Junio 2022

WP4-2022-004
N° de serie

DOCUMENTO DE TRABAJO

Announcement of the First Metro Line and its Impact on Housing Prices in Bogotá

Autores

Jeisson Cárdenas Juan Miguel Gallego María Alejandra Urrutia

COLOMBIA CIENTÍFICA
Conocimiento Global para el Desarrollo

ALIANZAEFI
Economía Formal e Inclusiva

Announcement of the First Metro Line and its Impact on Housing Prices in Bogotá

Jeisson Cárdenas [a], Juan Miguel Gallego [b], María Alejandra Urrutia [c]

[a] Institute for Employment Research, University of Warwick Coventry, United Kingdom. E-mail: jeisson.cardenas-rubio@warwick.ac.uk
[b] Facultad de Economía Universidad del Rosario, Bogotá, Colombia. E-mail: juan.gallego@urosario.edu.co
[c] Calle 25 # 68ª- 70, Bogotá, Colombia. E-mail: aleja.urrutia792@gmail.com

Abstract

Transport infrastructure investment not only impacts mobility but may also significantly affect the capitalisation of the areas surrounding these infrastructure projects, even before the projects are completed. This paper analyses the impact that the announcement of the First Metro Line has on housing prices in the city of Bogotá. It combines innovative databases, web scraping and Google Maps with administrative records—such as the Colombian Institute of Urban Development (IDU) databases—to assess how the announcement of the metro station construction impacts the rental or sales market prices for the houses surrounding the infrastructure projects. The results show that the housing sale prices increased since October 2019, the time when the awarding of the contract for construction of the First Metro Line was announced. The flats and houses on sale located within 1.5 km from the future metro station showed an increase of 10.5 and 6.5% in prices, respectively.

Keywords: metro, Bogotá, housing, prices, expectations, assessment.

Corresponding author:
Juan Miguel Gallego
juan.gallego@urosario.edu.co
Facultad de Economía Universidad del Rosario, Bogotá, Colombia

1. Introduction

The relation between transport infrastructure and housing price variations is a crucial element of public policy. Particularly, capitalisation of the areas surrounding projects may occur even before the construction takes place (Agostini and Palmucci, 2008). Between the announcement of a project and its completion, there is a window of time during which community expectations play an important role in the following way. Once the construction details are disclosed, the community may move first seeking an early advantage (Yiu and Wong, 2005; Ruan and Yin, 2014). For instance, proximity to a transport station provides easy access to means of transport, thus housing prices are expected to increase after the transport project is completed. However, people can look for houses and flats in advance to take advantage of these benefits at a low cost. As demand grows, prices will be pushed up, and there may be significant increases after the announcement, but before construction. This scenario allows us to infer not only the importance of the construction impact on prices but also the impacts of the transport infrastructure project announcements (Chen et al., 2019).

Therefore, this paper seeks to understand how the announcement of the Bogotá metro station construction project may impact surrounding housing prices. This case study is relevant for a specific reason: the metro construction project is not new. Although the project was awarded on October 17, 2019, and the actual construction plan was approved, there had been several announcements in last 70 years promising the metro construction[1] and that may impact how prices react to latest and more reliable announcement. In other words, the previous price increases derived from unofficial announcements may diminish the expected effect. However, given the expectations that the latest announcement will, in fact, be followed by the project fulfilment, a substantial increase in real estate market prices may occur (Chau and Ng, 1998; Yan et al., 2012).

Moreover, according to the Colombian National Administrative Department of Statistics, Bogotá has a population of more than seven million, which makes it the most populated city in the country and in Latin America, and one of the cities with the highest need for transport solutions. According to the World Economic Forum (2020) and the INRIX's global traffic index (2020), Bogotá was in 2019 the city with the highest congestion score, where an average person lost around 191 hours of the year in traffic. Besides, the fact that Bogotá is a city with a substantial informal economy and a high incidence of criminality gives rise to the growing concern that building the metro may exacerbate those problems (Forouhar, 2016; Dai et al. 2016). Given that the proposed metro is elevated, different spaces underneath will be created and can worsen two aforementioned problems (Poister, 1996; Bowes y Ihlanfeldt, 2001; Ridgeway y MacDonald, 2017; Mulley et al., 2018). This will affect the well-being of the community surrounding the project and impact housing prices. The construction of the Bogotá metro is one of the most important investments in public transport

[1] The first proposal for the construction of the metro in Bogotá dates back to 1942. Several unofficial announcements and construction proposals have been made in different periods by Bogotá's City Council. Please find a press release summarizing the announcements and proposals in the following link: https://www.elespectador.com/bogota/por-que-mas-de-70-anos-sin-metro-en-bogota-article-547135/

infrastructure in the country. It includes 24 km of railways make it one the most larger metro lines in Latin American, which will help to covers the demand of 12.7 million of travels per day in Bogotá.

Therefore, this study analyses how the announcement of the metro construction impacts housing prices through certified databases using several innovative information sources such as real state agencies' websites, Google Maps and administrative records. This methodology allows analysing around 673,567 of house and flat prices before and after the metro announcement. A difference-in-differences (DID) model is used to estimate the differential increase in dwellings (houses and flats) prices according to their proximity to the metro station, following the October 2019 construction announcement. With this methodology, the price of dwellings closer within a radio of 500m, 1km and 1.5km were compared with similar dwellings far away from the project new metro lines. Our methodology considers well-known external and internal dwelling conditions such as crime, green areas, number of bedrooms, household income among others.

The results revealed a significant impact on housing prices after the construction of Bogotá metro was announced. The average flat sale price increased 5.2% and 10.5% for those flats located within 1km and 1.5 km, respectively. Similarly, houses for sale had a positive impact on prices, particularly for houses located within 1km and 1,5kms, with an increase of 7.1% and 6.5%, respectively. That being said, houses for sale had a greater impact than flats for sale. Using a wealth residency property index in Bogotá, this paper measures the heterogenous impact of the metro announcement. Higher impacts on prices are found in the lowest and middle wealthy dwellings.

This paper is structured as follows: the first section contains an introduction to the issue being addressed. The second section comprises the literature review, and it explains the different study methodologies related to infrastructure and housing to put forward some hypotheses about the variables that influence the model. The third section provides details regarding the different information sources utilised, including a description of the data. The fourth section shows the empirical approach, while the fifth and sixth sections show the results, with their heterogeneous effects and placebo tests. The last section states the conclusions.

2. Literature Review

The importance of building transport infrastructure lies is that it serves as a strategic economic development tool. This relationship between transport infrastructure and economic development has been widely studied. Mohammad et al. (2013) performed a meta-analysis of papers that developed empirical estimates regarding the impact of train construction on land value change. The authors conclude that there are multiple factors that cause changes in the value, including land use, type of train facilities, system maturity, distance to the stations, geographic location and street accessibility. For instance, Billings (2011) studied the impact of light rail infrastructure construction on housing prices. In this regard, the author's results conclude that light rail investments could be used more as an economic development tool than an amenity. Accordingly, Yang et al. (2016) suggested that the government establishes stations in less developed locations to generate high growth rates and in turn more benefits through high housing values.

Efthymiou and Antoniou (2013) stated that proximity to transport infrastructures may have a positive or negative impact depending on the type of transport system. In this respect, they found that in Athens the train station facilities had a negative impact on housing prices because of increased noise in the area. Debrezion et al. (2011) arrived at a similar outcome regarding the Dutch housing market. Generally, even if the effects are positive, depending on the context, negative externalities may arise that can impact housing prices, that is, pulling prices down instead of raising them. Hence, the conditions of each context vary and may cause effects in both directions.

Most literature refer specifically to how the construction of projects itself impacts prices. However, there is relatively little research regarding how the announcement of projects involving this type of infrastructure impacts the change in real estate prices. Zhang et al. (2016) sought to analyse how the rail infrastructure construction impacts real estate prices and took into account the impact of transport project construction expectations (project announcement). The authors concluded that expectations do not have a significant impact on prices. Similarly, Devaux et al. (2017) analysed the anticipation impact of a metro extension on residential values in Laval, Canada. The authors concluded that on average the announcement of the construction of metro stations had no significant effect on property values.

On the other hand, Agostini and Palmucci (2008) study how the announcement of metro project construction and the engineering project disclosure impacted property values in Santiago, Chile. The results revealed that flat prices increased once the project was announced and stations locations were revealed. Additionally, proximity to the announced stations entails higher increases in prices. Along the same line of research, authors such as Knaap et al. (2001), Golub et al. (2012) and Bae et al. (2003) considered that announcing and planning the location of stations increases the value of neighbouring houses. Cohen and Brown (2017) suggested that the announcement of a new rail access line positively impacts sale prices of commercial establishments near the project. Thus, given that it is unclear which conditions have a positive impact on prices of the real estate surrounding the projects, this turns out to be an opportunity to study this topic further—in this case, the Bogotá metro.

Nevertheless, the lack of available data on the housing market makes the analysis difficult. Additionally, the lack of data with constant periodicity or with a high level of disaggregation makes it difficult to study the impact of transport infrastructure on prices. Likewise, the available data does not always provide complete information on prices (Caplin et al., 2008), the available housing stock and real estate infrastructure (Cubeddu et al. 2012). The aforementioned limitations may hinder the proper assessment of the impact on the price changes.

In response, the growth of data science tools and the expansion of the Internet real estate market have enabled the use of non-conventional tools to compensate for said lack of information in the real estate market. Efthymiou and Antoniou (2013) used a web-scrapping tool that allows gathering data using regular expressions and data mining techniques (Munzert et al., 2014). Through these techniques, the authors collect data, such as prices and property characteristics, published on real estate webpages. In turn, this data is used to analyse the factors that impact the property selling price. In the Colombian case, Cárdenas Rubio et al. (2019) resorted to data science tools to collect data for the Colombian market. Using this information,

the authors were able to carry out a case study based on geographic location and check data consistency. The researchers concluded that they could rely on said database to analyse the real estate market in Bogotá and others major cities in Colombia.

3. Data

To carry out this study, we created a single database by compiling and combing information from different sources. The first source contains information on housing prices in Bogotá from January 2017 to January 2020. This database, which was updated every fifteen days, was created by scraping data from advertisements listed in different real estate websites and leasing portals. The housing costs are standardised at constant 2017 prices. Aside from prices and the type of property, that is, flat or house, data includes property characteristics such as stratification[2], total number of rooms and bathrooms, total area, construction year and address and a key variable to include information on conditions of the property's surroundings.

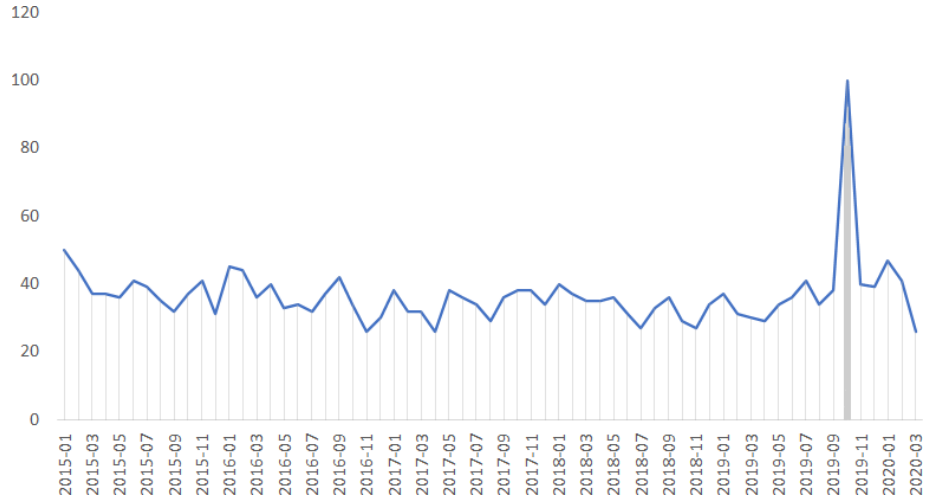
The second source of information stems from the Colombian Urban Development Institute (IDU), and it contains georeferenced characteristics of transport services, such as Transmilenio or bus stations; recreational or comfort areas, such as the number of restaurants or green areas; and total number of commercial establishments or public works[3].

This same database provides information regarding the future location of metro stations. Some studies (Gibbons y Machin, 2008; Bowes and Ihlanfeldt, 2001) have indicated that the perception of crime surrounding a property is also a crucial component of its value. Therefore, the crime rate in the property's neighbourhood—reported during the month of its publication on the sale database—must be added to the previously defined characteristics. The georeferenced crimes are retrieved from the National Police, based on data generated by the Monitoring Centres for Coexistence and Citizen Security. These characteristics are combined to the real estate sales information through their address.

[2] Socioeconomic stratification in Colombia is a classification in strata of residential property depending on the public services they receive and thus allow for differential charging. Besides, the said stratification enables the allocation of subsides and targeted contributions. In this way, the high-income users will pay more for public services, and this indirectly helps the population of lower income to pay their bills.

[3] The crossover between houses for sale and the number of Transmilenio stations is not perfect. However, only 5% of the houses could not be matched, and the estimates are robust, excluding this control.

The treatment variable is also created by using web-scraping tools that collect news regarding the metro, published between 2017 and 2020, on the most important online media[4]. Figure 1 shows the number of weekly news releases related to the metro construction or awarding. There is a clear peak during October 2019; thus, this month is identified as the most likely time for property owners to start updating their prices. The stated date matches the moment the First Metro Line was awarded.



Source: Google trends. Own calculations.

Figure 1: Google Searches Regarding the Metro in Bogotá

Table 1 shows descriptive statistics for flats for sale depending on the distance to the nearest future metro station. The further the property is away from a station, both the average building area and the average property value increase. For instance, the average price for flats increases from 473 million for flats within 500 metres to 672 million for all flats within 2 km. In the case of the average area, the change is between 135 m^2 and 142 $m^{2.}$. The same occurs with the distance to a Transmilenio station because the categories are built according to proximity to a metro station, which suggests that the two variables are correlated. The number of bathrooms, rooms and the number of restaurants, commercial facilities and bus stations are similar in the four categories. The number of public works and crime rates decrease when the group of flats is located farther away from the metro station.

Table 2 shows the distribution of the years of construction of flats for the same categories. It can be seen that further the distance, the more the composition of the flat complex is changing. For example, 29% of the flats located within 500 metres of the station are eight years old or less, while for those flats located more than 2 km away, this figure rises to 35%. The latter is in line with the expansion along the savannah that Bogotá has experienced in recent years.

[4] These media include *El Espectador*, *El Tiempo* or *Semana*.

Proximity	500 m		1 km		1.5 km		2 km	
	Yes	No	Yes	No	Yes	No	Yes	No
Per Square Metre Price	3,678	4,568	4,037	4,579	4,335	4,576	4,616	4,537
House Price*	473,219	586,534	494,795	589,328	583,679	584,142	672,306	567,512
Built Area	135	123	130	123	136	122	142	120
Number of Rooms	3	3	3	3	3	3	3	3
Bathrooms	2	3	2	3	3	3	3	3
Distance to the Nearest Station	177	6,115	377	6,316	642	6,655	889	6,944
Crime Rate	More than 40	34	More than 40	34	More than 40	34	38	34
Bus Stations	37	24	36	23	30	23	29	23
Restaurants	34	20	34	20	31	19	29	19
Commercial Establishments	32	30	34	30	34	30	34	29
Green Areas	36	30	35	30	34	30	34	30
Strata	4	4	4	4	4	4	4	4
Public Works	0.10	0.10	0.11	0.10	0.09	0.10	0.08	0.10
Notes	11,387	516,814	29,265	498,936	58,718	469,483	83,565	444,636

Source: Main Colombian online real estate portal. Own calculations.
*Note: The variable is in thousands in Colombian Pesos.

Table 1: Descriptive Statistics of Flats for Sale

Category According to Years	500 m		1 km		1.5 km		2 km	
	Yes	No	Yes	No	Yes	No	Yes	No
Less than 1 Year	0.07	0.07	0.08	0.07	0.08	0.07	0.09	0.07
1 to 8 Years	0.22	0.33	0.23	0.34	0.24	0.34	0.26	0.35
9 to 15 Years	0.15	0.21	0.15	0.22	0.16	0.22	0.15	0.22
16 to 30 Years	0.33	0.30	0.33	0.30	0.34	0.30	0.34	0.30
More than 30 Years	0.22	0.08	0.21	0.08	0.19	0.07	0.17	0.07

Source: Main Colombian online real estate portal. Own calculations.

Table 2: Descriptive Statistics based on Years of Flats for Sale

Table 3 shows the same characteristics as Table 1 but for houses for sale. However, for flats, the average area decreases as more distant houses are added, while the price per square metre increases. Likewise with flats, the amount of public works, restaurants and other features starts to decrease as the houses get further away from each other. Instead, for flats, no matter how far away the group of houses is, most of them were built 16 or more years ago. For houses within 500 metres, 71% are 16 years old or more, and for those within 2 km, the figure is 68% (see Table 4).

Proximity	500 m		1 km		1.5 km		2 km	
	Yes	No	Yes	No	Yes	No	Yes	No
Per Square Metre Price	2,642	3,290	2,717	3,315	2,794	3,340	2,752	3,374
House Price*	644,502	706,928	640,017	710,378	654,002	712,559	612,036	724,100
Built Area	250	210	243	209	236	208	226	209
Number of Rooms	6	4	6	4	6	4	6	4
Bathrooms	4	3	4	3	4	3	3	3
Distance to the Nearest Station	169	7,097	354	7,419	568	7,824	772	8,120
Crime Rate	36	29	37	29	35	28	35	28
Bus Stations	35	23	34	23	33	22	33	22
Restaurants	32	20	32	19	31	18	30	18
Commercial Establishments	28	24	30	24	29	24	28	24
Green Areas	37	28	36	28	35	27	34	27
Strata	3	4	3	4	3	4	3	4
Public Works	0.04	0.05	0.03	0.06	0.02	0.06	0.02	0.06
Notes	5,274	140,092	11,808	133,558	19,602	125,764	25,213	120,153

Source: Main Colombian online real estate portal. Own calculations

*Note: The variable is in thousands in Colombian Pesos.

Table 3: Descriptive Statistics of Houses for Sale

Category According to Years	500 m		1 km		1.5 km		2 km	
	Yes	No	Yes	No	Yes	No	Yes	No
Less than 1 Year	0.05	0.02	0.03	0.02	0.03	0.02	0.03	0.02
1 to 8 Years	0.10	0.13	0.10	0.13	0.10	0.13	0.10	0.13
9 to 15 Years	0.14	0.28	0.18	0.28	0.19	0.29	0.20	0.29
16 to 30 Years	0.34	0.36	0.35	0.36	0.34	0.36	0.36	0.36
More than 30 Years	0.37	0.22	0.34	0.21	0.33	0.20	0.32	0.20

Source: Main Colombian online real estate portal. Own calculations.

Table 4: Descriptive Statistics Based on Years of Houses for Sale

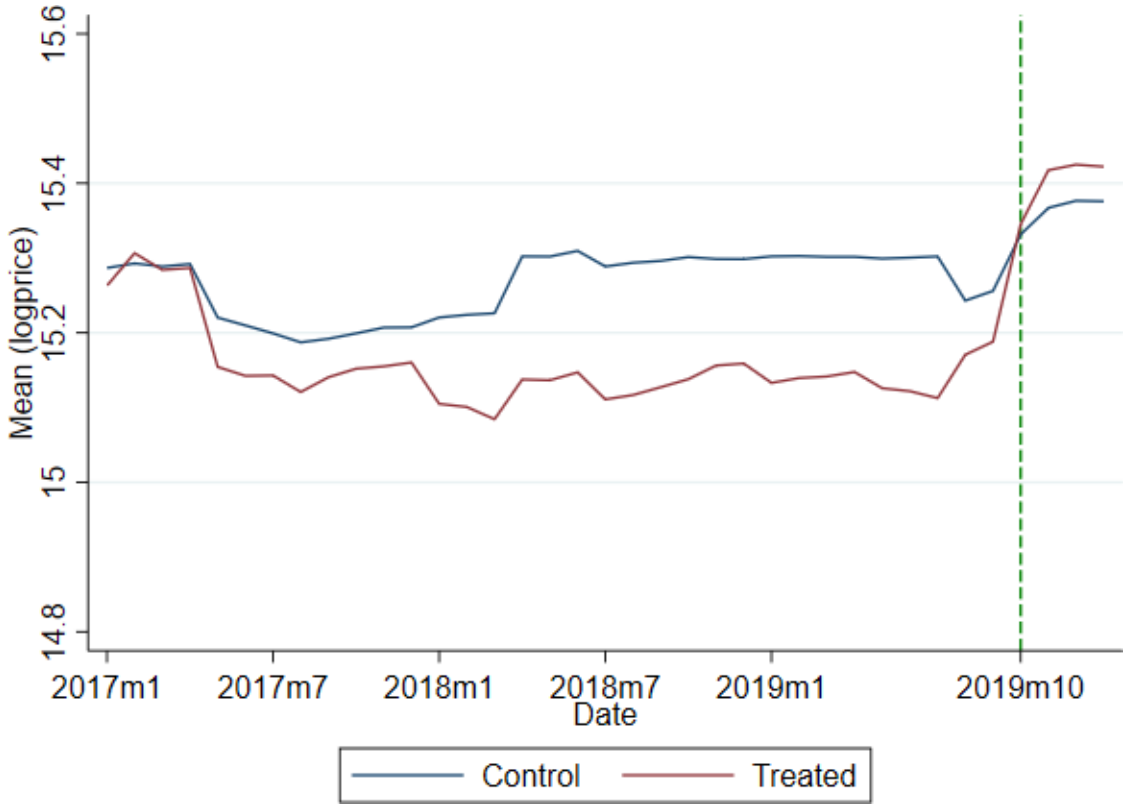
4. Empirical Approach

To estimate the impact of metro station construction on housing prices, we use the variation in the distance over time between the properties collected regarding each month and the nearest station (see Appendix A, Figure A.1). The level of impact would depend on this distance because houses located far enough will not be affected by the noise caused by the metro nor will be favoured by the enhanced connection to the city's transport systems. Therefore, the treatment group may be defined by those houses located within a specific radius from the future station, and the control group would be determined by those houses out of that radius. The empirical approach comprises a DID model used to estimate the differential increase in prices of houses near to the metro station in comparison to those that are not, after the project had wide media coverage, when probably people learned about the location of these stations. This method will take into account any difference existing before October 19 between the houses of the treatment group and those in the control group. The specification to be estimated follows the equation mentioned below:

$$Log(Price_{icmt}) = \alpha_c + \delta_{mt} + \theta(Proximity_i * Post_{mt}) + \beta^j \boldsymbol{X_i} + \gamma^j \boldsymbol{Z_{cmt}} + \varepsilon_{icmt} \quad (1)$$

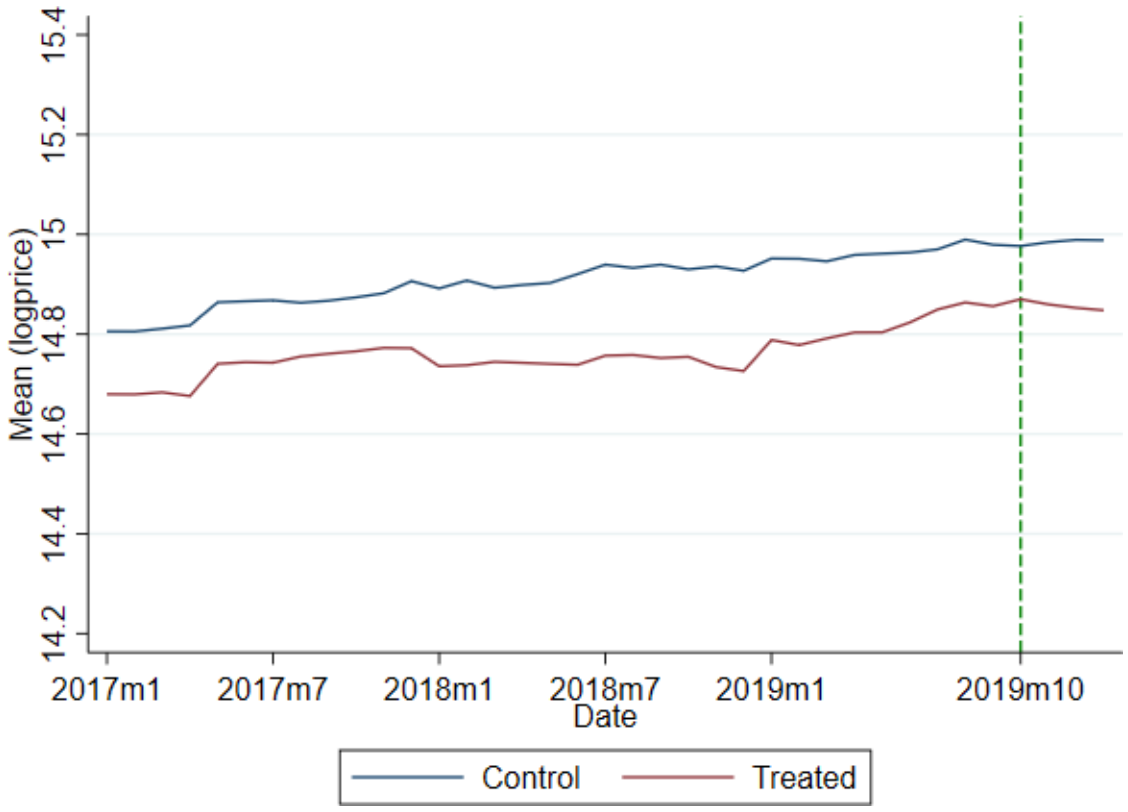
where $Log(Price_{icmt})$ is the logarithm of the price of the house i of the neighbourhood c in the month m of the year t. The estimation for flats and houses will be studied separately as they belong to different markets because of their physical characteristics and type of buyer. *Proximity* is a dummy variable that takes the value of either 1 if the house i is within the specific radius of the future metro station or 0 if the case is opposite. Considering that the exact distance is not perfectly known, we explore the possible value to estimate the specification of a radius of 500 metres, 1, 1.5 and 2 km on a separate basis The variable $Post_{mt}$ is a dummy variable that takes the value 1 for dates after October 2019 and 0 for previous dates. The interest coefficient θ captures the differential evolution of price levels, for both houses that may potentially be affected to a certain extent and houses that will not be affected after the information spike about the new metro. The terms α_c and δ_{mt} are neighbourhood and month–year fixed effects to control time-invariant neighbourhood-level characteristics and price shocks that all houses may suffer during a specific period.

Finally, vector $\boldsymbol{X_i}$ includes property characteristics, such as the number of rooms or bathrooms, location, stratus and distance, to the metro station. Additionally, vector $\boldsymbol{Z_{cmt}}$ is included, which contains characteristics of the neighbourhood *c* in month *m* of year *t*, such as the crime rate, Transmilenio and bus stations, restaurants, commercial facilities, parks and public works. The main assumption used to identify the impact is that—in absence of the increase in the metro information flow—the price level would have followed a trend similar to the one presented before October 2019, regardless of whether the house is located near the metro station or not. This assumption might be violated in case the unobserved selection of the metro station location is correlated with housing characteristics. Figures 2 and 3 show the differential price evolution for treatment and control groups, taking into account a 1 km radius, for flats and houses, respectively. The parallel trend assumption, at least visually, seems to be met for both property types as they show a similar movement prior to October 2019.



Source: Main Colombian online real estate portal. Authors' own calculations.
Note: The distance is within a 1.5 km radius.

Figure 2: Average Prices of Flats for Sale



Source: Main Colombian online real estate portal. Authors' own calculations.
Note: The distance is within a 1.5 km radius.

Figure 3: Average Prices of Houses for Sale

5. Results

Table 5 shows the results of the estimation explained in the above section, only with respect to flats for sale. Each column defines 'proximity' according to the heading at the top of the column. Thus, for Column 1, the variable takes the value of 1 if the property is located within 500 metres of the nearest metro station or 0 if it is outside of that radius. Moreover, all the specifications include the control vector for housing and neighbourhood and the month–year and neighbourhood fixed effects. The information spike related to metro construction had a positive impact on all model specifications. As Equation (1) estimates a log-linear model, the coefficient interpretation is derived by multiplying them by 100. The impact of the information spike causes an increase of 6.0% for flats located within a radius of 500 metres, while for those flats within 1 km, the impact is 5.2% These results suggest that the impact is not linear with respect to distance because for it to exist, the flat should not be located more than 2 km away.

	(1)	(2)	(3)	(4)
	500 m	1 km	1.5 km	2 km
VARIABLES	ln (price)	ln (price)	ln (price)	ln (price)
Impact	0.060***	0.052***	0.105***	0.124***
	(0.010)	(0.006)	(0.004)	(0.003)
Housing Characteristics	Yes	Yes	Yes	Yes
Neighbourhood Characteristics	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Neighbourhood Fixed Effects	Yes	Yes	Yes	Yes
Observations	528,201	528,201	528,201	528,201
R^2	0.711	0.670	0.711	0.713

Note: Robust standard errors (in parentheses); ***p < 0.01, **p < 0.05, *p < 0.1

Table 5: Impact of the Metro Announcement on Prices of Flats for Sale

Table 6 shows the estimation results with respect to houses for sale. The publication of the metro awarding has an important and positive impact on houses located in a radius of 500 metres and of 1, 1.5 and 2 km from a future metro station. As in Table 5, the coefficient showing a 7.1% price increase in real estate located within 1 km is the largest of all specifications. As for flats, the impact is not linear with respect to the distance from the house. Finally, the impact is greater on flats than on houses because, with the exception of the 1 km distance definition, all impacts on flats have a greater magnitude than on houses[5].

	(1)	(2)	(3)	(4)
	500 m	1 km	1.5 km	2 km
VARIABLES	ln (price)	ln (price)	ln (price)	Ln (price)
Impact	0.057**	0.071***	0.065***	0.053***
	(0.024)	(0.015)	(0.012)	(0.010)
Housing Characteristics	Yes	Yes	Yes	Yes
Neighbourhood Characteristics	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Neighbourhood Fixed Effects	Yes	Yes	Yes	Yes

[5] This same procedure was carried out to estimate the impact of future metro stations on the rental of flats and houses for the same period. In theory, the value of real estate should increase as the construction has a long-term impact, while rental contracts, which are temporary in nature, should not vary significantly. Tables A.1 and A.2 show the results of estimating Equation (1) using rental prices of flats and houses as dependent variables. It is found that, for flats and houses, the coefficients are not significant under any of the treatment group definitions.

Observations	145,366	145,366	145,366	145,366
R^2	0.635	0.635	0.635	0.635

Note: Robust standard errors (in parentheses); ***p < 0.01, **p < 0.05, *p < 0.1

Table 6: Impact of the Metro Announcement on Prices of Houses for Sale

5.1 Heterogeneous effects

The impact of proximity to a means of transport on housing is not identical across all types of properties. There are grounds supporting that households belonging to a lower strata could potentially be more likely to prefer such proximity due to the high dependence of individuals with no access to private means of transport on a public transport network. Consequently, Equation (1) is estimated for different samples, which are in turn divided in three categories: strata 1 and 2; strata 3 and 4; and strata 5 and 6. Table 7 shows the estimation results for the different treatment groups, types of houses and includes all control groups.

Once the treatment for housing is defined as 'being within a 500 metre or 1 kilometre radius', the estimated values are positive and highly statistically significant for the three samples of flats for sale. This confirms the reliability of this definition, since, the higher the stratum where the property is located, the lower the price increase, going from 14.2 % to 7.0 % (500 metre radius) and from 13.4 % to 5.9 % (1 km radius). Hence, the evidence suggests that the houses located in the lower strata have benefited more than those located in the higher strata. However, we should also mention that the houses located in the higher strata benefit more as they move farther away from a metro station, for instance, the houses located in strata 5 and 6 nearer to the metro (500 mt) have a 7.0% impact on their price while those that are farther away (2 km) have a 12.7% impact.

		500 m	1 km	1.5 km	2 km
		ln (price)	ln (price)	ln (price)	ln (price)
Impact on Flats for Sale	Strata 1-2	0.142*** (0.030)	0.134*** (0.019)	0.107*** (0.015)	0.139*** (0.012)
	Strata 3-4	0.018 (0.011)	0.008 (0.007)	0.088*** (0.005)	0.115*** (0.005)
	Strata 5-6	0.070*** (0.022)	0.059*** (0.011)	0.109*** (0.005)	0.127*** (0.004)
Housing Characteristics		Yes	Yes	Yes	Yes
Neighbourhood Characteristics		Yes	Yes	Yes	Yes
Time Fixed Effects		Yes	Yes	Yes	Yes
Neighbourhood Fixed Effects		Yes	Yes	Yes	Yes

Note: Robust standard errors (in parentheses); ***p < 0.01, **< 0.05, *p < 0.1
Number of remarks flat strata: 1-2: 32,529; 3-4: 260,595; 5-6: 235,077
Number of remarks houses strata: 1-2: 25,700; 3-4: 83,696; 5-6: 35,970

Table 7: Heterogeneous Impact by Strata

6. Placebo Tests

To provide further validity to the above findings, two placebo exercises were run for flats and houses for sale. The exercise involved estimating Equation (1) but changing the date we considered as the metro's news broadcasting spike. The news impact should be under this spike and not on any other date because the number of citizens informed about the location of stations would be much lower and therefore would not update their real estate prices. Thus, variable *post* takes the value of 1 from June 2019, when the International Public Tender for the construction of Bogotá's First Metro Line was opened. Table 8 shows no significant or relatively little (negative) impact on prices of flats in any specification if the treatment date is redefined to the one that precedes the spike perceived in Figure 1. This same exercise was conducted for houses for sale, the results obtained were statistically similar to the flat placebo test.

	(1)	(2)	(3)	(4)
	500 m	1 km	1.5 km	2 km
VARIABLES	ln (price)	ln (price)	ln (price)	ln (price)
Impact	-0.014	-0.028***	-0.009*	-0.007*
	(0.011)	(0.007)	(0.005)	(0.004)
Housing Characteristics	Yes	Yes	Yes	Yes
Neighbourhood Characteristics	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Neighbourhood Fixed Effects	Yes	Yes	Yes	Yes
Observations	479,056	479,056	479,056	479,056
R^2	0.710	0.671	0.710	0.711

Note: Robust standard errors (in parentheses); ***p < 0.01, **p < 0.05, *p < 0.1

Other dates of interest were considered for the placebo test, and the results were statistically similar

Table 8: Placebo Tests on Flats for Sale

7. Conclusions

This paper leverages the increase in the information flow regarding the metro construction as a natural experiment to assess how enhanced access to the public transport network impacts on the surrounding housing price levels. The construction of the metro is probably the most important public investment project in Bogotá since the Transmilenio transport service started running. Thus, it may be one of the most important determinants for the future assessment by the main Colombian online real estate portal.

Through web-scraping techniques, we collected data on property prices from January 2017 to January 2020 for several websites. This data also includes information regarding property characteristics. The degree of information that citizens are able to access about the metro is measured according to the trend in the amount of news posted on the leading news websites regarding said projects. The main outcome is that, after October 2019, the prices of flats and houses located within a 1.5 km radius from the future metro station experienced an increase of 10.5% and 6.5 %, respectively. Besides, the results suggest that the impact is somewhat heterogeneous because of the fact that flats for sale in the lower strata experience a greater impact on price increases compared with those in the upper strata. Such heterogeneity can also be found within the flats located in the higher strata because the farther they are from a future metro station, the more they benefit.

In contrast, the results for rentals are not statistically significant (see Appendix A, Tables A.1 and A.2). We argue that given that the metro construction will be completed after mid-2025, rentals should not increase as the prices are settled temporarily and not in the long term, as is the case with property prices.

To show that the results are robust, the same regression is estimated; however, now, the variable indicating

the date after the treatment is changed to October 2018 and June 2019, when the International Public Tender for the construction of the First Metro Line in Bogotá was opened. The coefficients are not significant, and this confirms that it is in October 2019 when the information flow about the metro construction actually increased.

Finally, these findings are essential for public policy makers, especially considering that the City Council of Bogotá could design a new and more complex property tax structure and that by taking into account the positive externalities derived from the metro construction it may generate certain tax relief or incentive measures intended for those individuals that were not reached by the positive impact, or the other way round—a heavier burden for those who are favoured by such externalities.

Acknowledgements
The authors would like to thank the Scientific Colombia Program – EFI Alliance for their financial support to the productive and social inclusion program: programs and policies for the promotion of a formal economy, code 60185, which are part of the EFI Alliance – Formal and Inclusive Economy, under the Contingent Recovery Agreement No. FP44842-220-2018. Furthermore, the authors are very grateful for Jonathan Sarango Iturralde's valuable assistance throughout this paper.

Funding
This work was supported by the Scientific Colombia Program – EFI Alliance: programs and policies for the promotion of a formal economy, code 60185, which are part of the EFI Alliance – Formal and Inclusive Economy, under the Contingent Recovery Agreement No. FP44842-220-2018.

References

Agostini, C. A. and Palmucci, G. A. (2008). The anticipated capitalisation effect of a new metro line on housing prices. *Fiscal studies*, 29(2):233–256.

Bae, C. H. C., Jun, M. J., & Park, H. (2003). The impact of Seoul's subway Line 5 on residential property values. *Transport policy*, 10(2), 85-94.

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

Bowes, D. R. and Ihlanfeldt, K. R. (2001). Identifying the impacts of rail transit stations on residential property values. *Journal of Urban Economics*, 50(1):1–25.

Caplin, A., Chopra, S., Leahy, J. V., LeCun, Y., and Thampy, T. (2008). Machine learning and the spatial structure of house prices and housing returns. *Available at SSRN 1316046*.

Cárdenas Rubio, J. A., Chaux Guzmán, F. J., and Otero, J. (2019). A database with prices and characteristics of the housing in Colombia created with information gathered from the Internet. *Revista de Economía del Rosario*, 22(1):25.

Chau, K. and Ng, F. (1998). The effects of improvement in public transportation capacity on

residential price gradient in hong kong. *Journal of property valuation and investment*.

Chen, Y., Yazdani, M., Mojtahedi, M., and Newton, S. (2019). The impact on neigh- bourhood residential property valuations of a newly proposed public transport pro- ject: The Sydney northwest metro case study. *Transportation Research Interdisciplinary Perspectives,* 3:100070.

Cohen, J. P. and Brown, M. (2017). Does a new rail rapid transit line announcement affect various commercial property prices differently? *Regional Science and Urban Economics*, 66:74–90.

Cubeddu, L., Tovar Mora, C. E., and Tsounta, E. (2012). Latin America: Vulnerabilities under construction? *International Monetary Fund*, Washington, DC.

Dai, X., Bai, X., and Xu, M. (2016). The influence of Beijing rail transfer stations on surrounding housing prices. *Habitat International*, 55:79–88.

Debrezion, G., Pels, E., and Rietveld, P. (2011). The impact of rail transport on real estate prices: an empirical analysis of the Dutch housing market. *Urban Studies*, 48(5):997–1015.

Devaux, N., Dubé, J., & Apparicio, P. (2017). Anticipation and post-construction impact of a metro extension on residential values: The case of Laval (Canada), 1995–2013. *Journal of Transport Geography*, 62, 8-19.

Efthymiou, D. and Antoniou, C. (2013). How do transport infrastructure and policies affect house prices and rents? evidence from Athens, Greece. *Transportation Research Part A: Policy and Practice*, 52:1–22.

Forouhar, A. (2016). Estimating the impact of metro rail stations on residential property values: evidence from Tehran. *Public Transport*, 8(3), 427-451.

Gibbons, S. and Machin, S. (2008). Valuing school quality, better transport, and lower crime: evidence from house prices. *Oxford Review of Economic Policy*, 24(1):99–119.

Golub, A., Guhathakurta, S., & Sollapuram, B. (2012). Spatial and temporal capitalization effects of light rail in Phoenix: From conception, planning, and construction to operation. Journal of Planning Education and Research, 32(4), 415-429.

INRIX (2020). Global Traffic Scorecard. Accessed September 29, 2020. https://inrix.com/scorecard/

Knaap, G. J., Ding, C., and Hopkins, L. D. (2001). Do plans matter? the effects oflight rail plans on land values in station areas. *Journal of Planning Education and Research*, 21(1):32–39.

Mohammad, S. I., Graham, D. J., Melo, P. C., and 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.

Mulley, C., Tsai, C.-H. P., and Ma, L. (2018). Does residential property price benefit from light rail in Sydney? *Research in Transportation Economics*, 67:3–10.

Munzert, S., Rubba, C., Meißner, P., and Nyhuis, D. (2014). *Automated data collection with R: A practical guide to web scraping and text mining*. John Wiley & Sons.

Poister, T. H. (1996). Transit-related crime in suburban areas. *Journal of Urban Affairs*,18(1):63–75.

Ridgeway, G. and MacDonald, J. M. (2017). Effect of rail transit on crime: a study of Los Angeles from 1988 to 2014. *Journal of quantitative criminology*, 33(2):277–291.

Ruan, L. F. and Yin, J. P. (2014). Impact of urban rail transit on residential property values in Hangzhou. In *Applied Mechanics and Materials*, volume 580, pages 1306– 1310. Trans Tech Publ.

World Economic Forum (2020). The countries with the worst traffic congestion - and ways to reduce it. Accessed September 29, 2020. https://www.weforum.org/agenda/2020/07/cities-congestion-brazil-colombia-united-kingdom/

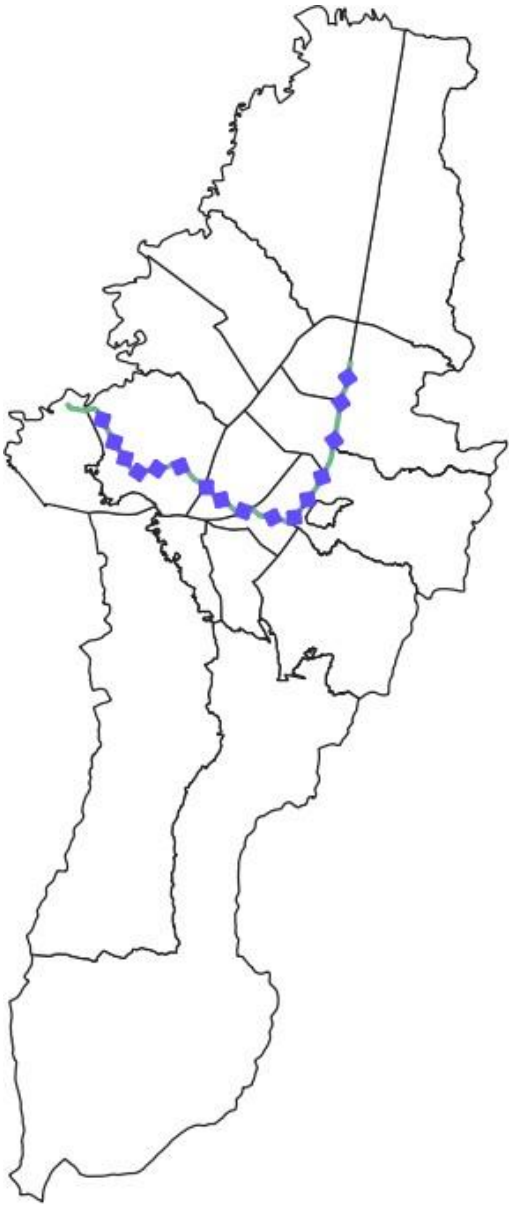
Yan, S., Delmelle, E., and Duncan, M. (2012). The impact of a new light rail system on single-family property values in charlotte, North Carolina. *Journal of Transport and Land Use*, 5(2):60–67.

Yang, J., Chen, J., Le, X., and Zhang, Q. (2016). Density-oriented versus development- oriented transit investment: Decoding metro station location selection in Shenzhen. *Transport Policy*, 51:93–102.

Yiu, C. Y. and Wong, S. K. (2005). The effects of expected transport improvements on housing prices. *Urban studies*, 42(1):113–125.

Zhang, X., Liu, X., Hang, J., Yao, D., and Shi, G. (2016). Do urban rail transit facilities affect housing prices? Evidence from China. *Sustainability*, 8(4):380.

Appendix A: Additional figures and tables



Source: Colombian Institute of Urban Development (IDU).
Created by the authors

Figure A.1: Metro Stations

	(1)	(2)	(3)	(4)
	500 m	1 km	1.5 km	2 km
VARIABLES	ln (price)	ln (price)	ln (price)	ln (price)
Impact	0.013	-0.005	-0.004	-0.007
	(0.014)	(0.008)	(0.006)	(0.005)
Housing Characteristics	Yes	Yes	Yes	Yes
Neighbourhood Characteristics	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Neighbourhood Fixed Effects	Yes	Yes	Yes	Yes
Observations	173,671	173,671	173,671	173,671
R^2	0.643	0.641	0.641	0.641

Note: Robust standard errors (in parentheses); ***p < 0.01, **p < 0.05, *p < 0.1

Table A.1: Impact of the Metro Announcement on Prices of Flats for Rent

	(1)	(2)	(3)	(4)
	500 m	1 km	1.5 km	2 km
VARIABLES	ln (price)	ln (price)	ln (price)	ln (price)
Impact	0.017	0.004	-0.029	-0.037
	(0.038)	(0.030)	(0.028)	(0.025)
Housing Characteristics	Yes	Yes	Yes	Yes
Neighbourhood Characteristics	Yes	Yes	Yes	Yes
Time Fixed Effects	Yes	Yes	Yes	Yes
Neighbourhood Fixed Effects	Yes	Yes	Yes	Yes
Observations	23,959	23,959	23,959	23,959
R^2	0.505	0.505	0.505	0.506

Note: Robust standard errors (in parentheses); ***p < 0.01, **p < 0.05, *p < 0.1

Table A.2: Impact of the Metro Announcement on Prices of Houses for Rent