The Journal of Transport and Land Use http://jtlu.org
Vol. 15 No. 1 [2022] pp. 613–634

Subway expansion, job accessibility improvements, and home value appreciation in four global cities: Considering both local and network effects

Adriano Borges Costa (corresponding author)
Arq.futuro Cities Lab
Insper Institute of Education and Research
AdrianoBFC@insper.edu.br

Camila Ramos
MIT Center for Real Estate
Massachusetts Institute of Technology
cramosy@mit.edu

Siqi Zheng
MIT Center for Real Estate
Massachusetts Institute of Technology
sqzheng @mit.edu

Abstract: We explore the potential of incorporating accessibility analysis in addressing the impact of subway expansions on the real estate market. We first demonstrate that by using increases in accessibility to jobs as a continuous treatment variable, rather than adopting a binary station dummy approach, we achieve better goodness-of-fit in a quasi-experimental econometric analysis. Furthermore, accessibility measures allow the exploration of impacts beyond the local effects around new subway stations, shedding light on a network impact that has been largely overlooked to date. To increase the external validity of our findings, we apply the same analysis to the cities of Santiago (Chile), Sao Paulo (Brazil), Singapore, and Barcelona (Spain). and then explore the emergent patterns. We argue that the integration of urban economics and transportation analysis via the use of accessibility measures constitutes an innovation in the empirical approach commonly adopted in the literature. The use of such measures in causal empirical studies on transportation impacts can yield more robust and comprehensive results and capture nuanced spatial heterogeneity effects.

Article history:
Received: December 14, 2021
Received in revised form: May 3, 2022
Accepted: July 1, 2022
Available online: October 5, 2022

1 Introduction

The impacts of transportation infrastructure investment upon real estate prices are commonly reported within the literature, with dozens of empirical studies reporting a positive effect of transit infrastructure extensions in both the short- and long-term. Though the magnitude of this appreciation varies significantly, depending on the kind of transportation technology implemented, meta-analyses consolidating the empirical literature show positive elasticities for the overall potential impact on real estate values (Debrezion et al., 2007; Mohammad et al., 2013). Methodologically, most of the studies examining transit impacts on real estate prices use either hedonic models in cross-sectional analysis (e.g., Benjamin

http://dx.doi.org/10.5198/jtlu.2022.2146

The *Journal of Transport and Land Use* is the official journal of the World Society for Transport and Land Use (WSTLUR) and is published and sponsored by the University of Minnesota Center for Transportation Studies. This paper is also published with additional sponsorship from WSTLUR.

& Sirmans, 1996; Hess & Almeida, 2007), difference-in-differences (with before and after comparison) (e.g., Chun-Chang et al., 2020), or a combination of both statistical methods (e.g., Gibbons & Machin, 2005; Trojanek & Gluszak, 2018). Variations and improvements in these methods include repeated sales data (e.g., Billings, 2011; McMillen & McDonald, 2004) to better control for time-invariant unobserved factors; and spatial models (e.g., Diao et al., 2017) to account for possible spatial dependencies in the dependent variable and the error term.

Empirical studies on the impact of transport on real estate prices have also made extensive use of Bernoulli variables to indicate the existence of a new station or transportation infrastructure. This station dummy treatment approach, combined with a Euclidean-based measure of proximity to that station, is employed in most papers on this topic and carries with it the underlying assumption that all new stations are equivalent and further implies a simplifying binary approach in that an urban area either receives a new station or it does not. In this paper we contend that this is an overly strong assumption, indeed one that lacks support from the empirical data. We investigate whether the employment of transportation accessibility measures within the analysis yields more comprehensive and robust results than the dummy alternative (Hypothesis 1). By measuring gains in accessibility to jobs, we adopt a continuous treatment variable that takes into account the fact that new subway stations can produce differential mobility improvements depending on their integration with the citywide transport system; the local land-use pattern; the area's location relative to city centers; the neighborhood's walkability (Lyu et al., 2020); plus other aspects that are further explored in the next section.

The use of such accessibility analysis also allows us to recognize that a new station not only affects nearby residents, but also those living in communities around existing stations within the same subway network who can now more readily reach urban opportunities located within the recently connected neighborhood. We investigate whether accessibility analysis can capture this type of network effect, one which is not adequately reflected in the station dummy approach (Hypothesis 2). To provide greater external validity to the results shown here, we apply the same methodology and analysis to the metropolises of Santiago (Chile), Sao Paulo (Brazil), Singapore, and Barcelona (Spain) and explore those commonalities arising.

Section 2 presents details about accessibility analysis and further discusses its correlations with home values. Section 3 describes our data and approach to its analysis. Section 4 shows the causal quasi-experimental empirical strategy selected for our methodological exercise. Sections 5 and 6 present our results, while section 7 estimates the total home value appreciation using the different approaches and compares the magnitude of local and network effects. Section 8 discusses our primary conclusions.

2 Access to urban opportunities and its capitalization in home values

The concept of accessibility has received increasing attention across the transportation and planning literature, as a deeper understanding of the interactions arising between the transportation system and land-use configurations has become of paramount importance in the planning of cities. From the point of view of accessibility, mobility is not an end in itself, but rather a way of reaching those opportunities a city offers, including jobs, social gatherings, leisure activities, and places of consumption. In order to reach urban opportunities, mobility is not the only parameter that matters, as the point of origin of a trip and the distance to the ultimate destination also drive residential patterns (Levinson & King, 2020). Therefore, accessibility is the product of the interaction between transport systems and land-use patterns. One of the most widely accepted and utilized definitions of accessibility was first proposed by Hansen (1959) as "the potential of opportunities for interaction." For this study, we shall understand accessibility to be a measure of potential, one which aggregates the number of urban opportunities accessible within different travel times, from a specific location, using the mobility systems available (He et al., 2019).

Subway expansions do not generate gains in accessibility that are strictly delimited to those areas bordering the new stations (termed the "local effect"), but rather such gains are diffuse and propagated throughout the entirety of the transit system ("network effect"). Residents living around extant transit stations are now more readily able to reach those opportunities available within the vicinity of the recently connected neighborhood, thereby increasing their own accessibility. However, these accessibility gains are not equally distributed across all the lines and stations within the network and essentially depend on the configuration of the subway network and the distribution of destinations of interest for those residents. Some extant stations will see high accessibility gains generated by the opening of a new station on the network, while others will not be appreciably impacted. This is the so-called network or "butterfly effect."

One of the primary reasons why people travel is for work, and commuting is the most important kind of trip in terms of explaining travel patterns using public transportation. Therefore, our analysis is centered on gains in accessibility to jobs generated by subway expansions. As the property effectively becomes "closer" to jobs via an increase in accessibility, so residents from adjacent areas will further reduce their commuting costs, particularly in relation to travel times. These, in turn, are priced into the real estate market given that accessibility is a valuable attribute. With accessibility-induced demand, real estate prices tend to go up, thereby creating additional incentives for the housing market to respond through the provision of further units. A commensurate densification of the surroundings might be accompanied by changes in land use, with an associated agglomeration of offices and industries, dependent on the location of the borough and socioeconomic profile of the affected neighborhood. Enhanced transit accessibility aids in the process of further concentrating demand, creating the condition for a more complex and heterogeneous environment of local amenities, a vital consumer benefit of urban agglomerations (Schiff, 2015). As new subway stations become gateways to jobs, amenities, and services that were previously out of easy reach, neighborhoods within walking distance of a station – typically 500 m or a quarter mile – are further benefited with accessibility gains, now by the attraction of potential new destinations (Zheng et al., 2016). Transit improvement might also help new firms benefit from agglomeration economies and, at the same time, provides workers with easy access the low-cost rental housing (Du & Zheng, 2020).

An accessibility-centered approach aids in clarifying that it is not the subway station itself that promotes real estate appreciation and other changes in the built environment, as implicit in the dummy station approach, but rather the accessibility to urban opportunities that it entails. Accessibility as a measurement has hitherto been rarely included within the real estate literature. Zheng et al. (2016) analyzed how local consumer amenities react to urban rail transit development and how this is subsequently reflected in the willingness of consumers to pay a premium to live in "subway neighborhoods." The authors incorporate accessibility within their analysis so as to capture changes in the network-level customer base generated by subway extension and go on to explain how these changes impact the formation of new amenities through a market size effect. The model that estimates the impact of new stations on the formation of consumer amenities uses an accessibility measure and yet, in the next step, when estimating the impact on home values, the paper was delimited to a station dummy strategy. Jing and Liao (2017) used a connectivity index that is essentially a gravity-based accessibility measure employing distances instead of travel times multiplied by a measure of the transit "quality" in an attempt to offset the absence of travel times. Using an empirical strategy similar to ours, the authors explored the network effect of subway expansions in Singapore. However, Jing and Liao (2017) did not consider multi-modal transit trips and theirs is a working paper as yet unpublished within a peer-reviewed journal.

Our paper, to the best of our knowledge, is the first to explore the use of multi-modal accessibility to estimate the effect of subway expansions on residential real estate prices. It is worth mentioning the work developed by Du and Zheng (2020), who investigated whether improved accessibility via expansions of the subway network benefited new firms from agglomeration economies and, at the same time, provided ready access to a wider pool of workers living in low-cost rental housing (Du & Zheng, 2020). To calculate accessibility to urban opportunities (in terms of business clusters and housing), the authors retrieved bilateral travel times and computed the commuting time-weighted urban opportunities following the standard approach within the transportation literature. In their paper, travel times were retrieved using Baidu (the Chinese equivalent to Google Maps), and these allow the authors to trace historical travel times from previous years (albeit constrained to the Chinese context). Transportation agencies in most global cities have gradually adopted the General Transit Feed Specification (GTFS) format to publish open public transportation data that allows scheduled travel times to be calculated. For this research, we worked with GTFS files and processed travel times using OpenTripPlanner, as detailed in the following section.

Dewees' seminal paper (1976) explored the impacts of subway expansion on residential property prices in Toronto and raised the question as to whether a transportation access variable has more explanatory power in regression analysis than a simple measure of distance to the city's central business district (CBD). The author used travel time access to the CBD across different transportation modes as the explanatory variable, an approach which does not fit within our accessibility concept because the distribution of urban opportunities is not considered, and the only destination considered is the city center. By comparing the R^2 values from an accompanying regression analysis, the author concluded that a "best" variable cannot be unambiguously identified (Dewees, 1976). Here we follow a similar comparative approach and find that gains in accessibility to jobs generated by the advent of new subway stations have a greater explanatory power than the classical dummy variable approach in regression analysis, as reflected in higher adjusted R^2 values and lower residual standard errors. This paper adds to the extensive literature on the impact of subway network expansions on housing prices by demonstrating the benefits of incorporating accessibility measures typically used in transportation analysis to explain the effects of subway stations on home values; incorporating both local and network effects; and comparatively exploring the impact of transit investment, thereby generalizing evidence in a cross-city perspective.

In summary, our study contributes to the literature by integrating, within a cross-cities analysis, an urban econometrics approach with a transportation accessibility analysis to bring innovation to the empirical approach commonly adopted within the literature. After a synthesis of existing studies on the topic of transit impacts on real estate values, we found that only two papers incorporate some manner of accessibility measure within their analysis (Jing & Liao, 2017; Zheng et al., 2016); only two papers explored network effects of new transit lines (Fesselmeyer & Lie, 2018; Jing & Liao, 2017); and only one performed a multi-city analysis (Baum-Snow & Kahn, 2000). We contribute to the literature by exploring the potential of a previously underutilized approach within an overstudied research topic. To the best of our knowledge, our analysis is the first to perform an international cross-city analysis in relation to the topic of transportation impacts on housing values.

3 Background and data

Three families of datasets were prerequisite for our analysis, specifically data on public transportation; housing transactions throughout the period of analysis; and points of interest (POIs) to which accessibility matters. We gathered comparable data from four different cities, namely Santiago (Chile), Singapore, Barcelona (Spain), and Sao Paulo (Brazil). Our intention is to augment the external validity of the findings in order to derive cross-cutting conclusions.

3.1 City subway network specificities

Santiago is a developing metropolitan area of nearly 7 million inhabitants (United Nations, 2019). Its first subway line began operations in 1975, covering the city's east-west axis, thereby connecting locations along Santiago's main avenue. Since then, six other lines have been inaugurated to reinforce the city's connectivity, the latest being lines 6 and 3 which began operations in 2017 and 2019, respectively. Around 30% of trips in Santiago are presently made via public transport and the subway network has an important structural role within the system, accounting for 46% of the trips made by public transport (IADB, 2019). It worth mentioning that, since 2006, public transport services were unified within the Transantiago system, with increasing fare integration between subway and bus networks and the adoption of electronic payments using smartcards. For this research, in particular, we will focus on the effect of Line 6, which was added to the network with the aim of providing greater interconnectivity between the capital's south-western industrial neighborhoods and the vibrant Providencia commune, an area that concentrates multiple amenities and offices and which has since become a second *de facto* CBD in Santiago (see Figure 1).

Sao Paulo's metropolitan region is the world's fourth most populous urban area, larger even than the metropolitan areas of London or New York. With a population of over 21 million residing in greater Sao Paulo (United Nations, 2019), the city also relies heavily on mass transit, with an approximately 30% modal partition for public transportation (IADB, 2019). Despite being the largest metropolitan region among those analyzed, it also has the shortest subway network. The city is thus highly dependent on its bus system, which is composed by more than 15,000 vehicles. Despite its limitations, the subway still plays an important structural role within the system and carries around 4 million passengers per day. The last significant wave of station openings in Sao Paulo's subway network took place in 2017 and 2018, respectively, adding a further 17 new stations, in total. These stations were all effective extensions of Line 5 (10 new stations), Line 15 (4 new stations), and Line 4 (3 new stations). The extension of Line 5 connected a previously existing section within the subway network to poorer peripheral neighborhoods lying in the southern sector of the city, thereby traversing the important tertiary centers *of Santo Amaro, Brooklin, Moema,* and *Vila Mariana*. The new stations on Line 15 further facilitated the connection of eastern peripheral neighborhoods to the already vibrant and previously well-served areas of Sao Paulo.

Barcelona's subway system is the oldest of the four cities under consideration. A city of around six million inhabitants (United Nations, 2019) has 12 established subway lines that have been developed since the end of the 19th century. Our research focuses on the latest addition to the network, the southern section of Line 9, which is planned to be the longest underground line in Europe upon completion. The so-called L9 South was opened in 2016 with 11 new stations connecting Barcelona's airport to the subway network, crossing the cities of *El Prat de Llobregat* and *L'Hospitalet de Llobregat*, and also a narrow southern area of Barcelona city called Parc Logistic. Table 1 and Figure 1 show that the environs of Barcelona's new stations have a similar concentration of jobs than the city as a whole, yet they have a higher population density. The areas surrounding L9 South are typically residential, but also specialized with the industrial and construction sectors, with location quotient[1] values of 1.28 and 1.44, respectively. Moreover, the section line surrounds crucial infrastructures such as the port, the airport, la Fira, la Ciutat de la Justícia, the FC Barcelona stadium and two Universities, thus addressing the demand from different urban areas which previously lacked a mass transit connection. This line was intended to interconnect various neighborhoods located at the fringes of the city as, prior to this development, commuters were forced to travel into the center of the city whenever they needed to change lines. We also

[1] Location quotient quantifies how concentrated a particular economic sector is in a region as compared to the city. It is calculated as the ratio of the number of firms of an industry located in an area with respect to the total number of firms in the same industry in the city.

included in our analysis a further two stations which were opened on the branch of L9 South known as L10 South.

Finally, Singapore has the most developed transit network amongst our four case studies. The city has recently invested in further expansions to improve network coverage in peripheral areas. According to the 2013 Land Transport Master Plan of Singapore, the new stations which opened between 2015 and 2017 – corresponding primarily to extensions of Downtown Line – are part of a strategy to raise the rail density from 34 km to 54 km of rail transit per million head of population, placing Singapore's subway rail density at the same level as London's Underground. The intention behind the extension of the Downtown Line was to connect to the East-West Line and the Eastern Region Line so as to better serve commuters in the eastern corridor of the city by enabling them to re-route their journeys more easily in the event of a disruption on the East-West Line.

Table 1. Subway network characteristics and stations' local urban context

CITY	Number of Subway Stations as in 2020	Number of New Stations From 2015 to 2018	New Stations Opened in	Jobs Density (per urbanized km²)		Population Density (per urbanized km²)	
				Citywide	Around New stations	Citywide	Around New stations
Santiago	118	10	2017	3,500	4,020	8,428	7,648
Singapore	182	36	2015, 2016, 2017	4,082	3,738	7,320	8,780
Barcelona	152	13	2016, 2018	3,452	3,736	8,232	11,132
Sao Paulo	92	17	2017, 2018	4,916	17,608	10,548	13,940

Source: The authors, using information on the transit system of each city and data from World Pop (www.worldpop.org), and Orbis Dataset.

Notes: Jobs and population densities are from the baseline period, prior to the opening of the stations considered in our analysis. Jobs and population densities around new stations consider a 15 min walk radius. Population density values may draw attention and be at odds with numbers available in other studies because here they were calculated for the main urbanized area of each metropolitan region. Barcelona is known as one of the densest cities in Europe, a fact restricted to the limits of the city itself, which has a population density of over 16,000 person/km^2, but not true for some of the other urban regions that compose Barcelona Metropolitan Area. Sao Paulo is the opposite case, with small variations in terms of population densities throughout the urban metropolitan area. The urbanized area of each metropolitan area was mapped from night light satellite imagery.

In summary, the analysis includes cities with very different urban and transportation realities. Located in global cities of two developed countries and two developing countries, Singapore's average monthly net salary is 7.5 times that of Sao Paulo, while Barcelona's mean income is 2.5 times that of Santiago. In terms of size, Sao Paulo has over 20 million inhabitants, some three to four times larger than that of the other cities considered. On the other hand, Singapore's subway system is two times bigger than Sao Paulo's in terms of the number of operational stations. Barcelona's subway extension comprises new stations located relatively far from the city center within a region characterized by less economic activity, while in the other cities selected, the new stations are located either in more central neighborhoods (or in new emerging centralities) and in more distant areas (see maps in Figure 1). Table 1 summarizes the number of stations considered in our analysis and the evaluation period for each case, as well as jobs and population densities around the new stations and citywide for the purpose of comparisons.

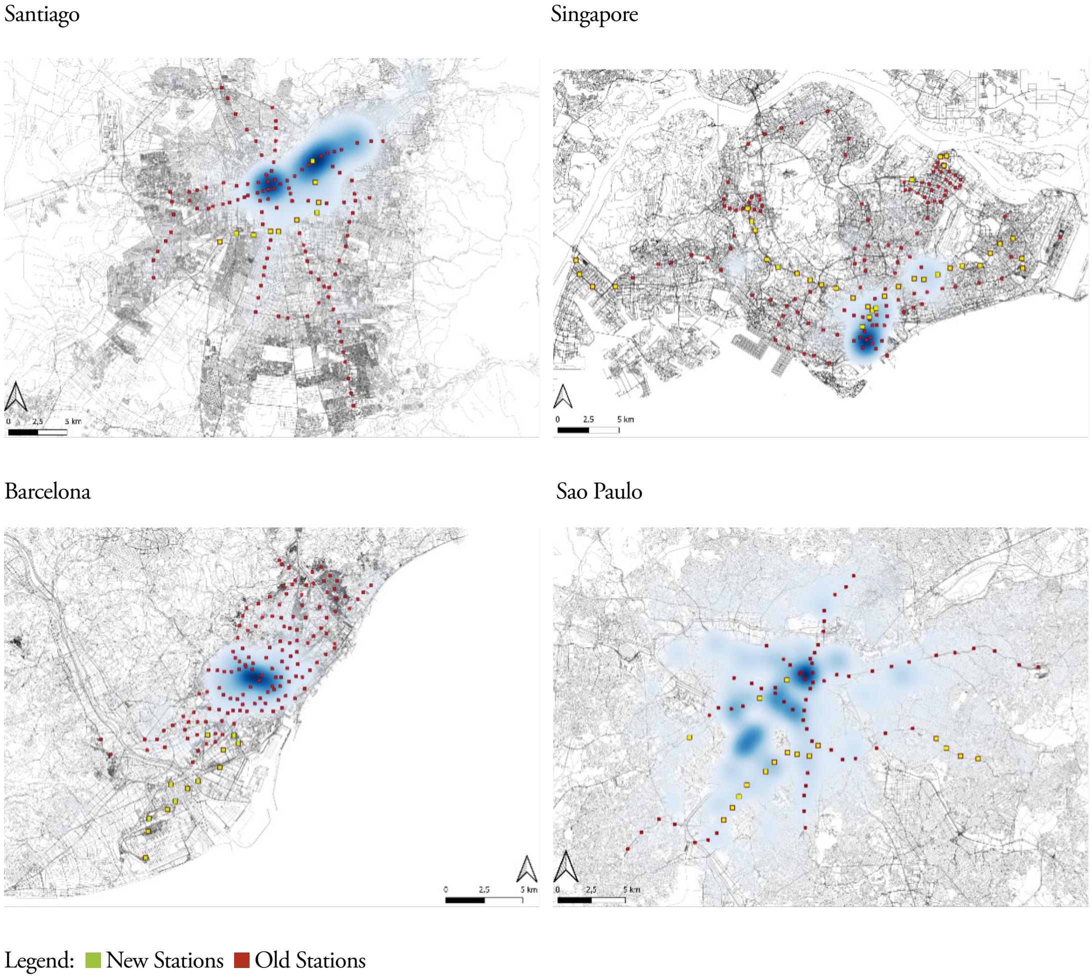

Figure 1. Jobs concentration and new subway stations
Source: The authors, using information on the transit system of each city, and the Orbis Dataset.
Note: Blue shades represent areas that agglomerate jobs.

3.2 Housing transactions

Collating real estate transaction data required special consideration, as very few cities provide public access to this type of information. In the case of Sao Paulo and Santiago, datasets were purchased from private institutions that gather this type of information for commercial use. For Barcelona, the data was acquired from a semi-private federal institution called the Colegio de Registradores de España (College of Registrars of Spain). In Singapore, data is openly available for research purposes. Each of these sources provided its datasets in different formats and the availability of key hedonic variables was disparate as were details of the real estate transactions. The datasets and their variables were thus standardized, and for the purposes of this study, we have employed the following variables: total price in USD of the real estate transactions; price per square meter in USD; the date of the transaction; and the geographic location of the property (latitude and longitude).

3.3 Points of Interest (POI)

POIs can refer to the location of any type of opportunity people may want to reach within a city. In transport theory, one of the primary reasons as to why people travel is work. The consideration of urban amenities has also been a recurrent feature of accessibility analysis, including access to restaurants and shops, green spaces, and basic services – another key focus of city planning (Levinson & King, 2020). All the mentioned types of POIs are highly valued and, therefore, are likely to drive property prices (Zheng et al., 2016). However, as commuting is the most important form of journey accounting for travel patterns within a given city, our analysis is centered on gains in accessibility to jobs generated by subway expansions.

Data on firms was obtained from Orbis, an international database with information on companies such as location, sector, and size, as well as other relevant financial and governance details that were not used for this research. For our model and accessibility analysis, we did not make a distinction nor selected particular types of firms, therefore our jobs data consider all types of jobs (financial, manufacturing, educational, retail, amongst others). Information on consumer amenities, used as control variable in our regression models, was retrieved from Foursquare, a platform that contains a global POI database, including geographic location, classification and other details from cafés, restaurants, local shops, and other types of amenities.[2] Consumer amenities data are used as control variables in our analysis, and data on firms are employed in the accessibility to jobs calculations described within the following section. Basic descriptive stats are systematized in Table 3.

3.4 Accessibility by public transportation

Despite its increasing popularity, it is important to acknowledge that accessibility fundamentally remains a construct and, therefore, may vary in its interpretation depending on the way it has been defined. In light of the diversity of the use of the concept of accessibility within the literature, any measure of accessibility will inevitably present advantages and limitations depending on its representation of preferences, data requirements, and communicability (He, et al., 2019). Cui and Levinson (2020) have categorized accessibility measures in two broad groups, namely primal measures (opportunity-denominated, which analyze how many opportunities can be reached in a given period of time), and dual measures (time-denominated, which consider the cost in terms of time of reaching a given set of opportunities).

Gravity-based accessibility is an aggregate measure of potential (or primal measure, according to Cui and Levinson (2020)'s categorization) that summarizes the opportunities for interaction and the travel impedance of accessing specified POIs from a given location (He et al., 2019). In practice, this measure identifies a certain type of destination – in our case it is jobs – which can be reached considering different travel times and the travel impedance function. This impedance function models how people perceptually penalize travel as they move further away from a specific location. The time impedance function can be formulated in a number of different ways. For a gravity-based indicator, the classic approach is a negative exponential function with a decay parameter λ, one which intuitively reflects that jobs located further away contribute less to the indicator as compared to those which can be reached in a shorter time. As λ reflects behavior, it can thus exhibit some variation across countries (Jing & Liao, 2017), a fact which poses a challenge for our cross-city analysis, as our cities do not provide comparable trip data for multiple years. In order to obtain an indicator that would allow us to compare results between cities and across time, we modelled the impedance effect by diminishing the weight of each time interval as they become larger. We evaluated a wide range of plausible λ values to test the robustness and sensitivity of our indicator and observed that, for values above 0.5, accessibility showed no significant variations. Thus, we set λ=0.5 to model the notion that jobs located 15 minutes away weigh half that of

[2] From FourSquare, we have used only the "Shops and Services" and "Restaurants" classes of POIs.

those located 30 minutes away.[3]

In addition to the time impedance function, for a gravity-based indicator the time intervals also need to be defined. In theory, time intervals can be measured in seconds to capture the finest granularity of the underlying travel and opportunity data but this inevitably leads to challenges in terms of computational capacity. For this study, we have set intervals of 15 min for all our cities, and we observed no significant changes in the indicator value by using shorter intervals. Travel times from a specific place were modeled as isochrones which represent those areas that can be reached within a specific time window (using one specific or a combination of transportation modes) when starting from a given location. We considered the fastest trip from an origin point to any point in a city made by subway, bus, walking, or any combinations of the above. In other words, we considered accessibility only in terms of public transportation, and not by private vehicles.

Isochrones differ from travel time radius in that they incorporate travel itineraries (i.e., the travel decisions and mode combinations needed to execute the planned trip). In this study, isochrones for all four cities were calculated using OpenTripPlanner, an open-source trip planner which requires GTFS data and Open Street Map data from a geographical area of interest. We obtained isochrones of up to 3 hours, considering time intervals of 15 min, for all the subway stations analyzed. To obtain the isochrones, we calculated all our travel times for a week day (Wednesday) at 8.30 in the morning, which would correspond to peak hours in all the cities included in this study. As our analysis requires the comparison of a sequence of years in which the subway network was being expanded and new stations opened, we have recreated these scenarios by modifying GTFS data that is currently available for each city. By using GTFS data we can calculate scheduled travel times for multi-modal trips – a combination of subway, bus, and walking – and, therefore, incorporate the changes into the whole network (network effect) as opposed to just the influence of the area physically next to a station. It is also important to note that the travel times we calculated using GTFS data do not incorporate real-time information, hence, our comparison of a sequence of years of subway expansion only considers planned services. We chose this approach, as opposed to using real-time data, to isolate the effects of modifications in the transport network (such as variations in bus routes) and variations in traffic conditions that can happen throughout the period of analysis.

Equation 1 shows the functional form for our accessibility indicator.

$$a_i = \sum_{j \in J_i} \lambda^{j-1} p_{i,j} \quad \forall \ \ i \in I \tag{1}$$

where a_i refers to the accessibility of station i; λ refers to a decay factor that models how the value of an opportunity decays with increasing travel impedance – in our case travel time – from a station; $p_{i,j}$ refers to the number of jobs that are located in each 15 min-isochrone area j for station i. J_i refers to the set of isochrones for station i, while I refers to the set of stations for a given city.

Table 2 summarizes the local and network accessibility gains generated by the subway extensions analyzed here following the framework and methodology as described above. Barcelona once again draws attention, in this instance because it registered the highest percentage gain in accessibility around new stations, yet showed a modest network impact compared to the number observed in Sao Paulo and Santiago. The Latin American cities draw attention for the relevance of the network impacts generated by the rail expansions analyzed.

---

[3] For this particular study, the value of the indicator itself is not relevant, as we analyze variations in accessibility. For this reason, the specifics of the functional form and the value of λ do not need to represent a city's behavior perfectly, but they do need to allow us to model sufficient granularity to capture these accessibility changes throughout a given period of time.

Table 2. Subway network characteristics and stations' local urban context

CITY	Accessibility gains around new stations (local effect)	Accessibility gains around old stations (network effect)	Total network effect / Total local effect
Santiago	7%	0.32%	34%
Singapore	9%	0.42%	19%
Barcelona	16%	1.01%	23%
Sao Paulo	6%	0.69%	39%

Source: The authors.
Note: Values for accessibility gains shown correspond to average values across new and old stations respectively.

4 Empirical strategy

To verify the impact of a new subway station on the network as a whole and its environs, we need to establish a counterfactual situation, one that allows us to consider a hypothetical scenario in which the area under analysis did not receive accessibility improvements. The ideal experiment would be to randomly select neighborhoods to implement transportation projects and, after a period, the altered real estate metrics in randomly selected areas can be compared with metrics from those areas which were not selected. Differences arising in real estate prices would thus serve as an actual measure of the impact of transportation investments. However, we cannot perform this idealized experiment, as transportation investments are implemented rationally and not in a random manner. Consequently, we must adopt a quasi-experimental control strategy, one that can produce unbiased results. Previously, when evaluating the impact of transportation investments on the characteristics of the urban built environment, this challenge has been commonly addressed using difference-in-differences (DID) models which employ a distance buffer around new transportation facilities to define "treated areas," an approach that is commonly combined with matching techniques to select control areas (e.g., Du & Zheng, 2020; Jing & Liao, 2017).

In keeping with standardized empirical procedures in the literature, we divided our cities into a grid of square cells measuring 0.5 km on each side. Neighborhood variables were used as controls, and the number of amenities, jobs, and population were aggregated at a cellular level. We also used the grid cells to define the treated areas as those whose centroids fall within the 15-minute walking distance buffer from a new subway station. Control cells were selected using propensity score matching, which was estimated using the *matchit* function in R which incorporated the following variables for the grid cells: distance to the city center; density of consumer amenities; density of jobs; the number of real estate transactions in the area; and population density. Table 3 presents basic descriptive stats for the treated and control grid cells. Additional material (available in the Appendices) shows that the matching procedure performed as anticipated and created a "common ground" between treatment and control groups.

Although some neighborhood characteristics have been calculated for grid cells *i*, the observation unit in our model is the real estate transaction k. More specifically, we will run regressions with the specification detailed in Equation 2.

$$ln(Price_{k,i,t}) = \beta_0 + \beta_1 Treatment + \beta_2\, ln(NJobs_{i,t0}) + \beta_3\, ln(NAmenities_{i,t0}) + \beta_4\, DistCBD_i + \beta_5\, ln(Pop_{i,t0}) + \beta_6\, MeanPrice_{i,t0} + FE : SaleYear_{k,t} + FE : Station_k + FE : Treat_k \quad (2)$$

Where $\boldsymbol{Price_{k,i,t}}$ is the outcome of interest, namely the price per square meter of the unit transacted k located in the cell i in the year t. Here, t_0 refers to the baseline period and $t\text{-}1$ to the year before t. In station dummy models, ***Treatment*** is said to be 1 if the transaction arises in the treated area and occurred in the year after the station opened; and in accessibility gain models, ***Treatment*** is the logarithm of the increment in accessibility to jobs promoted by the new subway station during the period *t-1* if the transaction arises in the treated area and occurred after the station opened. In other words, in terms of accessibility treatment models, ***Treatment*** is $\boldsymbol{ln(\Delta AccesJobs_{i,t-1})}$. We control for a set of variables which are known determinants of housing prices, specifically the number of jobs and consumer amenities (***NJobs*** and ***NAmenities***); the distance from the city center (***DistCBD***); the number of residents (***Pop***); and the average value of real estate transactions (***MeanPrice***). These control variables are calculated for the grid cell i during the baseline period and represent the neighborhood characteristics. It is noteworthy that a set of neighborhood characteristics that we cannot observe, such as the existence of green areas or proximity to good schools, are somehow reflected in the $\boldsymbol{MeanPrice_{i,t0}}$ variable. Models that use accessibility measures also control for the baseline level of accessibility to consumer amenities, via $\boldsymbol{ln(AccessAmenities_{i,t0})}$, to reflect local transit service conditions. The specification as outlined essentially follows the DID approach (Wooldridge, 2013), and considers fixed effects for year, treatment, and the subway station itself. We report the results with standard beta coefficients to allow for magnitude comparisons, as the coefficients of impact on access gain models are not easy to interpret.

Table 3 presents some basic descriptive statistics on the accessibility to jobs, home values, and local POI variables for both the selected treated and control areas of each city.

Table 3. Descriptive stats of treated and control cells by city

		Baseline Treated Cells	Std. Dev Treated Cells	Change Treated Cells	Baseline Control Cells	Std. Dev Control Cells	Change Control Cells
Santiago	Access to Jobs	54,153.43	19,940.39	3,996.34	50,609.25	23,668.34	131.62
	Price Sqm (U$)	1,560.35	1,131.88	426.81	1,592.93	1,243.02	380.20
	Number of Consumer Amenities	35.14	37.64	3.86	27.11	35.40	2.86
	Number of Jobs	1,005.11	547.12	39.43	1,376.96	678.36	20.45
Singapore	Access to Jobs	103,954.62	75,410.64	9,660.97	68,253.46	51,847.00	1,053.63
	Price Sqm (U$)	9,509.53	2,492.80	1,114.33	8,744.22	2,227.99	867.98
	Number of Consumer Amenities	40.74	56.94	5.14	35.98	60.84	4.22
	Number of Jobs	934.74	767.17	32.49	987.25	872.19	30.38
Barcelona	Access to Jobs	16,822.45	6,308.24	2,656.59	10,123.25	9,487.17	323.35
	Price Sqm (U$)	2,145.05	875.05	292.76	2,198.17	708.07	363.53
	Number of Consumer Amenities	11.41	18.08	1.50	16.52	32.15	1.72
	Number of Jobs	933.53	990.19	28.36	927.02	634.67	7.67
Sao Paulo	Access to Jobs	107,492.68	75,716.58	685.91	57,949.87	64,516.71	126.97
	Price Sqm (U$)	1,370.32	967.27	234.84	1,457.43	1,292.94	335.00
	Number of Consumer Amenities	89.31	97.44	8.50	88.94	101.93	8.22
	Number of Jobs	4,402.06	4,280.75	546.28	4,594.12	11,042.02	520.40

Source: The authors

5 Accessibility in lieu of the station dummy treatment variable

Using the specifications detailed above, our analysis aimed to compare the results and goodness-of-fit between station dummy and accessibility gain models. The most widely used parameter for comparing non-nested models is the Adjusted R^2 (Wooldridge, 2013). The residual standard error is also a good measure of goodness-of-fit for models with the same dependent variable. Table 4 shows the results of the regressions, as previously specified, for the different cities under evaluation. The first four columns report those parameters yielded when using the dummy strategy, and the following columns show the results derived for the accessibility-to-jobs continuous treatment variable.

Table 4. Comparing local effects on housing prices using station dummy and accessibility to jobs

Dependent variable: ln(PriceUS)| Accessibility and Station Dummy Treatment Variable | Local Effect

	Station Dummy				Accessibility Continuous Variable			
	STG	SGP	BCN	SPL	STG	SGP	BCN	SPL
Ln Number of Amenities	0.053***	0.008	0.000	0.054***	0.016***	0.042***	-0.022	0.018***
Baseline, cell	(0.007)	(0.002)	(0.004)	(0.006)	(0.002)	(0.001)	(0.003)	(0.002)
Ln Number Jobs	-0.051***	-0.083***	-0.039	-0.037***	-0.010**	-0.098***	-0.005	-0.007
Baseline, cell	(0.004)	(0.003)	(0.010)	(0.003)	(0.002)	(0.002)	(0.010)	(0.002)
DistCBD	-0.060**	-1.183***	0.369***	-0.047**	-0.032**	-0.703***	0.326***	-0.043***
Cell	(0.053)	(0.101)	(0.206)	(0.044)	(0.030)	(0.070)	(0.177)	(0.025)
Ln Population	-0.057***	0.051	0.034	-0.063***	-0.031***	0.011	-0.029	-0.041***
Baseline, cell	(0.036)	(0.020)	(0.038)	(0.030)	(0.021)	(0.010)	(0.030)	(0.017)
Ln Mean Price	0.520***	0.682***	0.561***	0.472***	0.609***	0.654***	0.521***	0.604***
Baseline, cell	(0.060)	(0.017)	(0.036)	(0.063)	(0.013)	(0.014)	(0.027)	(0.011)
Ln Access Jobs	0.046***	0.010	0.018	0.040***	0.006	0.017	0.092***	-0.010
Baseline, cell	(0.033)	(0.017)	(0.018)	(0.023)	(0.025)	(0.013)	(0.013)	(0.019)
ΔTreatment Variable	0.030***	0.048***	0.012	0.021**	0.033***	0.011**	-0.022	0.037***
Ln Access to Jobs or Station dummy	(0.028)	(0.016)	(0.057)	(0.025)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.000***	0.000***	0.000	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.780)	(0.440)	(0.734)	(0.715)	(0.399)	(0.316)	(0.510)	(0.314)
Year, station, and treatment Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

	Station Dummy				Accessibility Continuous Variable			
	STG	SGP	BCN	SPL	STG	SGP	BCN	SPL
Observations	26,474	12,876	5,078	37,069	26,474	12,876	5,078	37,069
R^2	0.522	0.526	0.391	0.507	0.558	0.557	0.336	0.564
Adjusted R^2	0.521	0.523	0.386	0.506	0.557	0.554	0.330	0.564
Residual Std. Error	0.764	0.584	0.681	0.771	0.640	0.337	0.438	0.637
F Statistic	389.598***	189.304***	77.104***	513.615***	450.957***	214.138***	60.623***	647.566***

Note:
*p**p***p<0.01 | Clustered Robust Standard Errors for stations

First, we note that both the station dummy and accessibility gain models reported significant coefficient impacts (*Δ Treatment Variable*) in Santiago, Singapore, and Sao Paulo, although yielding different standard beta coefficient values. For these three cities, the values of the adjusted R^2 are constantly higher (ranging around 0.55) and the residual standard errors lower in the accessibility gain model than for the station dummy one, suggesting that the continuous treatment variable consistently provides a better fit of the data, where relevant, thereby explaining the dependent variable.

Interestingly, Barcelona differs from the pattern we observed for other cities, as both models reported a non-significant impact and, while the station dummy model reported a higher R^2, it also reported higher residual standard error values in comparison to the accessibility gain model. Note that Barcelona's regressions returned the lowest R^2 values as compared to the other cities and the coefficients' standard errors were consistently the highest. Standard error values, as reported in regression models, are dependent on the number of independent variables included in the model; the number of observations available to fit the chosen model; and the deviation of the data set from the assumed regression model. Although Barcelona has the smallest sample size among the cities analyzed, there were more than 5,000 data points available to fit the model which is, in general, not considered to be a small data set. The number of regressors was the same among the models used for the other cities, and thus it cannot be considered a viable explanation for the high standard deviation observed within Barcelona's results.

What remains to be determined is the exploration of different functional forms in our search for the best goodness-of-fit. This approach is especially relevant in order to recognize the specificity of Barcelona's new subway stations as previously discussed in Section 3. Even though Barcelona's subway system is used daily by divergent social groups, those families living in more affordable neighborhoods are typically more mass transit-dependent. This led us to explore (among other several functional forms) the interaction of our treatment variables with the average value of traded properties within the same grid cell. In this approach, the model with the interaction term yielded lower values of residual standard error and a slight increase in R^2, and not only for Barcelona, as reported in Table 5. The coefficient standard errors remained the highest in Barcelona, yet the model exploring the heterogeneous impacts produced more meaningful and robust results.

Table 5. Comparing heterogeneous local effects on housing prices using station dummy and accessibility to jobs

Dependent variable: ln(PriceUS) | Accessibility and Station Dummy Treatment Variable

	Station Dummy				Accessibility Continuous Variable			
	STG	SGP	BCN	SPL	STG	SGP	BCN	SPL
Ln Number of Amenities	0.051***	0.017	-0.010	0.053***	0.016***	0.035***	-0.027*	0.018***
Baseline, cell	(0.007)	(0.002)	(0.004)	(0.006)	(0.002)	(0.001)	(0.003)	(0.002)
Ln Number Jobs	-0.051***	-0.086***	-0.034	-0.037***	-0.013**	-0.113***	0.019	-0.008*
Baseline, cell	(0.004)	(0.003)	(0.009)	(0.003)	(0.002)	(0.002)	(0.009)	(0.002)
DistCBD	-0.059**	-1.175***	0.176*	-0.047**	-0.033**	-0.653***	0.334***	-0.045***
Cell	(0.053)	(0.101)	(0.202)	(0.044)	(0.030)	(0.070)	(0.177)	(0.025)
Ln Population	-0.055***	0.046	0.076*	-0.061***	-0.031***	0.013	-0.038	-0.041***
Baseline, cell	(0.036)	(0.020)	(0.039)	(0.030)	(0.021)	(0.010)	(0.031)	(0.017)
Ln Mean Price	0.524***	0.693***	0.606***	0.476***	0.609***	0.622***	0.518***	0.602***
Baseline, cell	(0.062)	(0.017)	(0.038)	(0.064)	(0.014)	(0.015)	(0.028)	(0.011)
Ln Access Jobs	0.046***	0.014	-0.021	0.040***	0.006	0.015	0.095***	-0.010
Baseline, cell	(0.033)	(0.017)	(0.018)	(0.023)	(0.025)	(0.013)	(0.013)	(0.019)
ΔTreatment Variable	0.335**	1.983***	3.837***	0.303**	0.154*	1.710***	1.941***	0.213***
Ln Access to Jobs or Station dummy	(0.402)	(0.393)	(0.918)	(0.363)	(0.012)	(0.015)	(0.050)	(0.010)
ΔTreatment * Ln Mean Price	-0.306**	-1.941***	-3.818***	-0.283**	-0.121*	-1.691***	-1.961***	-0.177**
Ln Access to Jobs or Station dummy	(0.034)	(0.028)	(0.077)	(0.030)	(0.001)	(0.001)	(0.004)	(0.001)
Constant	0.000***	0.000***	0.000	0.000***	0.000***	0.000***	0.000***	0.000***
	(0.798)	(0.446)	(0.734)	(0.738)	(0.405)	(0.319)	(0.510)	(0.319)
Year, station, and treatment Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	26,474	12,876	5,078	37,069	26,474	12,876	5,078	37,069
R^2	0.522	0.527	0.394	0.507	0.558	0.560	0.398	0.564
Adjusted R^2	0.521	0.524	0.389	0.506	0.557	0.558	0.392	0.564
Residual Std. Error	0.764	0.583	0.680	0.770	0.640	0.336	0.438	0.637
F Statistic	384.525***	187.463***	76.143***	506.904***	450.199***	214.703***	59.660***	639.115***

Note:
*p**p***p<0.01 | Clustered Robust Standard Errors for stations

Table 5 reports significant impact parameters (Δ*Treatment Variable* and Δ*Treatment**Ln Mean Price*) across all the cities, including Barcelona, revealing that the proportional impact becomes consistently higher as the neighborhood becomes more affordable. In the model incorporating the interaction term, the comparison between the station dummy and access treatment remained as previously described, yielding a lower residual standard error and higher adjusted R^2 in access gain models. The evidence presented confirms Hypothesis 1, as previously stated, and reveals that the use of accessibility analysis in the econometric causal inference can yield better goodness-of-fits than the station dummy variable, an approach commonly used within the extant literature. Accessibility-centered analyzes also allow for the inclusion of more accurate transportation variables, those which consider multi-modal trips and reflect the level of services more precisely than the proximity to a station binary variable.

6 Accessibility measures reveal hidden network impacts

Hypothesis 2 is also related to the potential of using accessibility-to-jobs continuous treatment variables and states that network effects are not taken into account in station dummy causal identification strategies, thereby underestimating the impacts of a new subway station. As previously detailed, expansions in the subway network do not generate gains in accessibility to jobs which are solely restricted to those neighborhoods that border the new stations, but rather such gains are spread throughout the transport network. However, these network gains are not equally distributed amongst all lines and stations, and such gains essentially depend on the configuration of the subway network and the spatial distribution of jobs. Some previously existing stations will see high accessibility gains generated by the opening of a new station on the network, while others will not be impacted. Thus, a continuous treatment variable might better reflect the heterogeneity of network gains and do so more accurately than the binary alternative. Measuring network effects using the station dummy strategy means assigning to the ***Treatment*** variable the value 1 to all those transactions proximal to any old subway station in the system. This, however, is arbitrary and imprecise, as several stations might be unaffected by a new line or station in the transportation system.

Table 6 reports the results for the network effects using the station dummy strategy and accessibility analysis. The regression models and other methodological procedures are otherwise the same as described in the previous section, but now only real estate transactions near old subway stations were considered as treated units. In other words, treatment and control areas of the new stations analyzed in the previous section were not considered in this network effect analysis. Another specificity of the network effect model is the use of level values of accessibility gains rather than logarithm ones. As network accessibility gains are less intense and more homogeneously distributed, the logarithmic form of this variable reduces variations and produces less relevant results.

Table 6 confirms that accessibility measures are capable of capturing network impacts that are not otherwise possible using the station dummy variable strategy. Statistically significant network impacts were found in Santiago, Barcelona and Sao Paulo, those cities in which network accessibility gains are more relevant (see Table 2). In Singapore, no significant network impact was found, but this result is expected when we verify that the effects on the network represent only 19% of the total accessibility gain generated. The same proportionality measure represented a 23% increase in Barcelona, 34% in Santiago, and 39% in Sao Paulo. Such network impacts are occluded within the station dummy model. This is because the network effect is extremely heterogeneous, and new stations open elsewhere do not affect several old subway neighborhoods. Therefore, the average impact is minimal, as shown in the next section. The binary alternative cannot capture the very little average effect and reflect how heterogeneous is the butterfly treatment. In other words, the dummy approach introduces too much noise and cannot show statistical significance in the network analysis.

Table 6. Network effect on housing prices for existing subway stations

Dependent variable: ln(PriceUS)| Accessibility and Station Dummy Treatment Variable | Network Effect

	Station Dummy				Accessibility Continuous Variable			
	STG	SGP	BCN	SPL	STG	SGP	BCN	SPL
Ln Number of Amenities	0.013***	-0.040***	-0.048***	0.013***	0.013***	-0.038***	-0.048***	0.013***
Baseline, cell	(0.001)	(0.003)	(0.005)	(0.001)	(0.001)	(0.003)	(0.005)	(0.001)
Ln Number Jobs	-0.014**	0.023***	0.015	-0.004	-0.014**	0.024***	0.016	-0.004
Baseline, cell	(0.001)	(0.001)	(0.009)	(0.001)	(0.001)	(0.001)	(0.009)	(0.001)
DistCBD	-0.012	0.004	-0.021	-0.001	-0.012	0.003	-0.023	-0.002
Cell	(0.016)	(0.078)	(0.206)	(0.011)	(0.016)	(0.078)	(0.206)	(0.011)
Ln Population	-0.072***	-0.056**	0.026	-0.067***	-0.071***	-0.060**	0.027	-0.067***
Baseline, cell	(0.015)	(0.020)	(0.036)	(0.013)	(0.015)	(0.020)	(0.035)	(0.013)
Ln Mean Price	0.620***	0.575***	0.658***	0.627***	0.620***	0.579***	0.659***	0.628***
Baseline, cell	(0.009)	(0.019)	(0.035)	(0.008)	(0.010)	(0.019)	(0.035)	(0.008)
Ln Access Jobs	0.020**	0.038**	-0.000	0.017**	0.020**	0.038**	-0.001	0.017**
Baseline, cell	(0.012)	(0.011)	(0.020)	(0.010)	(0.012)	(0.011)	(0.020)	(0.010)
ΔTreatment	-0.003	0.053	0.048	-0.003	0.009***	0.005	0.007**	0.006**
Access to Jobs or Station dummy	(0.014)	(0.031)	(0.046)	(0.012)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.000***	0.000***	0.000	0.000***	0.000***	0.000***	0.000	0.000***
	(0.238)	(0.381)	(0.716)	(0.203)	(0.238)	(0.383)	(0.718)	(0.203)
Year, station, and treatment	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effect								
Observations	96,283	26,188	7,393	126,085	96,283	26,188	7,655	126,085
R^2	0.568	0.508	0.486	0.570	0.568	0.508	0.481	0.570
Adjusted R^2	0.568	0.506	0.482	0.569	0.568	0.505	0.476	0.569
Residual Std. Error	0.682	0.479	0.674	0.686	0.682	0.480	0.670	0.686
F Statistic	1,182.986***	228.164***	115.428***	1,543.702***	1,172.761***	227.633***	108.154***	1,543.788***

Note: *p**p***p<0.01 | Clustered Robust Standard Errors for stations

The number of observations now is higher than observed in previous models, as more areas are considered for network effects, and not only those regions bordering the new stations. Within such network effect models, adjusted R^2 and residual standard error values were found to be similar between both approaches.

The results presented so far serve as evidence that using accessibility to jobs as a continuous treatment variable instead of its binary alternative, the station dummy approach, yields a better goodness-of-fit in the selected functional form based on DID quasi-experimental strategy. Further, accessibility measures allow for the exploration of impacts beyond the local effects around new subway stations, shedding light on network impact that has hitherto been largely overlooked in the literature. In addition, the verification of network effects also serve as a robustness test to confirm the impact of new subway stations on home prices because network accessibility gains are plausibly exogenous, as they are generated by stations opened elsewhere and are not likely to be associated with any changes or characteristics in the local neighborhoods of pre-existing stations (for a strong case exploring the exogeneity of the network effects for causal verification, see Du & Zheng, 2020).

7 Monetizing the impact of new subway stations

To allow a comparison of the coefficients we have thus far used the beta standardized coefficients. The downside of this approach is that its interpretation is not intuitive for the dummy station approach. In the accessibility approach, the beta coefficient can be interpreted in terms of one standard deviation change in accessibility values, whereas in the dummy approach the coefficient's interpretation would be as one standard deviation of a "new station," which does not make logical sense. The upside is that the effects of scale are thereby eliminated and, while having different units is not in of itself a problem, there is no technical obstacle to using beta coefficients together with binary independent variables. As our primary research question is to compare the two models, the standardized option was chosen. For a more intuitive interpretation of our results, we estimated (in USD) the total home value appreciation produced by the subway investments using regular regression coefficients from the accessibility model, specifically its corresponding treatment variable. This exercise enabled us to compare the magnitude of local and network effects using an accessibility analysis strategy.

Using the regular coefficients of impact, we calculated the predicted values added in those residential transactions that occurred in the year following the opening of new subway stations. We must be aware that the quantities reported in Table 7 do not represent the total value added to the real estate market from the mass transit investments under analysis, but rather only the predicted appreciation in home value transactions observed over the year following the station's opening. Table 7 presents both the local and network average percentage increases in home values (1st and 3rd rows) produced by new subway stations, as well as the sum of the values added (2nd and 4th rows) in the aforementioned transactions for each city. Lastly, Table 7 (5th row) compares the real estate values added by local and network effects by simply dividing the last by the former (4th over 2nd rows).

Access gain models yield an impact between 2.5% and 3% increase in home values based on the average accessibility gain produced by a new station. Singapore's homeowners most highly value the utility of the subway. For this metropolitan area, we observed almost 3% average increase in residential values around the new stations which represent, at least in Singapore's booming real estate market, an average appreciation of around thirty-five thousand USD per unit, amounting almost one-hundred and eighty million dollars added in terms of the real estate transactions that occurred in the year after the new stations opened. In Santiago and Sao Paulo, both cities in developing countries, we found a similar percentage in terms of appreciation (2.52% and 2.80%, respectively), but a much lower average appreciation per unit (US$ 2,590 and US$ 4,105, respectively). When comparing the total appreciation in USD shown in Table 7, we should consider that Singapore's subway expansion included 36 new stations, while Santiago's and Sao Paulo's new investments included only 10 and 17 new stations, respectively.

Table 7. Local and network effects of new subway stations on housing prices in USD

	Accessibility Continuous Variable			
	STG	SGP	BCN	SPL
Local percentage impact Average increase	2.52%	2.98%	2.83%	2.80%
Local added value impact Total US$ added	53,347,671	179,826,420	13,637,395	70,208,240
Network percentage impact Average increase	0.18%		0.06%	0.11%
Network added value impact Total US$ added	20,688,762		1,060,339	17,823,565
Network / Local added value impact Percentage	38.78%	0.00%	7.78%	25.39%

Note: These are predicted values considering only those residential transactions that occurred in the year following the opening of the stations analyzed.

This monetization exercise also allows us to intuitively compare the magnitudes of local and network effects. Significant network effects were captured in Santiago, Barcelona and Sao Paulo using access gain models. The average percentage appreciation produced by network effects is thus much lower than the local appreciation produced by the new stations, as anticipated. However, the rather faint appreciation induced by network effects is dispersed over a much larger area than local effects and, therefore, the added value becomes quite significant. In Santiago and Sao Paulo, the network effect represents 38% and 25%, respectively, of the local value increase, while in Barcelona it is circa 8%.

Finally, the accessibility approach also allows us to capture the more nuanced spatial heterogeneity of the impacts generated by new stations, as different locations experience different increases in accessibility. The series of maps in Figure 2 illustrate how heterogeneous the impact predicted by the access gain models is within those cities analyzed. The heatmaps depict those units with a relatively higher predicted increase in home values due to accessibility gains generated by the subway extension by local and network effects. Blue shades represent concentration of home values generated around new subway stations by local accessibility gains, and red shades show areas with significant real estate appreciation created by network effect.

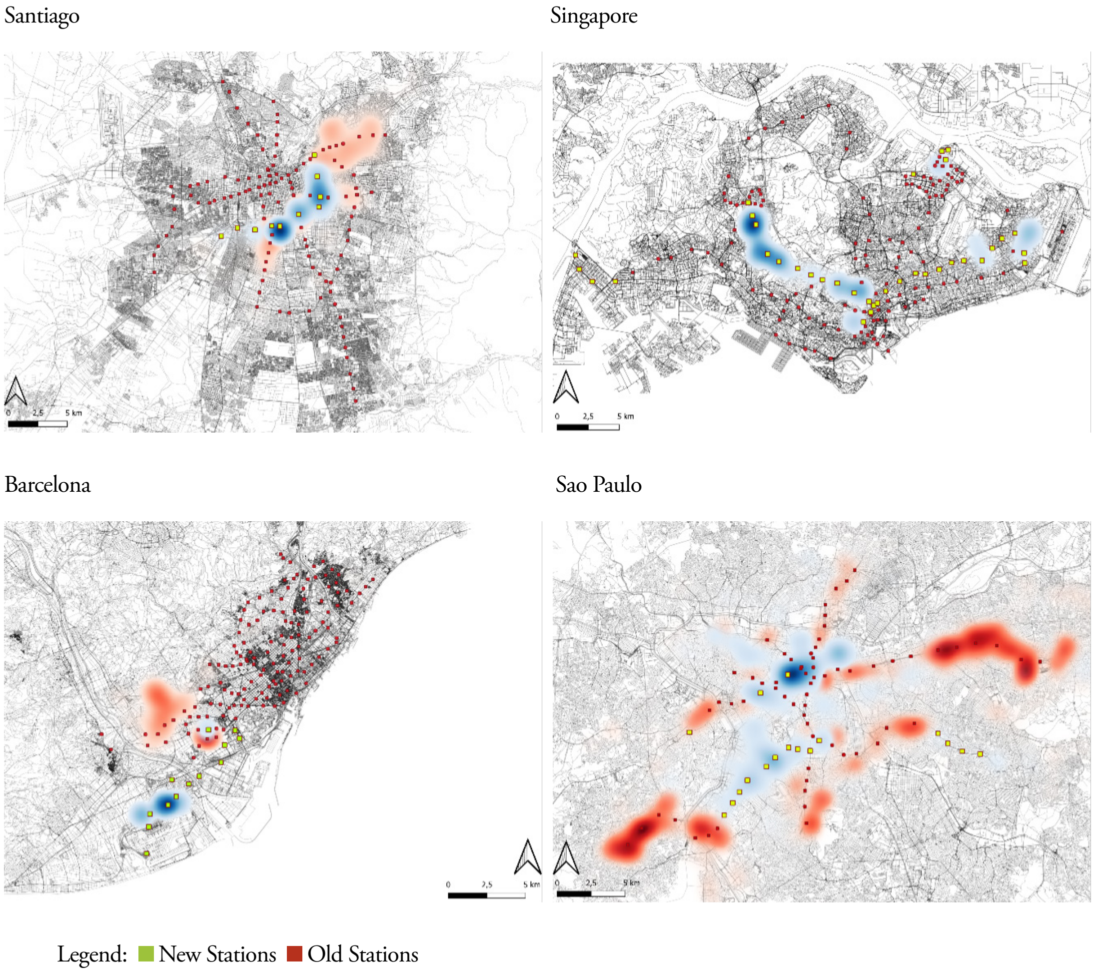

Figure 2. Predicted residential prices appreciation generated by the new subway stations
Source: The authors.
Note: Blue and red shades represent concentration of home values generated by local and network accessibility gains, respectively.

8 Conclusion

The empirical results demonstrate the viability of using accessibility analysis for studies of the impact of new transport infrastructure on the real estate market. We first show that the use of a continuous accessibility gains variable in place of its binary alternative, the station dummy approach, yields a better goodness-of-fit with a higher adjusted R^2 and lower residual standard error. Further, such accessibility measurements allow us to explore the impact of transportation investments beyond their local effects, thereby elucidating a network impact that has, to date, been largely overlooked within the literature. We explore how the integration of urban economics and knowledge of transportation analysis and their associated methodologies can lead to innovations in the empirical approach commonly adopted within the literature and yield more robust and comprehensive results. To extend greater external validity to the results shown here, we applied the same methodology to four cities, namely Santiago (Chile), Singapore, Barcelona (Spain), and Sao Paulo (Brazil).

This paper has sought to contribute methodologically to a substantial branch of the literature whose research findings have relevant ramifications for transportation decision-makers who need to technically support substantial investment decisions on transportation policies in terms of municipal resources.

From the perspective of public policymakers, it is essential to consider robust and accurate measurements of the potential impact of investments of high social and urban relevance, such as a new subway line. Cost-benefit and welfare analyzes are thus of substantial importance to public decision-making processes and are highly dependent on those parameters calculated in academic studies. Such studies, therefore, should seek the most accurate measures available. In the context of this study, our findings suggest that network impacts can constitute a relevant fraction of the total appreciation generated by subway expansion – up to 40% among our case studies – thereby opening a new horizon of opportunities for the implementation of land value capture programs, an aspect which we contend merits greater attention in future research.

Our results are also of interest to multilateral institutions and international banks, including the World Bank Group, Inter-American Development Bank, and Asian Development Bank, all of whom provide loans to both national and local governments as well as to private companies implementing transportation projects. Naturally, such institutions require strictly accurate impact assessment studies. Although we have focused on subway infrastructure for the purposes of this study, our methodology can readily be adapted to consider other types of transportation infrastructure investment. Finally, although our findings have suggested generalized trends across our four cities, we believe that differences observed in both local and network effects across stations may be explained by differing neighborhood characteristics that set the initial scenario for analysis. These parameters, for instance, include land-use regulations, income geographic distribution, urban design, inter alia, and also the level of maturity of the transportation network in terms of its network extension, multi-modality, affordability of the transit system, etc. Thus, understanding the role that these initial conditions play in shaping later local and network effects may also be of wider interest within the field of urban economics.

Acknowledgments

We gratefully acknowledge Kristopher Stephen Steele and Tabea Sonnenschein for their superb research assistance; He He and Fabio Duarte from MIT for their valuable comments received during early versions of this paper; and the participants of the Sustainable Urbanization Lab Seminar Series.

Appendix

Appendix available as a supplemental file at www.jtlu.org/index.php/jtlu/rt/suppFiles/2146/0.

References

Baum-Snow, N., & Kahn, M. E. (2000). The effects of new public projects to expand urban rail transit. *Journal of Public Economics, 77*(2), 241–263.

Benjamin, J. D., & Sirmasn, G. (1996). Mass transportation, apartment rent and property values. *Journal of Real Estate Research, 12*(1), 1–8.

Billings, S. B. (2011). Estimating the value of a new transit option. *Regional Science and Urban Economics, 41*(6), 525–536.

Cui, M., & Levinson, D. (2020). Multi-activity access: How activity choice affects opportunity. *Transportation Research Part D: Transport and Environment, 85,* 102364.

Chun-Chang, L., Chi-Ming, L., & Hui-Chuan, H. (2020). The impact of mass rapid transit system on neighborhood housing prices: An application of difference-in-difference and spatial econometrics. *Real Estate Management and Valuation, 28*(1), 28–40.

Dewees, D. N. (1976). The effect of a subway on residential property values in Toronto. *Journal of Urban Economics, 3*(4), 357–369.

Debrezion, G., Pels, E., & Rietveld, P. (2007). The impact of railway stations on residential and commercial property value: A meta-analysis. *The Journal of Real Estate Finance and Economics, 35*(2), 161–180.

Diao, M., Leonard, D., & Sing, T. F. (2017). Spatial-difference-in-differences models for impact of new mass rapid transit line on private housing values. *Regional Science and Urban Economics, 67*, 64–77.

Du, R., & Zheng, S. (2020). Agglomeration, housing affordability, and new firm formation: The role of subway network. *Journal of Housing Economics, 48*, 101668.

Fesselmeyer, E., & Liu, H. (2018). How much do users value a network expansion? Evidence from the public transit system in Singapore. *Regional Science and Urban Economics, 71*, 46–61.

Gibbons, S., & Machin, S. (2005). Valuing rail access using transport innovations. *Journal of Urban Economics, 57*(1), 148–169.

He, H., Ponce-Lopez, R., Shaw, J., Le, D.-T., Ferreira, J., & Zegras, P. C. (2019). Representing accessibility: Evidence from vehicle ownership choices and property valuations in Singapore. *Transportation Research Record, 2673*(2), 724–733.

Hansen, W. G. (1959). How accessibility shapes land use. *Journal of the American Planning Association, 25*(2), 73–76.

Hess, D. B., & Almeida, T. M. (2007). Impact of proximity to light rail rapid transit on station-area property values in Buffalo, New York. *Urban Studies, 44*(5–6), 1041–1068.

IADB. (2019). *Hechos estilizados de transporte urbano en América Latina y el Caribe* (IDB-TN-1640, Rep.). Washington, DC: Inter-American Development Bank.

Jing, K., & Liao, W. (2017). *The butterfly effect of transport network connectivity and welfare implication.* Retrieved from https://ssrn.com/abstract=3089360 or http://dx.doi.org/10.2139/ssrn.3089360

Levinson, D., & King, D. (2020). *Transport access manual: A guide for measuring connection between people and places.* Sydney: Committee of the Transport Access Manual, University of Sydney.

Lyu, G., Bertolini, L., & Pfeffer, K. (2020). How does transit-oriented development contribute to station area accessibility? A study in Beijing. *International Journal of Sustainable Transportation, 14*(7), 533–543.

McMillen, D. P., & McDonald, J. (2004). Reaction of house prices to a new rapid transit line: Chicago's Midway Line, 1983–1999. *Real Estate Economics, 32*, 463–486.

Mohammad, S. I., Graham, D. J., Melo, P. C., & Anderson, R. J. (2013). A meta-analysis of the impact of rail projects on land and property values. *Transportation Research Part A: Policy and Practice, 50*, 158–170.

Schiff, N. (2015). Cities and product variety: Evidence from restaurants. *Journal of Economic Geography, 15*(6), 1085–1123.

Trojanek, R., & Gluszak, M. (2018). Spatial and time effect of subway on property prices. *Journal of Housing and the Built Environment, 33*, 359–384.

United Nations. (2019). *World urbanization prospects: The 2018 revision* (ST/ESA/SER.A/366). New York: United Nations, Department of Economic and Social Affairs, Population Division.

Wooldridge, J. M. (2013). I*ntroductory econometrics: A modern approach* (5th ed., Vols. 120–121). Boston: Cengage Learning.

Zheng, S., Xu, Y., Zhang, X., & Wang, R. (2016). Transit development, consumer amenities and home values: Evidence from Beijing's subway neighborhoods. *Journal of Housing Economics, 33,* 22–33.