The impact of bike-share on real-estate transaction prices

Vincent Thorne*

v0.1
July 2023

Click here for the latest version

Abstract

Bike-share programs have been introduced in more than two thousand cities around the world, but little is know about their impact on cities. Given their potential to act as an local amenity (providing new transport options to commuters), agents may increase their valuation property units nearby bike-share stations. In this paper, I test this hypothesis in New York City using the universe of real-estate transactions. I find that transactions within 150 meters of a bike-share station are sold at prices up to 5.8% higher than properties between 150 and 500 meters of the same station. This result suggests that bike-share is valued by urban dwellers and that it initiated important value creation.

Keywords: Real estate, Urban transportation, Cycling

*Department of Economics, Trinity College, Dublin · thornev@tcd.ie

1 Introduction

In the past thirty years, cities around the world have invested significantly in policies to encourage cycling. Praised for its potential to reduce traffic congestion, relieve pressure on public transport and curtail air pollution, the implementation of cycling policies has been widespread. On the one hand, cities have improved and expanded their cycling infrastructure (cycling paths, bike parking, etc), making cycling safer and more convenient. On the other hand, they have also sought to make cycling more accessible by providing public bike-share schemes to their inhabitants. More than two thousand such programs now exist around the world, providing an estimated 66 million trips in North America alone in 2021.

The advent of an affordable and practical cycling option thanks to bike-share programs has the potential to significantly change the transport landscape of cities. From previous research, we know that these changes in mobility options may have important impacts on commuting patterns, neighbourhood composition, and real-estate prices [citations]. Despite their popularity in the past two decades, there is a notable scarcity of research on the impact that bike-share programs may have on cities.

In this paper, I test whether the initial roll-out of North America's largest bike-share scheme in New York City in May 2013 had an impact on real-estate transaction prices. I geocode the universe of real-estate transactions from the New York City Department of Finance and match each of them to a bike-share station opened in the first wave (i.e., between May 2013 and June 2015). I assign transactions within 150 meters of a bike-share station to the treatment group (or ring), and transactions between 150 and 500 meters to the control group. I use a two-ring difference-in-differences approach, comparing transactions in the treatment ring with those in the control ring, before and after the opening of a bike-share station, and including station fixed effects. I find that transaction prices in the treatment ring increased by 5.8% after bike-share was implemented compared to transactions in the control ring.

This study contributes to a large body of research documenting the effects of transportation on real-estate prices. For example, [citations]. A recent paper by Shr et al. (2023) examines the short-term impact of a bike-share program expansion on rental prices. The extension of a bike-share system in Taiwan increased rental prices by 2% for rental units within 150 meters of bike-share stations, six months after implementation. These results suggest that bike-share programs have a positive impact on real-estate prices for units in their vicinity, but say little about their potential long-term effects. My paper adds to our understanding of these effects by estimating the impact of bike-share implementation and roll-out over the medium run (2.5 years) and for transaction (sale) prices, which is important for at least two reasons. First, transport habits may be sticky and individuals may take time to adjust to new transport options: as such, one might expect that commuting changes brought by cycling policies take time to materialise into real-estate transaction prices, motivating a long-term analysis. Second, rental and sale markets might differ in how they respond to changes in cycling policies.

These results are important in several regards. First, they show that cycling policies behave similarly to other transportation policies: they have an impact on real estate markets, even in a city where only about 1% of workers commute to work by bicycle. Second, it highlights the importance of taking into account the distributional impacts of transport policies, as they may change neighbourhood attractiveness and eventually price out some segments of the population. Third, the increase in property value due to bike-share documented here may be partially captured by the city through property taxes, which could go towards financing these investments, and improve policymakers' and voters' support for them (Gupta et al., 2022).

2 Data

This section describes the data used in the paper. I begin by describing the outcome data, which originates from the NYC transaction records, the primary operations performed on the data, and the sample creation process. I then transition to describing the bike-share data, and how I define control and treatment units.

2.1 Real-estate transaction records

To assess the impact of bike share on real-estate prices, I utilise the universe of transaction records collected by the New York City Department of Finance (NYCDF). These records report key variables such as sale price, surface area, tax lot identifiers, and building classes for every real estate transaction in NYC since 2003. I have gathered transaction records spanning from January 2011 to April 2015, and executed a series of cleaning and transformation procedures to ensure the data is ready for analysis.

First, I geocode each transaction using tax lot identifiers. The NYCDF maintains its own property identifier, uniquely locating every piece of property in the city. Using their Digital Tax Map, which associates each property identifier with a geographical polygon, I determine the location of a transaction as the centroid of the polygon.[1]

As mentioned above, I concentrate on the first wave of bike-share implementation in NYC (i.e., stations opened between May 2013 and April 2014) to keep the analysis manageable. The estimation strategy is based on a two-ring approach (see the [identification strategy] section), using transactions up to 500 meters away from first-wave bike-share stations, and I retain transactions within that range (significantly reducing computational complexity). Approximately 80 thousand transactions fall within that spatial-temporal range — I outline the precise construction of the treatment variable in subsection 2.3.

[1] Most polygons are identified using a borough + tax block + tax lot identifier. Condos and other communal property types are uniquely identified at the borough + tax block level only, and their location is therefore the centre of the tax block.

Next, I retain only sales with non-zero prices and non-missing surface areas, deflate the sale price to December 2015 levels, and compute the surface area per unit.[2] I identify price outliers, first using the definition in (Gupta et al., 2022) (sale prices greater than 400 thousand dollars and less than ten million dollars), but also price-per-square-foot outliers (greater than $50 and less than 20 thousand dollars).[3] Finally, I extract building attributes (residential/commercial, elevator, condo, etc: see subsection 2.4 for the complete list of attributes) using the building class category reported for each transaction. I also compute distances to main amenities for each transaction: distance to the nearest subway entrance and bus stop,[4] and distance to the nearest park.[5]

The analysis presented later will concentrate on residential units excluding price and price-per-square-foot outliers, but results including outliers and commercial units are reported in the appendix.

2.2 Bike-share data

This paper estimates the impact of the first wave of bike-share stations on real-estate prices. I use the universe of bike-share trips[6] to identify the opening (and occasionally closing) date of each station. The first wave is defined by the first spatial extent of the system, i.e. the initial area of the city that the bike-share system covered. The bike-share system in NYC was launched in May 2013 and the subsequent spatial expansion took place in July 2015. During that period, stations opened in three different months: May and June 2013, and March 2014. Upon visual inspection, it appears that some stations were closed and others opened within the same calendar month, and very near to each other. Since the estimation strategy relies on bike-share station fixed effects (see the section 3), it is critical to identify the correct set of stations, so I match those that opened and closed within a month and within 50 meters of each other as the same station.

2.3 Treatment construction

As detailed later in the paper, the estimation strategy compares real-estate transactions close to a bike-share station (within 150 meters) to those further away (between 150

[2]Surface area is given for the whole building, not the transacted unit: I take the surface area per building unit in order to correctly measure the size of a transacted unit.

[3]The later outlier definition is based on anecdotal evidence of top and bottom prices in NYC (see for example `https://therealdeal.com/new-york/2018/03/21/these-are-nycs-most-expensive-homes-by-price-per-square-foot/` and `https://www.nyrentownsell.com/blog/best-price-per-sq-feet-in-ny-to-rent-and-buy/`, both accessed 2023-07-13). The range is wide by design, as its goal is to remove to most unlikely prices per square foot.

[4]Historical data on subway entrances and bus stops locations is collected by Baruch College at the City University of New York, and freely available at `https://www.baruch.cuny.edu/confluence/display/geoportal/NYC+Mass+Transit+Spatial+Layers+Archive` (accessed 2023-07-13).

[5]The current location of parks is provided by the New York City Parks at `https://nycopendata.socrata.com/Recreation/Parks-Properties/enfh-gkve` (accessed 2023-07-13). I was not able to find historical locations of parks — the assumption is that the location of new parks is not correlated with the deployment of bike-share stations, which is plausible.

[6]Available on the bike-share provider's website: `https://citibikenyc.com/system-data` (accessed 2023-07-13).

and 500 meters from the station), before and after the opening of the station. Each transaction thus has to be matched with one (or more, depending on the case) bike-share station. In this subsection, I detail the steps and decisions made in matching transactions to stations.

I start by computing, for each transaction, all the bike-share stations within 500 meters. Each transaction may be matched to multiple stations: indeed, it is not uncommon for a transaction to fall within multiple 500-meter rings around bike-share stations, with a median of ten stations matched and a maximum of 17. At this stage, every row in the dataset is a transaction-station pair, with as many rows per transaction as it matches bike-share stations, and includes a measure of distance to the station (between zero and 500 meters by construction). However, not all matches are valid for estimation: indeed, a transaction cannot be treated by one station and act as a control for another. I therefore select station matches according to the following algorithm:

Case 1 The transaction matches only one station: keep that match.

Case 2.1 The transaction matches multiple stations, and all are further than 150 meters (i.e., the transaction is always a control): keep all matches. This allows the transaction to act as a control for multiple bike-share stations.

Case 2.2 The transaction matches multiple stations, and all are within 150 meters (i.e., the transaction is always treated): keep the earliest and closest matched station (in that order). If the transaction is impacted by treatment, the first station within 150 meters likely started affecting its value first, and it is probable (if the stations within 150 meters opened at the same time) that the closest one has the most impact.

Case 2.3 The transaction matches multiple stations, some of them within 150 meters, some of them between 150 and 500 meters: keep the station (1) within 150 meters, (2) opened earliest, and (3) closest (in that order). Once treated, a station should not be considered a control (it would violate SUTVA), and the station matches between 150 and 500 meters are discarded. If there are multiple stations within 150 meters, the same criteria as in Case 2.2 are used.

By allowing for multiple matches as described above (and after cleaning transactions as outlined earlier in subsection 2.1), I end up with about 20 thousand transaction-station pairs. Transactions which are within 150 meters of their matched stations are coded as within the treatment ring (those who are between 150 and 500 meters are coded as belonging to the control ring), and transactions taking place after the opening of their matched station are coded as post-period (those before as pre-period). As detailed below, the treatment effect will be identified by the interaction between the treatment ring and post-period dummies.

2.4 Descriptive statistics

Here I report summary statistics for the variables used in the estimation, as well as balance tables by treated vs control rings before the opening of bike-share.

Table 1: Summary statistics, numeric variables

	Mean	SD	Min	Median	Max	Miss
Sale price (2015 $)	3,138,831	2,419,373	400,920.9	2,430,962	1e+07	0
Log sale price (2015 $)	14.63	0.86	12.9	14.7	16.12	0
Sale price per sqft (2015 $)	3,501.5	3,692.28	112.86	1,955.51	19,761.28	0
Residential units (count)	8.42	45.2	0	3	1,681	0
Commercial units (count)	0.42	0.85	0	0	15	0
Total units (count)	8.84	45.4	0	3	1,684	0
Built surface (sqft)	7,425.46	35,257.67	0	3,680	1,231,250	0
Land surface (sqft)	2,463.39	7,666.63	0	2,000	298,550	0
Final surface (sqft)	7,464.32	35,259.12	680	3,712	1,231,250	0
Surface per unit (sqft)	1,258.99	903.44	191.62	1,003	9,155	0
Building age	99.51	27.3	0	110	217	3
Year built	1,913.32	27.28	1,798	1,901	2,015	3
Distance to bus stop (m)	100.45	61.07	7.57	91.28	479.44	0
Distance to subway entrance (m)	293.28	166.4	12.08	264.06	1,073.82	0
Distance to bike-share station (m)	307.13	129.75	2.09	324.61	499.93	0
Distance to park (m)	347.12	222.98	2.51	303.73	1,137.15	0
Sale quarter	9.78	4.95	1	10	18	0

Table 2: Balance table treated vs control ring, numeric variables, pre-treatment period

	Control ring 0 (N=3871)		Treated ring 1 (N=669)			
	Mean	Std. Dev.	Mean	Std. Dev.	Diff. in Means	p-value
Sale price (2015 $)	2,971,322.44	2,348,748.57	3,309,722.13	2,407,715.54	338,399.69***	0.00
Log sale price (2015 $)	14.56	0.87	14.70	0.85	0.14***	0.00
Sale price per sqft (2015 $)	3,484.51	3,628.14	3,759.74	3,955.44	275.23*	0.09
Residential units (count)	11.10	64.93	8.12	24.06	-2.97**	0.03
Commercial units (count)	0.47	0.90	0.55	1.05	0.08*	0.08
Total units (count)	11.57	65.14	8.67	24.69	-2.90**	0.04
Built surface (sqft)	9,369.99	49,368.65	7,619.01	24,514.66	-1,750.98	0.16
Land surface (sqft)	2,831.98	11,061.97	2,262.38	2,144.94	-569.60***	0.00
Final surface (sqft)	9,451.63	49,368.31	7,625.89	24,512.91	-1,825.74	0.14
Surface per unit (sqft)	1,218.60	888.58	1,272.60	917.00	54.00	0.16
Building age	98.83	25.17	101.43	23.54	2.60***	0.01
Year built	1,912.99	25.11	1,910.40	23.49	-2.59***	0.01
Distance to bus stop (m)	103.74	60.57	92.49	49.11	-11.25***	0.00
Distance to subway entrance (m)	291.77	159.61	297.35	202.16	5.58	0.50
Distance to bike-share station (m)	345.16	100.95	98.80	34.40	-246.36***	0.00
Distance to park (m)	337.13	215.64	311.22	206.45	-25.91***	0.00
Sale quarter	5.78	2.73	5.76	2.71	-0.02	0.84

Table 3: Balance table treated vs control ring, categorical variables, whole sample period

		Control ring 0 (N=7245)		Treated ring 1 (N=1370)		Total	
		N	Pct.	N	Pct.	N	Pct.
Post-period	0	3871	53.4	669	48.8	4540	52.7
	1	3374	46.6	701	51.2	4075	47.3
Treated (treatment ring × post)	0	7245	100.0	669	48.8	7914	91.9
	1	0	0.0	701	51.2	701	8.1
Elevator	0	7066	97.5	1314	95.9	8380	97.3
	1	179	2.5	56	4.1	235	2.7
Walkup	0	4832	66.7	877	64.0	5709	66.3
	1	2413	33.3	493	36.0	2906	33.7
Condo	0	7195	99.3	1364	99.6	8559	99.3
	1	50	0.7	6	0.4	56	0.7
Coop	0	7130	98.4	1354	98.8	8484	98.5
	1	115	1.6	16	1.2	131	1.5
Rental	0	4048	55.9	676	49.3	4724	54.8
	1	3197	44.1	694	50.7	3891	45.2

Table 4: Balance table treated vs control ring, building class categories, whole sample period

	Control ring 0 (N=7245)		Treated ring 1 (N=1370)		Total	
Building class category	N	Pct.	N	Pct.	N	Pct.
01 One Family Dwellings	789	10.9	162	11.8	951	11.0
02 Two Family Dwellings	1955	27.0	317	23.1	2272	26.4
03 Three Family Dwellings	1180	16.3	179	13.1	1359	15.8
07 Rentals - Walkup Apartments	2391	33.0	488	35.6	2879	33.4
08 Rentals - Elevator Apartments	117	1.6	46	3.4	163	1.9
09 Coops - Walkup Apartments	22	0.3	5	0.4	27	0.3
10 Coops - Elevator Apartments	53	0.7	8	0.6	61	0.7
11a Condo-Rentals	1	0.0	1	0.1	2	0.0
13 Condos - Elevator Apartments	9	0.1	2	0.1	11	0.1
14 Rentals - 4-10 Unit	688	9.5	159	11.6	847	9.8
17 Condo Coops	40	0.6	3	0.2	43	0.5

3 Empirical strategy

Identifying the impact of bike-share on real-estate transaction prices is lined with several empirical challenges. First, transaction prices are determined by a multitude of well-known factors (size, unit attributes, distance to amenities, etc.). Second, other dynamics might be at play that could determine the evolution of real-estate prices concurrently to bike-share stations. In this section, I discuss these challenges and how I address them with my estimation strategy.

The primary obstacle to determining the causal relationship between bike share and real-estate transaction prices is the non-random placement of bike share stations throughout the city. This deliberate selection process is logical from a city planning perspective, as it aims to ensure the success of the bike share program by targeting areas with a significant transportation market. In the case of NYC, this resulted in the initial deployment of bike share stations in Manhattan south of 60th Street and downtown Brooklyn.

The second challenge lies in the vast and diverse nature of NYC, which is continually evolving. Each year, numerous policies are implemented that may impact real-estate transaction prices, potentially confounding the estimation of the bike share's impact.

I address these challenges using a two-ring difference-in-differences (TRDD) strategy. The empirical strategy compares transactions taking place close to bike-share stations (inner or treatment ring) to those further away (outer or control ring), before and after the first deployment of bike-share. The two-ring construction identifies treatment (ring from zero to 150 meters around the station) and control (ring from 150 to 500 meters) transactions for each bike-share station, and estimates the difference between treatment and control transactions after bike-share implementation, net of station and time fixed effects.

The TRDD addresses the concerns around the identification by carefully selecting a credible control group of transactions (the outer ring). By focusing on potential impacts at a very small geographical scale, this identification strategy gives more ex-ante credibility to the parallel trends assumption required by difference-in-differences: everything else equal, it is reasonable to assume that transactions in the control and treatment ring follow similar sale-price trends in the absence of bikes share. I will provide evidence of this in later sections using an event study design.

The second main assumption of difference-in-differences is that no other concomitant policy that could have affected sale prices was enacted at the same time and place where bike-share was rolled out. I address this assumption by narrowing down the sample period to four and a half years (January 2011 to June 2015), which should limit the number of potential policies that might affect the estimates. Moreover, to be a significant threat to identification, other policies would have to exactly follow the spatial and temporal pattern of bike-share stations, which I control for to the best of my abilities.

The econometric specification used to estimate the TRDD is described by the following equation:

$$\begin{aligned} ln(P_{ijct}) =& \delta_{<150} D_{ij}^{<150} \times Post_{ijt} + \lambda_{<150} D_{ij}^{<150} + \beta' X_{it} + \gamma' Z_{it} \\ &+ \kappa_j + \varphi_c + \tau_t + \kappa_j \times t + \varepsilon_{ijct}, \end{aligned} \tag{1}$$

where P_{ijct} is the real sale price (base December 2015) of transaction i, matched to bike-share station j, at time t; $D < 150_{ij}$ is a dummy variable which equals one if unit i is within 150 meters of station j (i.e., in the treatment ring), and zero otherwise (i.e., in the control ring); $Post_{ijt}$ is a dummy variable indicating whether a unit is transacted after the opening of bike-share station j; X_{it} is a vector of unit attributes (elevator, building age, etc); Z_{it} is a vector of distances to nearby (dis)amenities (subway station, bus stop); κ_j and τ_t are station and year-month fixed effects, respectively; $\kappa_j \times t$ are station-specific linear time trends (some specifications); ε_{ijct} standard errors clustered at station level.

The coefficient of interest is $\delta_{<150}$, which represents the average treatment effect of bike-share on the treated transactions. In practice, it is the average change in log sales for a transaction within the treatment ring of a bike-share station after the opening of the station. Alternatively, I replace the treatment ring dummy $D < 150_{ij}$ by a continuous measure of distance (in hundreds of meters) to the matched bike-share station D_{ij}. The coefficient now reports, for a transaction, the average effect (in percent) of being 100 meters further from its matched bike-share station.

Finally, I also run a dynamic TWFE model in order to investigate the dynamic effect of bike-share with respect to the timing of treatment. The dynamic DD specification, also known as event study, plots the treatment effect for all periods. The dynamic specification also allows us to test for differential pretends between groups: by plotting the difference between treatment and control in the pre-treatment period, we will be able to evaluate the validity of the parallel trends assumption.

The specification for the dynamic DD is given in equation 2:

$$\begin{aligned} ln(P_{ijct}) = \sum_{k=-6}^{-2} \beta_k \cdot D_{ij}^{<150} + \sum_{k=0}^{6} \beta_k \cdot D_{ij}^{<150} + \beta' X_{it} + \gamma' Z_{it} \\ + \kappa_j + \varphi_c + \tau_t + \kappa_j \times t + \varepsilon_{ijct}, \end{aligned} \tag{2}$$

where k denotes the relative time to the first year of treatment, the other terms being the same as in the previous specification. The coefficients of interest are β_k, which are then plotted against relative time. In this setting, the reference period is relative time $k = -1$, therefore the plotted β_ks denote the relative difference between treatment and control groups compared to the period right before treatment.

4 Results

This section presents the results of the analysis. First, I report the results for the hedonic regressions, then the impact of the opening of bike-share stations on sale prices, and additional results using the distance to the station as the treatment variable.

Columns 1 to 3 of Table 5 display the results for the hedonic model. The goal of the hedonic model is to predict the sale price with unit and building attributes as accurately as possible. Good predictions from the hedonic model serve as both (1) a check that the data is sound and reliable (behaves predictably), and (2) a test of whether the unit attributes are correctly specified and constructed.

In Column 1 of Table 5, I regress the log sale price (in real terms, based on December

Table 5: Treatment impact on transaction sale prices

	Log sale price (2015 $)					
	(1)	(2)	(3)	(4)	(5)	(6)
Post-period				-0.1617***	-0.1759***	-0.1595***
				(0.0604)	(0.0617)	(0.0583)
Treated ring				0.0065	0.0137	-0.0175
				(0.0246)	(0.0234)	(0.0241)
Treated ring × Post-period				0.0481	0.0499	0.0581*
				(0.0349)	(0.0345)	(0.0338)
Surface per unit (100s sqft)	0.0089***	0.0092***	0.0244***		0.0091***	0.0243***
	(0.0015)	(0.0009)	(0.0011)		(0.0009)	(0.0011)
Building age (10s years)	0.0142**	-0.0029	-0.0017		-0.0032	-0.0019
	(0.0056)	(0.0036)	(0.0026)		(0.0036)	(0.0026)
Distance to bus stop (100s m)	0.0716***	0.0891***	0.0910***		0.0921***	0.0927***
	(0.0189)	(0.0162)	(0.0153)		(0.0166)	(0.0156)
Distance to subway (100s m)	-0.1046***	-0.0378***	-0.0328***		-0.0374***	-0.0327***
	(0.0165)	(0.0073)	(0.0069)		(0.0073)	(0.0069)
Distance to park (100s m)	-0.1634***	-0.0379***	-0.0307***		-0.0386***	-0.0312***
	(0.0119)	(0.0073)	(0.0070)		(0.0074)	(0.0071)
Bike-share station FE		Yes	Yes	Yes	Yes	Yes
Sale year-quarter FE		Yes	Yes	Yes	Yes	Yes
Building class category FE			Yes			Yes
Varying Slopes						
Sale year-quarter (Bike-share station)		Yes	Yes	Yes	Yes	Yes
Mean outcome pre-period	3,022,065	3,022,065	3,022,065	3,021,188	3,022,065	3,022,065
Observations	8,612	8,612	8,612	8,615	8,612	8,612
R^2	0.234	0.637	0.680	0.622	0.638	0.681
Within R^2		0.041	0.088	0.002	0.043	0.090
RMSE	0.753	0.519	0.487	0.529	0.518	0.486

Note: Significance codes: *: 0.1, **: 0.05, ***: 0.01. Standard errors clustered at the bike-share station level.

2015 prices) on a unit's surface area, building age, and distances to bus stops, subway entrances, and the nearest park. Most of the independent variables have the expected sign: the larger a unit's surface, the closer to a subway station and a park, the more expensive it is. Two variables seem to behave less obviously: the closer to a bus stop and the younger a building, the lower the transaction price. This, however, may make sense: everything else being equal, being closer to a bus stop probably means being closer to busy roads, a disamenity. Older buildings are associated with cachet and prestige, or might be located in areas where architecture was preserved, which may be associated with higher transaction prices.

In Column 2, I introduce bike-share station and year-quarter fixed effects, and allow the slope of year-quarter to vary across stations. From section 2, you will recall that each transaction is matched with at least one bike-share station (multiple if the transaction only falls into control rings). Introducing station and time fixed effects controls for space-specific and time-specific unobservable factors. The coefficient of unit surface remains stable, while the magnitude of distances to subway and parks decreases. Interestingly, building age is not a statistically significant predictor of transaction price. Adding building class categories in effect compares transactions within the same category, which might share common, unobservable traits associated with price. I do that in Column 3, and the magnitude of unit surface increases: every additional hundred square feet increase the price by 2.4%. The magnitude and significance of the other explanatory variables remain stable.

Moving on to Columns 4 to 6, I introduce the treatment and post-period dummies, our main coefficients of interest. There are no controls in Column 4, and the model is estimated with bike-share station and year-quarter fixed effects, allowing each station to have time-varying slopes. The main coefficient of interest, the interaction between the treatment and post-period dummies, is positive but not statistically significant at conventional levels (p-value 0.168). Controlling for unit attributes and distance to amenities in Column 5 does not sensibly change either the magnitude or the statistical significance of the coefficient (p-value 0.148). When adding building class fixed effects, however, the coefficient turns statistically significant (p-value 0.0869). Under the model in Column 6, the impact of having a bike-share station within 150 meters of a property increases its transaction price by 5.8% with respect to property transacted between 150 and 500 meters of the same bike-share station. By adding building class category fixed effects, we effectively compare transactions within the same building category, within the same bike-share station 500-meter radius, and within a given year-quarter, controlling for unobservable common factors within each of these dimensions. Since building categories within a given area might share a lot of characteristics in common, which are unobservable with the current data, comparing within building class might thus be important to remove heterogeneity and sharpen the estimation of our coefficient of interest. Moreover, the coefficient for the treated ring is statistically insignificant, indicating that bike-share stations are not placed in more ex-ante expensive areas with respect to their entire 500-meter radius.

The data would allow for more disaggregated fixed effects, using for example year-month of sale (instead of year-quarter), or building class (instead of building class category). While the appeal of employing more detailed fixed effects is clear, it is important to note that they also bring their own set of disadvantages. In particular, allowing for finer levels of fixed effects would greatly reduce the number of observations used in the estimation. The model would compare observations within each fixed-effect "cell", i.e., it would only compare, with each other, the observations that share the same values across all fixed effects. As we introduce more granular fixed effects, it becomes increasingly likely that a given observation (transaction) is unique in a given fixed-effect cell, and has no other observation to be compared with. In this case, the estimation ignores that observation and the coefficients are estimated using fixed-effect cells with at least two observations.[7] Therefore, I decide not to introduce more granular fixed effects, as those would restrict the number of observations used by the estimation.[8]

I check the consistency of this result in Table 6, Columns 3 to 5. Replacing the binary treatment variable with a measure of distance to the matched bike-share station, I find that every additional hundred meters away from the station reduces transaction prices by 2.1% (preferred specification, Column 5). This is in line with the result obtained with the dichotomous variable, and shows there exists a downward-sloping gradient of transaction prices around bike-share stations.

[7] It also follows that the model may underestimate the standard errors of the estimated coefficients, which would lead us to over-reject the null hypothesis.

[8] With the current set of fixed effects as in column 6, there are about five thousand observations (out of eight thousand) that belong to a fixed-effect cell with more than two observations. Introducing building class and year-month fixed effects lowers the usable observations down to two thousand.

Table 6: Additional specifications and continuous treatment

	Log sale price (2015 $)				
	(1)	(2)	(3)	(4)	(5)
Post-period	0.0870***	0.1332***	-0.0938	-0.1047	-0.0843
	(0.0253)	(0.0243)	(0.0693)	(0.0693)	(0.0645)
Treated ring	0.1408***	0.1050**			
	(0.0491)	(0.0416)			
Treated ring × Post-period	0.0532	0.0548			
	(0.0472)	(0.0403)			
Distance to bike-share station (100s m)			0.0025	0.0002	0.0066
			(0.0084)	(0.0080)	(0.0078)
Distance to bike-share station (100s m) × Post-period			-0.0191*	-0.0200*	-0.0213**
			(0.0113)	(0.0111)	(0.0102)
Surface per unit (100s sqft)		0.0084***		0.0091***	0.0243***
		(0.0015)		(0.0009)	(0.0011)
Building age (10s years)		0.0130**		-0.0030	-0.0018
		(0.0056)		(0.0036)	(0.0026)
Distance to bus stop (100s m)		0.0810***		0.0917***	0.0928***
		(0.0187)		(0.0167)	(0.0158)
Distance to subway (100s m)		-0.1059***		-0.0372***	-0.0326***
		(0.0168)		(0.0072)	(0.0069)
Distance to park (100s m)		-0.1644***		-0.0391***	-0.0314***
		(0.0120)		(0.0075)	(0.0072)
Bike-share station FE			Yes	Yes	Yes
Sale year-quarter FE			Yes	Yes	Yes
Building class category FE					Yes
Varying Slopes					
Sale year-quarter (Bike-share station)			Yes	Yes	Yes
Mean outcome pre-period	3,021,188	3,022,065	3,021,188	3,022,065	3,022,065
Observations	8,615	8,612	8,615	8,612	8,612
R^2	0.009	0.244	0.622	0.638	0.681
Within R^2			0.002	0.043	0.090
RMSE	0.857	0.748	0.529	0.518	0.486

Note: Significance codes: *: 0.1, **: 0.05, ***: 0.01. Standard errors clustered at the bike-share station level.

4.1 Discussion

The results displayed above indicate that the introduction of bike-share stations had a positive impact on residential real-estate transaction prices. Properties within 150 meters of a station saw their transaction price increase by 5.8% on average, compared to property transactions further away (between 150 and 500 meters), and within a given building class category. Moreover, every additional hundred meters away from a bike-share station reduces transaction price by 2.1% on average, compared to before the opening of the system.

These impacts on real-estate transaction prices are relatively significant, but within the range we would expect from this type of intervention (see [citations]). Of particular interest to us are the estimated effects of bike-share on rents determined by Shr et al. for Taiwan. They find that six months after the opening of bike-share, rents increase on average by 1.71% for units within 150 meters of stations. While of the same order of magnitude, my results are relatively larger. [Should I try to give some ex-post rationalisations?].

[Should I discuss the limited statistical significance of the main coefficient more in-depth?]

References

Gupta, Arpit, Stijn Van Nieuwerburgh, and Constantine Kontokosta (2022). "Take the Q Train: Value Capture of Public Infrastructure Projects." In: *Journal of Urban Economics* 129, p. 103422. DOI: 10.1016/j.jue.2021.103422.

Shr, Yau-Huo et al. (2023). "Air Quality and Risky Behaviors on Roads." In: *Journal of Environmental Economics and Management* 118, p. 102786. DOI: 10.1016/j.jeem.2023.102786.