísticamente significativos, la probabilidad de ser informal se reduce en 5 p.p aproximadamente para trabajadores asalariados en las comunas de intervención. No obstante, el efecto diferencial según tipo de ocupación indica que los cuenta propia son 5,3 p.p más propensos a ocuparse

informalmente que los asalariados en las comunas tratadas. Por su parte, la columna 2 a 4 muestra que el efecto tanto para asalariados como para el diferencial es decreciente con la distancia a la Línea. Para los cuenta propia el impacto sigue siendo nulo con la distancia. Específicamente, se observa que son decrecientes en el orden de 18 p.p (500 metros), 8 p.p (1000 metros) y 4,5 p.p (1500 metros) para asalariados y, de 11 p.p (500 metros) y 6,5 p.p (1500 metros) para el diferencial.

De acuerdo con lo anterior, la reducción de la informalidad observada en la Tabla 8 parece estar impulsada por el empleo asalariado, cuyo efecto dado en la columna 1 implicaría una reducción de 40,5% (antes de la intervención) a 35,5% luego de la puesta en marcha de la Línea J. Los efectos en la columna 2 a 4 implican que al estar a 500 metros de la estación se reduciría la informalidad en asalariados hasta 23% respecto a su tasa de informalidad anterior al tratamiento, mientras que estar a 1500 metros solo la reduciría hasta 32,5%. Respecto al diferencial, solo sugiere que la diferencia entre los efectos de ambos grupos se hace cada vez más pequeña con la distancia. Esto va en línea con estudios que también 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 con diferencias de 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 mejor las mejoras en el transporte público. Según datos de la ECV para 2008⁷, es posible que este resultado este explicado por el hecho de que, 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 ratifica que los cuenta propia tienden a valorar los empleos informales cercanos y los asalariados aprovechan la conexión con las zonas centrales de empleo para transitar hacia la formalidad.

⁷ 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

	DDD informal	E(Informal = 1 X)		
		DDD 500	DDD 1000	DDD 1500
Tratamiento	0.069 *** (0.016)	0.207 *** (0.036)	0.103 *** (0.026)	0.106 *** (0.021)
Tiempo	-0.027 ** (0.012)	-0.038 *** (0.009)	-0.040 *** (0.010)	-0.044 *** (0.010)
Tipo empleo	0.457 *** (0.016)	0.452 *** (0.013)	0.448 *** (0.014)	0.458 *** (0.014)
Tiempo x Tratamiento (asalariados)	-0.04 *** (0.018)	-0.176 *** (0.039)	-0.077 *** (0.029)	-0.045 * (0.024)
Tratamiento x Tipo empleo	-0.027 (0.026)	-0.116 ** (0.054)	-0.029 (0.040)	-0.083 ** (0.034)
Tiempo x Tipo empleo	-0.006 (0.020)	0.011 (0.016)	0.012 (0.016)	0.009 (0.017)
Tiempo x Tratamiento x Tipo empleo	0.053 * (0.031)	0.110 * (0.061)	0.042 (0.045)	0.065 * (0.039)
Constante	0.676 *** (0.018)	0.697 *** (0.017)	0.696 *** (0.018)	0.684 *** (0.018)
Tiempo x Tratamiento (cuenta propia)	0.003 (0.025)	-0.067 (0.046)	-0.035 (0.035)	0.02 (0.031)
Controles	Si	Si	Si	Si
N	16440	16440	16440	16440
R-squared	0.272	0.273	0.272	0.274
Adj. R-squared	0.272	0.273	0.272	0.273
Residual Std. Error	0.425	0.424	0.425	0.424
F Statistic	512.728 ***	514.194 ***	512.115 ***	516.417 ***

Nota: ***p < .01; **p < .05; *p < .1

Fuente: Cálculos propios. DDD se refiere a triple diferencia

Conclusiones

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.

La literatura general ha revelado tres hechos empíricos relevantes respecto a la relación entre la accesibilidad y la informalidad hasta el momento: los trabajadores informales son más sensibles a los costos de desplazamiento que sus contrapartes formales, los altos costos de desplazamiento inducen a los trabajadores de las periferias a elegir trabajar en la informalidad cerca

de su residencia y la infraestructura de tránsito reduce las tasas de informalidad en las áreas tratadas. Mediante una estrategia de identificación que combina control sintético con un enfoque de triple diferencia, confirmo por un lado que, la probabilidad de trabajar en un empleo informal se reduce significativamente en las comunas intervenidas (San Javier y Robledo). Por otro, muestro que la infraestructura de transporte tiene efectos diferenciales por tipo de ocupación, la informalidad se reduce para los trabajadores asalariados mientras que es nulo para los cuenta propia, con diferencial significativo entre ambos impactos que sugiere que los asalariados aprovechan mejor la infraestructura de transporte y los cuenta propia, al parecer tienden a valorar más el empleo informal cercano. Asimismo, estos efectos van disminuyendo con la distancia a la estación. Por último, muestro 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 tres importantes contribuciones. Primero, los resultados muestran evidencia en favor de que las políticas de transporte en zonas segregadas espacialmente pueden 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 no en los cuenta propia, sugiere que se podría complementar la política de transporte con políticas de vocación laboral que fomente la transición hacia empleos asalariados. Finalmente, según hallazgos revisados, las políticas de subsidios al transporte parecen no funcionar bien en los países en desarrollo y es un hecho que los efectos del transporte en el bienestar dependen en gran medida de la cercanía al transporte, por lo tanto, en este contexto parece haber una mayor ganancia si se priorizan proyectos de transporte en las zonas pobres y desconectadas del centro de la ciudad.

Por último, los nuevos rumbos de investigación deberían considerar analizar la relación abordada aquí con datos de panel y metodologías más novedosas como el Estimador Sintético de diferencias en diferencias (SDID en inglés) propuesto por (Arkhangelsky et al., 2020).

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/10.1257/00028280321455188>
- Abramo, P. (2003). La teoría económica de la favela: cuatro notas sobre la localización residencial de los pobres y el mercado inmobiliario informal. *Ciudad y Territorio: Estudios Territoriales*, 136, 273–294.
- 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.
- 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>
- 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>
- 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>
- 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., & Oyvat, 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/10.17533/udea.le.n89a04>
- 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>
- 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>
- 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>
- 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>
- 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 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>
- 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>

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:

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

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

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

Anexo

1. Diferencia de media por cuenta propia y asalariados

Tabla 5: Resumen estadísticas cuenta propia 2007 con buffer 1500

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

Nota: w es el peso de las comunas de CS, Sig.: Significancia
Fuente: Elaboración propia ECV 2007.

Tabla 6. Resumen estadísticas asalariados 2007 con buffer 1500

	Control Sintético	Tratamiento	P-Valor	Sig.
% Hombre	0.3813 (0.4861)	0.4240 (0.4943)	0.050	**
% Hombre (w)	0.3813 (0.4861)	0.4157 (0.4847)	0.112	
% Casados	0.4795 (0.5000)	0.5258 (0.4994)	0.038	**
% Casados (w)	0.4795 (0.5000)	0.5155 (0.4897)	0.105	
% Jefes	0.5074 (0.5004)	0.4326 (0.4955)	0.001	**
% Jefes (w)	0.5074 (0.5004)	0.4244 (0.4862)	<0.001	**
Edad promedio	36.4975 (12.0693)	36.6519 (11.2791)	0.772	
Edad promedio (w)	36.4975 (12.0693)	35.9338 (11.0502)	0.288	
Educación promedio	8.0753 (4.3688)	9.5338 (4.6150)	<0.001	**
Educación promedio (w)	8.0753 (4.3688)	9.3470 (4.5225)	<0.001	**

Nota: w es el peso de las comunas de CS, Sig.: Significancia
Fuente: Elaboración propia ECV 2007.

2. Estudio de eventos y efecto causal con DD estandar del Metrocable linea J

Tabla 7. Estudio de eventos sobre la informalidad laboral

	E(Informal = 1 X)	
	DDD informal	DDD informal
$I_{2005} * T$	-0.021 (0.017)	-0.018 (0.018)
$I_{2008} * T$	0.001 (0.018)	0.014 (0.019)
$I_{2013} * T$	-0.050 ** (0.023)	-0.027 (0.024)
$I_{2018} * T$	-0.048 ** (0.021)	-0.063 *** (0.022)
Controles	Si	No
N	24124	24124
R-squared	0.160	0.067
Adj. R-squared	0.159	0.067
Residual Std. Error	0.458	0.483
F Statistic	327.616 ***	192.657 ***

Nota: ***p < .01; **p < .05; *p < .1.

DD se refiere a Diferencias en Diferencias

Fuente: Cálculos propios

Tabla 8. Impacto causal estandar del Metrocable Linea J sobre la informalidad laboral

	E(Informal = 1 X)			
	DD informal	DD 500	DD 1000	DD 1500
Tratamiento	0.039 *** (0.014)	0.135 *** (0.031)	0.069 *** (0.022)	0.065 *** (0.018)
Tiempo	-0.030 *** (0.011)	-0.037 *** (0.008)	-0.037 *** (0.009)	-0.041 *** (0.009)
Tiempo x Tratamiento	-0.026 (0.016)	-0.101 *** (0.034)	-0.044 * (0.025)	-0.020 (0.021)
Constante	0.728 *** (0.019)	0.739 *** (0.018)	0.738 *** (0.019)	0.730 *** (0.019)
Controles	Si	Si	Si	Si
N	16440	16440	16440	16440
R-squared	0.116	0.116	0.116	0.117
Adj. R-squared	0.115	0.116	0.115	0.116
Residual Std. Error	0.468	0.468	0.468	0.468
F Statistic	268.343 ***	270.441 ***	268.928 ***	271.644 ***

Nota: ***p < .01; **p < .05; *p < .1. DD se refiere a Diferencias en Diferencias

Fuente: Elaboración propia

