Housing and Income Effects of Rapid Transit Expansion: Evidence from Houston's Green and Purple Lines

Hunter Tyszka

July 27, 2025

Abstract

In May of 2015, Houston opened the long-awaited Green and Purple lines of the METRORail light rail transit system. In an era plagued by major issues that affect large metropolitan areas, such as roadway congestion, affordability, and wealth inequality, these lines could pave the way to solving some of these issues at a local level. However, the effects of these new lines on housing prices and local incomes are not well understood. This study will use a hedonic and spatial difference-in-differences approach to estimate the impacts of the new lines of housing price capitalization and incomes at the census tract level. This project is expected to be capitalized positively in prices and increase incomes, as people reap the benefits of better urban accessibility and connectivity. Thus, through this study, we should be able to answer whether the construction of major public transit projects is able to positively benefit the communities through which it is built.

1 Introduction

1.1 Public transit investment - an expensive tool for change

North American cities are facing challenges that are top of mind for all levels of government, the business community, and especially residents. Chronic roadway congestion, affordability, stagnating wages, and rising wealth and income inequality are creating real and salient concerns in communities across Canada and the USA. Fortunately, cities are equipped with the ingenuity and policy tools to address these major concerns. One possible remedy in the arsenal of municipal and state-level governments is investment in transit infrastructure. Public transit projects, especially high-capacity rapid transit, including commuter rail, subways, and light rail transit, have the potential to transform communities and create better living conditions for residents.

By offering fast and reliable alternatives to automobile travel, transit agencies are able to connect communities and unlock opportunities that are limited by difficult commutes. As a result, it is possible that residents and business owners along these corridors can make long-term decisions that are anchored on the permanence that rapid transit provides. This could mean choosing to live along a corridor for greater access to education and employment opportunities or opening a business in response to higher numbers of nearby customers and residents. Then it seems possible that incomes could grow in adjacent communities as a result of these projects. In a similar vein, the value of nearby real estate can increase as the benefits of nearby rapid transit are capitalized into the housing market. This could result in lifting both income and wealth of transit-adjacent neighbourhoods. It is critical to understand whether these possible benefits are actually achieved as a result of transit investment. These projects are often major public expenditures that result in years of planning and construction. Determining whether they are effective is critical to optimizing their implementation.

In this paper, I will focus on a recent expansion of the METRORail rapid transit system in Houston, Texas. In 2004, the city of Houston opened its first light rail line, the Red Line, along a downtown corridor. In the following years, plans were drawn up to significantly expand the system across the region, with a referendum being passed in 2003 to build an additional five lines (METRO, 2003). However, due to funding issues and local opposition, only the Green and Purple lines were fully implemented. Although ground technically broke in 2009, funding was withheld from the project until the federal government intervened in 2011 (Samuel, 2011). In 2015, the city opened the Green and Purple lines in the southeastern region of the city. I will focus on both the housing market capitalization effects of the project and the impacts on incomes in the adjacent census tracts. Using the two major milestone dates of 2011, when funding was allocated, and 2015, when the lines were opened, I will seek to evaluate the ability of these projects to affect the incomes and real estate markets of the adjacent communities.

1.2 Green and Purple Lines - a hole in the literature

Assessing the impacts of transit infrastructure on real estate markets has been approached from many directions in the literature. However, the impacts of the Houston Green and Purple lines have not been studied, and few papers have sought to analyze both the effects on housing markets and incomes in the same communities. Assessing both simultaneously sets the foundation for understanding both income and wealth effects and the distribution of benefits across residents, as a result of transit infrastructure investment.

1.3 Relationship with existing literature

Although the Green and Purple lines have not been studied in Houston, previous work in the literature has focused on unpacking the housing price capitalization of new high-order public transit in major cities around the world. Pan, 2019, studied the impacts of the Red Line in Houston on residential property values, with an ordinary least squares hedonic model. Using housing data from 1982-2010, they find that the metro has significant positive effects on property values after several years of metro service. While my paper will follow a similar approach, using a hedonic model to understand capitalization, I will use a difference-in-differences model to compare real estate values before and after both the construction start date and the opening date to analyze impacts over time.

Alternatively, Keeler, Stephens, 2022, focused their study on the LA Metro system and the capitalization effects of the Gold and Expo lines. They used a hedonic spatial difference-in-differences model to quantify the effect, and varied treatment by distance from individual properties to the nearest metro stop. They find that there is a distance decay effect of the metro and that the magnitude and direction of capitalization vary based on the characteristics of the neighbourhood and treatment timing. In my paper, I will follow a similar methodological approach and will similarly vary treatment timing and distance to investigate the possibility of decay and anticipation effects as well. To build on their analysis, I will investigate changes to different quantiles of housing types, and will carry the analysis into census tract level income variation.

1.4 Brief overview of this paper

This paper will analyze two outcome variables. First, it will investigate the relationship between metro construction and the capitalization of this investment into housing prices. I will design a hedonic difference-in-differences approach that will control for housing and neighbourhood characteristics to tease out the capitalized value of proximity to newly built rapid transit. The treated properties will be those within varying quarter-mile distances from Green and Purple metro stations, and the control properties will be those outside a quarter-mile buffer from the treated group.

The second outcome of interest will be the relationship between light rail transit construction and mean incomes in census tracts near the constructed lines. I will use a difference-in-differences approach, specifying tracts within various distances from metro stations as treated groups. The control groups will be census tracts outside of this treated threshold.

For both sets of outcome variables, I will focus on two treatment dates to analyze anticipation effects. First, the date when funding was officially allocated to the projects, in December 2011, and second, the date of their openings in May 2015.

1.5 Summary of results

In terms of housing capitalization, both at the opening of the Green and Purple lines in 2015 and the release of federal funds in 2011 appear to result in statistically significant increases in housing prices. These increases appear to decay at properties farther away from the stations but appear between zero and one mile from the station.

For neighbourhood effects, the impact of the metro opening on mean incomes at the census tract level is not clear. For 2011, it appears that the start of construction increased income, but not in a statistically significant way. For 2015, after the opening of the Green and Purple lines, mean incomes in adjacent census tracts appear to fall, although these estimates are statistically non-significant. The direction of effects at both of these treatment dates is surprising.

2 Data

To analyze the outcomes of the construction of the Green and Purple light rail lines on housing capitalization, I will use data that tracks housing characteristics, tax assessments, and transaction amounts over time. Similarly, to isolate the impacts on income, income data from the US Census, American Community Survey.

2.1 Attom Data - Property Transfer Recording Data

For property-level transaction and characteristics information, I used Attom Recorder (Attom Data, 2024) and Attom Assessor (Attom Data, 2023) data to consolidate an unbalanced panel dataset. The Attom Recorder data tracks property transactions that have occurred throughout the United States between 2010 and 2024, and has more than one hundred variables that describe the properties in question. The set has been filtered for transactions that were conducted at arms length to remove lower-than-market value transfers to related parties. I have then used this data to construct the key outcome variable 'log_transfer_amount', the log of the most recent price a property was traded for. The transfer amount is right-skewed, and thus by taking the natural log, I am able to remove the skew and focus on percent changes in valuation. To construct the 'treated' variable in the difference-in-differences model, the 'transfer_date' variable was used as the day a trade occurred. Unfortunately, very few properties were sold more than once in this range of time, so I have not been able to construct a balanced panel. This means that the outcome variable will measure a change in average prices over properties within a neighbourhood, rather than an average change at the property level. However, there are still more than 10,000 properties within a three-mile buffer of the metro system, generating a large enough sample to provide significant statistical power.

The Attom Assessor data contains tax assessment information for the year 2023 for residential properties in the United States. This dataset contains very detailed records of all residential properties within Harris County, Texas, which I have used to control for the number of bedrooms, bathrooms, property size, and build date, factors that drive housing prices. I have merged this data with the recorder data to construct a detailed property-level dataset which describes the transaction prices for transferred homes and their specific features. For use in my hedonic analysis, this will enable me to pin down the capitalized value of the metro system in residential values.

2.2 American Community Survey - US Census Data

For census-level data on income, education, and employment, I have compiled a panel of American Community Survey (ACS) data from the US Census Bureau (U.S. Census Bureau). For the years 2010 through 2020, I collected the census tract level data for Harris County, Texas from the ACS relating to commuting characteristics, education, and income over 12 months. The key variable of interest is the log of mean income over 12 months. The remaining data will be used as a control in the regression analysis. I merged these datasets and filtered them to focus only on the census tracts within 3 miles of the Houston metro system, to center on the city of Houston. The smallest possible observational unit for this data is the census tract level. Thus, my further analysis is at this lowest level of aggregation. The data is organized as 5-year moving average groups and hence, when using the data for difference in differences regression, I will use one observation before and after each treatment date. I used a yearly panel of these moving averages only when assembling event study designs for use in a common trends analysis. Due to the limitations of moving average data, it is possible that outlying events may skew results, and annual changes in income are impossible to identify individually.

2.3 Census Tract and Houston METRORail - Geographic Data

For the purpose of analyzing the census tract level information and for assigning properties to specific census tracts, I used tract level shape files from the US Census Tiger database (U.S. Census Bureau) and from Houston METRORail's shape file service (Houston-Galveston Area Council, 2025). Using GIS software, I computed a network walking distance from each property to the nearest metro station and Green/Purple station, for use in determining the threshold for assignment to the treated group. Furthermore, since some metro lines were planned but were not constructed, I created shape files using archived planning documents (USDOT, 2010) that described the University and Uptown Houston metro lines that were approved in the 2003 referendum, for use in robustness checks.

3 Empirical Approaches

I will use two difference-in-differences designs to analyze the impacts of the Houston METRORail expansion on housing capitalization and on incomes in neighbouring census tracts.

For both housing capitalization and income data, I will use two treatment timings. First, December 2011 is the date when funding for the Green and Purple line was finalized and confirmed by the US Federal government. I hypothesize that participants in the housing market may only have internalized the potential benefits or drawbacks of the metro after it was confirmed that it would be constructed, which was officially determined in 2011. Hence, I believe this is a logical assessment as to when the effects may start to be realized. Moreover, I am limited in the Attom data to a time period of 2010 - 2024, and thus 2011 is the earliest feasible treatment date to have housing data in both the pre and post-2011 period.

Secondly, I will consider 2015 as a treatment date. This date coincides with the opening of the Green and Purple lines in May of 2015. I suspect that by the time the metro line opens, forward-looking participants of the housing market would have already capitalized most of the potential amenity value of the metro lines into property values. However, there may be additional unrealized positive or negative effects such as faster than expected increases in commute times, changes to neighbourhood dynamics, or increases in noise and crime that could be the cause for additional adjustments to housing values.

3.1 Housing Capitalization - Methodology

First, I will estimate the impacts of the construction of the METRORail Green and Purple lines on housing price capitalization by using a hedonic difference-in-differences model. The outcome variable of interest is the log of property transfer amount. The treatment variable in the model will vary based on quarter-mile or 400 meter distances from each property to the nearest Green and Purple metro station. Typically, a distance between a quarter and half mile walk to a station is considered reasonable for the average commuter, so this is a logical treatment assignment. To differentiate between the treated and control regions, a quarter mile buffer between treated and control groups is used. This is a similar approach to that used by Keeler and Stephens (2022). To account for possible distance decaying effects as properties get further away from the metro stations, I will use various treatment thresholds to analyze effects at different treatment distances.

These distances and treatment definitions are as follows for Houston properties:

- Treated: up 1/4 mile from the nearest Green or Purple station. Control: beyond 1/2 mile.
- Treated: up to 1/2 mile from the nearest Green or Purple station. Control: beyond 3/4 mile.
- Treated: up to 3/4 mile from the nearest Green or Purple station. Control: beyond 1 mile.
- Treated: up to 1 mile from the nearest Green or Purple station. Control: beyond 1 mile.

The preferred specification in future comparison and robustness checks will be the 1/2 mile treatment buffer from the Green and Purple line.

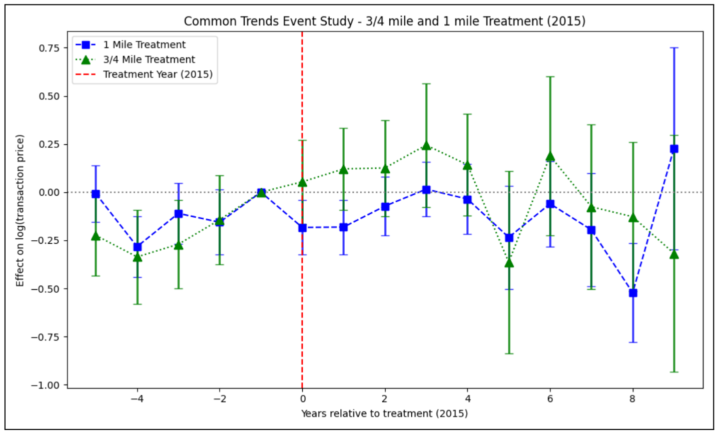
3.1.1 Identifying assumptions and challenges

The identifying assumption for this difference-in-differences approach is that, in the counterfactual scenario where the Green and Purple lines were not constructed, the prices for properties that are nearest to the Green and Purple line stations would have grown at a similar rate to those that are further away or not affected by the new light rail lines. It is possible for this assumption to hold, if the characteristics of those properties closest to the metro alignment are similar to those that may be slightly further away. Similarly, if the metro alignment was arbitrarily selected along given residential areas, it is likely that all homes within that neighbourhood would progress in values at a similar pace.

Given the spatial nature of this analysis, certain neighbourhoods and regions of Houston may have confounding characteristics that may introduce challenges to this assumption. For example, certain areas may have higher densities, differing levels of education, and incomes. Moreover, the fact that the city decided to build along this corridor suggests that there are some observed or unobserved characteristics that may have led to different trends with or without treatment. Thus, the plausibility of this assumption will vary based on the treatment distance definition. Properties that are close to the metro alignment may have underlying unobserved characteristics that differ from properties further away. However, considering properties that are nearby these treated groups, but not extremely far away, it is likely that they share common characteristics that could drive similar trends in the absence of treatment. Moreover, this plausbility may differ using the 2011 treatment date from the 2015 treatment date. In 2011, there are no major simultaneous events that confound with the beginning of the light rail construction, except for the recovery from the financial crisis. However, by 2015, certain properties have experienced a permanent adjacency to the future light rail for several years during construction, and thus may develop different trends in property prices.

Figure 1 shows the event study trends for the 3/4 mile and 1 mile treatment categories. These trends represent the upper and lower bound of the four treatment specifications, with the 1/4 and 1/2 mile definitions fitting in between.

Figure 1: 2015 Housing Common Trends Event Study



Notes: This figure plots the event study coefficients for the 1 mile and 3/4 mile treatment categories. It does not appear that common pre-trends hold with great confidence.

As seen in Figure 1, it does not appear that the parallel trends hold with great confidence, with statistically different values in several of the pre-periods. Figure 4 in the Appendix demonstrates that the treated group tends to grow at a slower or negative rate and only catches up to the control group after treatment. Table 5 shows a balance table for the half-mile treated group in the Appendix shows that control properties tend to be slightly newer builds with more bedrooms and held by wealthier individuals. Although characteristics are relatively similar for the transfer price, there is some lack of balance in building size and income.

For the case of 2011 as a treatment year, given there is only one pre-period in the dataset, Table 6 in the Appendix presents balance statistics to bolster confidence in the parallel trends. The control properties tend to be transfered at slightly higher values, built more recently, and have slightly more bathrooms and bedrooms. These characteristics suggest there may be some lack of balance, but there is little data available to show the full pre-trends.

Due to the spatial nature of the analysis, it is possible that spillovers exist which may violate the stable unit treatment value assumption as well. For example, in building the metro, migration could occur between neighbourhoods as people move toward or away from the metro. This means that control properties potential outcomes also change as a result of treatment, which can bias the estimates upward or downward depending on the direction of migration.

Having analyzed these assumptions, and with the risk of parallel trends being violated, I will elect to also employ an Honest DiD strategy on top of the baseline specifications. This strategy will project the pre-trends onto the post-treated period, to create a range of possible values which depend on the degree to which the parallel trend assumption may be violated. This will help to qualify the possible true magnitude of any treatment effect.

3.1.2 Estimating equation - Hedonic Difference in Difference

The nature of a hedonic model is to decompose the price of a property into the various attributes that contribute to it. Thus, in my model, I include a series of controls that relate to first-order characteristics that drive housing prices, such as size, number of rooms, and various amenities. In one specification, I will also include fixed effects, to understand how controlling for factors fixed in time and over census tracts may affect the estimates.

The estimating equation for this model is as follows:

$$Log(PropertyPrice) = \beta_0 + \beta_1 Post_t + \beta_2 Treated_i + \beta_3 Post_t * Treated_i + \gamma H_{it} + \delta_i + \delta_t + \epsilon_{it}$$

β_3 is the coefficient of interest, which measures the capitalization effect to the Green and Purple stations for properties designated as treated, for observations post 2011 or post 2015. The index i refers to individual residential properties. The index t refers to the dates on which these properties were transferred. H_{it} is a vector of housing characteristics that tend to drive prices, including number of bedrooms, bathrooms, size in square feet, build date, and distances to major employment centers. These first-order conditions help to refine the estimate of the impact of the metro on prices. Finally, δ_i and δ_t are fixed effects for census tract and year which account for factors that vary between census tracts and between years. Although β_3 measures the change between treated and control, and pre and post treatment, the sum $\beta_3 + \beta_1$ represents the total change between the pre and post periods for the treated group. This can be informative in understanding the overall trend for the treated group, not just the treatment effect relative to the controls.

3.1.3 Expected Results

With the treatment year as 2015, I expect there to be a positive but relatively small effect on housing price capitalization, represented by a positive β_3 and a positive $\beta_3 + \beta_1$. Most of the capitalization likely occurred leading up to the metro opening, as anticipation of the project was priced into local residential prices. It also follows logically that the houses closest to the metro rose at a higher rate, than those further away, as defined in the treated and control group. This aligns with the theory that consumers and participants of the market are forward looking, and take into consideration all available information. Therefore, it is unlikely that there existed any significant effects that were not already priced in. However, there could still be some externalities and unforeseen benefits or drawbacks that were only visible after opening, such as even faster than expected commute times, increased accessibility to friends and family, or higher levels of noise and crime. However, I suspect that these benefits are likely positive, especially over time as more economic opportunities are created in the area.

To account for heterogeneity that may occur at different levels of property values, I will also report coefficients from a quantile regression. This analysis will help to inform whether the metro system affects properties differently based on their existing valuations. I expect properties with lower existing values to increase at a higher rate, as there is greater potential and drive for redevelopment on these lands.

For the 2011 treatment year, I expect to see much larger and more significant effects on housing prices, expressed in β_3. Since 2011 is the date at which certainty of the project was evident, I expect forward-looking market participants to price the amenity value of the metro system as construction continues. I suspect prices to increase faster in properties closest to the stations, with the expectation that the metro will bring greater positive benefits to communities than negative externalities. That being said, given construction on these large-scale projects is highly visible, cumbersome, lengthy, and challenging, it would not be surprising to see some degree of decline for treatment areas directly along the corridor in the short term. However, I expect over the long time frame that there will be net positive effects. It is important to qualify here, that the quantity of data available before 2011 may present problems in accurately estimating the results for this treatment date.

3.2 Neighbourhood Impacts - Methodology

Next, I will estimate the impacts of the Green and Purple lines on driving changes to neighbourhood level incomes. I will use a spatial difference in differences approach to complete this analysis. The outcome variable of interest is the log of mean income from the census tract-level data of the American Community Survey. Given that some census tracts tend to be very high income, while others are very low income, I have used the log of mean income to smoothen the distribution and account for outliers. Similar to the property data, I will use several treatment definitions in order to account for distance decaying effects for the system.

The treatment definitions for census tracts are specified below:

- Treated: 1 mile from metro lines. Control: between 1 and 2 miles from metro lines.

- Treated: 1 mile from metro lines. Control: between 1 and 3 miles from metro lines.

- Treated: 1.5 miles from metro lines. Control: between 1.5 and 3 miles from metro lines.

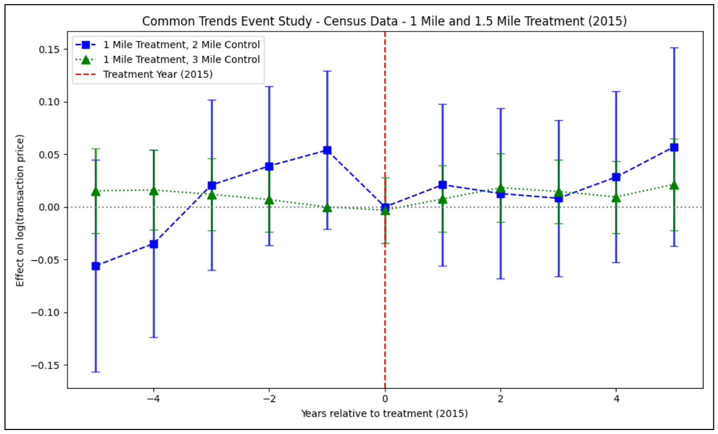
The preferred specification is the second definition, selecting treated tracts as 1 mile from metro lines and control tracts as those between 1 and 3 miles away.

3.2.1 Identifying assumptions and challenges

The identifying assumption for this approach is that in the absence of the construction of the Green and Purple lines, the mean incomes for the tracts nearest to the new light rail lines would have followed a similar trend as those far away and outside of reasonable walking distance. It is plausible that this assumption holds, as census tracts are drawn to attempt to maintain stable population levels and socioeconomic characteristics. Thus, census tracts that have similar populations and comparable levels of income before the construction of the metro, may very well follow similar trends, that could differ only after residents within them are exposed to closer and faster transit. Especially for 2011, when little changed before and after treatment other than an allocation of funds, I would not expect the assumption to be violated. For 2015, there is some cause for concern given that construction before treatment in the census tracts may create negative externalities that change trends. It is also possible that some tracts that have higher levels of income or education, may follow different trends than those with lower levels. To account for some of these concerns, I will employ matching on coarse census tract characteristics to create a treatment and control population that is balanced on observables, as a robustness check.

Figure 2 demonstrates the trend in the differences between the treated and control groups for the American Community Survey log mean income variable, using 2015 as a treatment date. For the preferred specification, the pre-trend is very close to zero, which gives confidence that the common trends may hold. For the 1 mile treated, 2 mile control however, the trend is much more volatile, suggesting that common trends may not hold. The control groups in this case may be affected by spillovers from the rail construction.

Figure 2: 2015 Census Common Trends Event Study



Notes: This figure plots the event study coefficients for the two 1 mile treated categories. The category with the 3 mile control tends to have common pre-trends, while these trends are more volatile for the 2 mile control group.

For the case of 2011, given census tract boundaries changed after 2010, there is little consistent data available to study the pre-trends for 2011. However, the Appendix tables, Table 8, Table 9, and Table 10 compare observable statistics between the treated and control groups. Generally, they show log_mean_income is relatively balanced, but there are differences in the unemployment rate and education levels. To overcome these differences, I will use a matching design in the robustness checks to boost confidence in my estimates.

3.2.2 Estimating equation

Included in my spatial difference-in-differences model will be census tract level fixed effects and the percentage of the population that has a bachelors degree. This will help to account for variation between census tracts that is specific to that tract, as well as variation in income that is dependent on the level of education. This will help to ensure that tracts are more comparable throughout the difference-in-differences model.

Since the data I am using is 5-year moving average data, I will use only one observation before and after each treatment date, to avoid using a particular year more than once. For 2015, the data will include 2010-2014 for the pre-period and 2015-2019 for the post period. For 2011 this will be 2007-2011 for the pre period and 2012-2016 for the post period.

The estimating equation for this model is as follows:

$$Log(MeanIncome) = \beta_0 Post_t + \beta_1 Treated_i + \beta_3 Post_t * Treated_i + \gamma N_i + \delta_i + \epsilon_{it}$$

β_3 is the coefficient of interest, which measures the proximity effect to the Green and Purple lines on log mean income for census tracts within the treated group, after the treatment date. The index i refers to the census tract, and index t refers to the 5-year moving average group of data. N_i represents a control for the level of education and accounts for variation in income that can be attributed to education. δ_i represents census tract fixed effects, characteristics that vary between census tracts but are fixed over time. Since there are only two observations, one each for a pre and post-treatment period, fixed effects for year are not included as they would be collinear with the $Post_t$ variable.

3.2.3 Expected Results

For the 2011 treatment date, when funding was granted for the project, I do not expect a statistically significant increase in income, represented by β_3. Since the metro would not have been operational at this time, it would not have decreased commute times and led to increased levels of business activity along the corridor, unless there were anticipation effects. In terms of business activity, these anticipation effects are unlikely as prolonged construction create difficulties for customers trying to access businesses in this project phase. This could even lead to falling incomes if layoffs and business solvency issues are prevalent. However, it is possible that in anticipation of the future light rail line, individuals migrated to these census tracts, which could either drive up average incomes if higher income individuals moved in, or vice versa. If there is inmigration of higher-income individuals, this would have resulted in a negative coefficient for β_1 as mean income in the control tracts would have decreased. To help account for this, I have included education as a control to account for abrupt changes in education and therefore income, if individuals engaged in migration.

I expect that after 2015, effects on income may appear, and thus β_3 may be positive and statistically significant. The presence of these new light rail lines in the neighbourhoods may create opportunities for residents. First, residents may be able to access a greater range of employment opportunities in the same or shorter commute time, increasing their chance of lowering unemployment and increasing household incomes. Secondly, with a higher-order transit line nearby, businesses may seek to locate in these tracts knowing that there can be larger volumes of customers within reach. This, too, may drive up employment activity and increase incomes in these neighbourhoods. In the years after the metro opening, these factors may have the tendency to lock-in, and thus may drive up mean incomes in the census tracts. It is also important to consider that there is a possibility of spatial spillover. That is, higher income individuals in tracts that are considered controls, may migrate toward or away from the Green and Purple lines, which can drive up or lower average incomes in these regions, without raising incomes for existing residents. These forces could violate the stable unit treatment value assumption and contaminate the estimated effects of the regression.

4 Results

4.1 Housing Capitalization - Results

The following figures are structured using five columns, with the outcome variable as the log transfer price of properties. The first column represents a treatment area of 1/4 mile, with control properties beyond 1/2 mile from the nearest Green or Purple station. Each subsequent column defines treatments as an additional 1/4 mile buffer from the metro stations.

According to Table 1, there is a positive effect on housing prices near the newly opened Houston METRORail Green and Purple stations. For properties within a quarter mile of a station, column (1), the effect is statistically significant at the 0.05 level and is positive with an estimate of about 32.5% increase after 2015. In column (2), for properties within a half mile distance from the station, there is a statistically significant 30.9% increase in transfer prices after 2015. Column (3) repeats the 1/2 mile treatment model, with the inclusion of fixed effects, which marginally reduces the treatment effect to 29.3%. There appears to be a distance decaying effect as well. In columns (4) and (5), where treatment is defined as a 3/4 mile and 1 mile distance from the station, the increase in prices is statistically insignificant and is of a smaller magnitude of 18.1% and -0.53% respectively. In all models, the sign of the coefficients for controls, especially those that are likely to heavily influence property prices, such as bedrooms, bathrooms, and building size, are all positive.

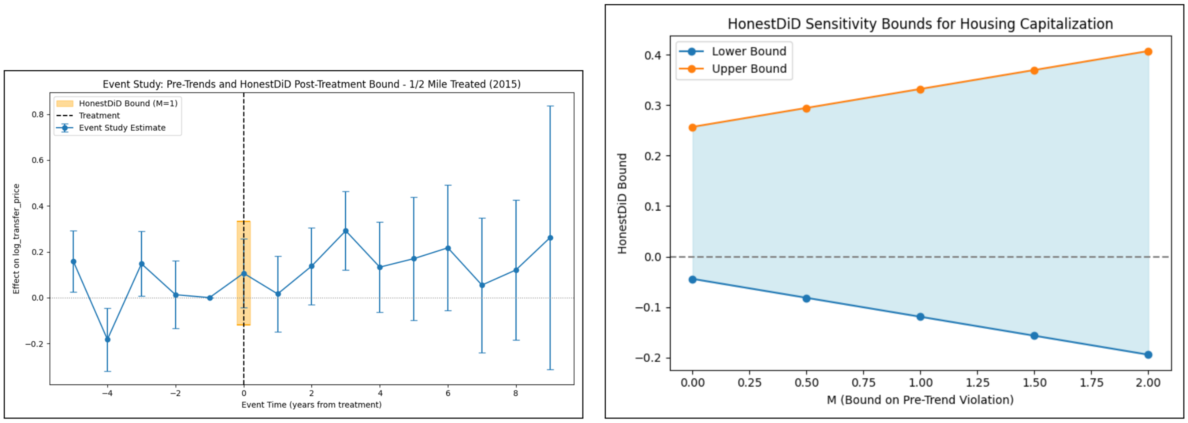
Table 1: Housing Price Capitalization by Treatment Distance - (2015) Treatment Year

	(1)	(2)	(3)	(4)	(5)
Intercept	17.774***	18.195***	18.640***	15.480***	25.862***
	(4.335)	(4.258)	(2.294)	(4.086)	(3.046)
Treated	0.101	0.059	-0.431	-0.066	0.250*
	(0.229)	(0.191)	(0.334)	(0.186)	(0.144)
Post (post)	0.366***	0.358***	0.071	0.368***	0.511***
	(0.067)	(0.073)	(0.080)	(0.086)	(0.075)
Treatment × Post (treated_post)	0.325**	0.309***	0.293***	0.181	-0.053
	(0.146)	(0.099)	(0.084)	(0.111)	(0.097)
Bedrooms Count	0.029	0.063	0.094**	0.068	-0.096***
	(0.068)	(0.066)	(0.036)	(0.073)	(0.031)
Bath Count	0.131**	0.130**	0.022	0.121**	0.190***
	(0.056)	(0.055)	(0.035)	(0.058)	(0.039)
Area Lot (1000s sq. ft.)	0.004	0.004	0.010***	0.006	-0.003
	(0.005)	(0.004)	(0.003)	(0.004)	(0.007)
Area Building (1000s sq. ft.)	0.421***	0.393***	0.263***	0.389***	0.454***
	(0.057)	(0.051)	(0.030)	(0.052)	(0.053)
Year Built	-0.004	-0.004*	-0.005***	-0.002	-0.008***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Distance to Downtown (1000 ft.)	0.042**	0.040**	-0.011	0.033	0.072***
	(0.019)	(0.020)	(0.079)	(0.020)	(0.023)
Distance to Medical Centre (1000 ft.)	-0.031	-0.028	0.204***	-0.037*	0.005
	(0.019)	(0.021)	(0.073)	(0.019)	(0.018)
Distance to Uptown (1000 ft.)	-0.074***	-0.076***	-0.155**	-0.075***	-0.134***
	(0.011)	(0.011)	(0.070)	(0.011)	(0.016)
Year FE	No	No	Yes	No	No
Tract FE	No	No	Yes	No	No
R^2	0.26	0.27	0.53	0.27	0.30
Control Obs.	12099	10707	10707	8684	4926
Treated Obs.	591	1695	1695	3087	5110
Total Obs.	12,690	12,402	12,402	11,771	10,036

Notes: The outcomes variable of interest in this table is the log transfer price. All coefficients can be interpreted as approximate percentage changes. Column (1) corresponds to the model with properties 1/4 mile from a Green/Purple station. Column (2) is the 1/2 mile treated specification. Column (3) is the 1/2 mile specification with fixed effects. Column (4) is the 3/4 mile treated specification. Column (5) is the 1 mile treated specification. Clustered standard errors at the census tract level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In Figure 3 the Honest Difference-in-Differences event study shows the range of reasonable coefficients of the difference-in-differences regression under several scenarios if the parallel trends assumption does not hold. As opposed to the 29.3% change expressed in the preferred specification for β_3, the event study suggests a treated coefficient between -10% and 30%. This effectively qualifies the 29.3% coefficient as an upper bound on reasonable results.

Figure 3: 2015 Property Prices - Honest Difference in Differences



(a) Event Study (b) Sensitivity Bounds

Notes: Plot (a) shows the Honest DiD event study demonstrating the possible range for the treatment period coefficient in yellow. Panel (b) shows the sensitivity bound for the treatment effect in the post period depending on the degree of parallel trend violation.

To analyze the heterogeneity that could be present in the capitalization of the metro into housing prices, Table 11 shows outcomes at different quantiles of transfer price. It appears that the capitalization effect is generally lower, when broken down by quantile, and houses at the median tend to benefit the most from the metro, with an increase in price of about 21.8%. The most expensive houses, in the 75th percentile, experience an increase of 15.9%, and those in the 25th percentile experience an increase of 20.4%.

For the 2011 treatment date, according to Table 2, there is a positive effect on log transfer prices for properties adjacent to the future metro system. This effect also displays a distance decay effect similar to that of the 2015 treatment date. For the quarter-mile treatment distance, column (1), the effect is statistically significant at the 0.01 level with an increase of prices of 45.4%. The half-mile distance, column (2) shows a smaller effect of 28.0%, and including fixed effects in column (3) further reduces this amount to 26.2%. The wider treatment distances of 3/4 miles and 1 mile are both statistically insignificant with price increases of 19.4% and -5.9%. The sign and coefficients for the major controls are in the expected directions.

Table 2: Housing Price Capitalization by Treatment Distance - 2011 Treatment Year

	(1)	(2)	(3)	(4)	(5)
Intercept	19.174***	19.714***	18.687***	16.879***	26.746***
	(4.024)	(3.930)	(2.271)	(3.796)	(2.922)
Treated	-0.166	-0.080	-0.534	-0.162	0.264
	(0.263)	(0.208)	(0.357)	(0.218)	(0.173)
Post (post)	-0.231**	-0.246**	-0.191	-0.233*	-0.041
	(0.110)	(0.119)	(0.176)	(0.136)	(0.075)
Treatment × Post (treated_post)	0.454***	0.280**	0.262***	0.194	-0.059
	(0.162)	(0.135)	(0.099)	(0.150)	(0.122)
Bedrooms Count	0.026	0.058	0.093**	0.061	-0.095***
	(0.063)	(0.060)	(0.036)	(0.066)	(0.030)
Bath Count	0.147***	0.147***	0.023	0.138**	0.203***
	(0.055)	(0.053)	(0.035)	(0.056)	(0.040)
Area Lot (1000s sq. ft.)	0.004	0.004	0.010***	0.005	-0.004
	(0.004)	(0.004)	(0.003)	(0.004)	(0.007)
Area Building (1000s sq. ft.)	0.417***	0.388***	0.262***	0.385***	0.453***
	(0.057)	(0.051)	(0.030)	(0.052)	(0.052)
Year Built	-0.004**	-0.005**	-0.005***	-0.003	-0.008***
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Distance to Downtown (1000 ft.)	0.041**	0.038**	-0.011	0.031*	0.074***
	(0.018)	(0.019)	(0.079)	(0.019)	(0.023)
Distance to Medical Centre (1000 ft.)	-0.032*	-0.028	0.204***	-0.037**	0.004
	(0.018)	(0.020)	(0.072)	(0.018)	(0.017)
Distance to Uptown (1000 ft.)	-0.070***	-0.072***	-0.155**	-0.071***	-0.129***
	(0.011)	(0.011)	(0.069)	(0.010)	(0.016)
Year FE	No	No	Yes	No	No
Tract FE	No	No	Yes	No	No
R^2	0.25	0.26	0.53	0.26	0.27
Control Obs.	12099	10707	10707	8684	4926
Treated Obs.	591	1695	1695	3087	5110
Total Obs.	12,690	12,402	12,402	11,771	10,036

Notes: The outcome variable of interest in this table is the log transfer price. All coefficients can be interpreted as approximate percentage changes. Column (1) corresponds to the model wtih properties 1/4 mile from a Green/Purple station. Column (2) is the 1/2 mile treated specification. Column (3) is the 1/2 mile specification with fixed effects. Column (4) is the 3/4 mile treated specification. Column (5) is the 1 mile treated specification. Clustered standard errors at the census tract level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.2 Neighbourhood Impacts - Results

The following figures are structured using three columns, with the outcome variable as log mean incomes at the census tract level. The first column represents treated tracts as those 1 mile from the metro, and control from 1-2 miles further. The second column has treated tracts as 1 mile from the metro, control as 1-3 miles further. The third column has treated tracts as 1.5 miles from the metro and control as 1.5-3 miles further.

According to Table 3 for all treatment distances in 2015, the effect is negative and non-significant. The point estimates for β_3 are -8.5%, -7.0% and -4.0% respectively per treatment group. The sign of these coefficients is surprising, and suggests that after the opening of the metro, individuals actually earned less, as measured in the 2015-2019 American Community Survey, than they did from 2010-2014. The statistically significant coefficient on the *Post* variable however, of 13.9%, 15.2% and 15.6% in each of the specifications, suggest that the source of the negative β_3 coefficient is an increase in the control tracts, rather than a decrease in incomes of the treated tracts.

Table 3: Impact of Treatment on Census Tract Log Mean Income - 2015

	(1)	(2)	(3)
Intercept	10.212***	10.494***	10.471***
	(0.100)	(0.106)	(0.108)
Treated	0.893***	0.794***	0.844***
	(0.080)	(0.066)	(0.062)
Post (post)	0.139***	0.152***	0.156***
	(0.040)	(0.029)	(0.032)
Treatment × Post (treated_post)	-0.085	-0.070	-0.040
	(0.059)	(0.053)	(0.046)
% Bachelors Degree	0.015***	0.012***	0.011***
	(0.003)	(0.003)	(0.003)
Tract FE	Yes	Yes	Yes
Year FE	No	No	No
R^2	0.98	0.98	0.98
Control Tracts	58	128	92
Treated Tracts	50	50	86
Total Obs.	108	178	178

Notes: The outcome of interest in this table is the log mean income, at the census tract level. All coefficients can be interpreted as approximate percentage change. Column (1) represents the specification with 1 mile treated tracts and 1-2 mile control. Column (2) is 1 mile treated, 1-3 mile control. Column (3) is 1.5 mile treated, 1.5-3 mile control. Clustered standard errors at the census tract level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

In Table 4 for all treatment distances in 2011, the effect is positive and statistically non-significant. The results show an increase of about 5.1%, 6.1% and 4.5% for the two 1 mile treatment distances and the 1.5 mile treatment distance. This suggests that there could be a slight distance decay effect as the treated buffer grows in size, indicating that census tracts closest to the metro corridor are increasing in income fastest. It is also possible that resident individuals of these tracts have found new sources of higher-paying employment. Notably, there could have been a broad recovery in incomes after the Great Recession that disproportionally affected those census tracts along the construction corridor, given that this pre-treatment period contains the worst portion of the financial crisis.

Table 4: Impact of Treatment on Census Tract Log Mean Income - 2011

	(1)	(2)	(3)
Intercept	10.284***	10.464***	10.407***
	(0.185)	(0.130)	(0.139)
Treated	0.990***	0.805***	0.857***
	(0.108)	(0.062)	(0.060)
Post (post)	0.051	0.033	0.027
	(0.053)	(0.033)	(0.038)
Treatment × Post (treated_post)	0.045	0.061	0.045
	(0.076)	(0.066)	(0.054)
% Bachelors Degree	0.013*	0.013***	0.013***
	(0.007)	(0.004)	(0.005)
Tract FE	Yes	Yes	Yes
Year FE	No	No	No
R^2	0.96	0.97	0.97
Control Tracts	58	128	92
Treated Tracts	50	50	86
Total Obs.	108	178	178

Notes: The outcome of interest in this table is the log mean income, at the census tract level. All coefficients can be interpreted as approximate percentage change. Column (1) represents the specification with 1 mile treated tracts and 1-2 mile control. Column (2) is 1 mile treated, 1-3 mile control. Column (3) is 1.5 mile treated, 1.5-3 mile control. Clustered standard errors at the census tract level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.3 Robustness Checks

4.3.1 Housing Capitalization - Uptown University Control

First, I have repeated the analysis of both the hedonic difference-in-differences regression on property prices using an alternative control group of metro lines that were not built. Going back to the 2003 referendum where Houston voters approved the construction of five additional lines, including the Uptown and University lines that were to be built in the west and northwest parts of the city. Although this referendum led to the construction of the Green and Purple lines, the Uptown and University lines were included in that package but not built. Thus, city officials recognized that these two corridors had the necessary characteristics, similar to the Green and Purple line, for building higher-order transit, which may suggest that common trends could hold. I have run the four quarter-mile difference-in-differences models for the housing capitalization specification using distances from these lines as the control group. I would expect to see positive results in housing price capitalization when using this setup, which would bolster confidence in the main regression results.

Appendix Table 12 presents the regression results. The results for the preferred specification in column (2), the 1/2 mile treated group shows an increase in property prices for treated groups after 2015 of 33.3%. This coefficient is significant only at the 0.1 level. This is a magnitude and sign similar to the main specification of 30.9%, which drives confidence in the results, as this new control group has little overlap with the original controls. However, it must be noted that if there are issues with the common trends assumption, this robustness check may not credibly override those concerns.

4.3.2 Housing Capitalization and Neighbourhood Effects - Matching

Due to the nature of the spatial distribution of economic activity in a city, certain socioeconomic characteristics may agglomerate within the region, which may prevent randomization and balance between treated and control areas. To account for these spatial differences between census tracts and neighbourhoods, I have used a matching approach for both the property data and census tract data to construct a well-balanced treated and control sample. For property data, I have constructed a series of coarsened variables, price decile, counts of bathrooms, bedrooms, decade built, and a coarse building area to match on broader characteristics. This provides a larger sample of matched pairs and selects control properties that correspond to treated properties on relevant variables. Similarly, I have constructed matching on census tracts by coarse, decile variables including unemployment, commuting time, educational attainment, and income. Once again, this provides a larger sample match between treated and control groups on relevant characteristics.

For housing capitalization, Appendix Table 13 demonstrates that the observable housing characteristics were successfully balanced between the treated and control groups. Appendix Figure 5 attempts to build confidence in the pre-trends for the matched group using an event study. Although the groups are balanced, it appears that a treated effect is more likely to occur in 2011, however data is not available in enough pre-treatment periods to analyze this hypothesis effectively. That being said, the pre-trends seem to be balanced except for the 2011 year. The event study shows that a negative outcome is likely. Appendix Table 14 presents the difference-in-differences statistics, where column 1 represents the 1/2 mile matched treated specification. Under matching, the coefficient on β_3 is negative and non-significant, suggesting that after opening the metro, these matched properties experienced a decline in price. This suggests that there are some unobserved spatial characteristics within the original specification which could be causing prices to rise at a high pace, and could indicate that control groups are actually experiencing spillover treated effects by their proximity to the metro as well. As a result, the resulting coefficients from the previous results should be interpreted as an upper bound for benefit, and should be interpreted with caution.

For neighbourhood characteristics, Appendix Table 15 demonstrates that the regression coefficient β_3 is negative with a value of -8.8% but is non-significant. This coefficient does not vary greatly from the original specification in 2015 which showed that after the opening of the Green and Purple lines, incomes fell in treated tracts from between 4.0% to 8.5%. This provides a higher degree of confidence in my earlier findings, and suggests there could be cause for the metro opening to result in factors that lowered the average incomes of the tracts adjacent to the newly opened metro lines.

5 Discussion

Public transit projects are often hailed as transformational endeavours by cities, sub-national and national level governments that have the power to reshape and kickstart growth and opportunities for local communities. At the same time, they are often contentious and require enormous sums of public and private funds to properly construct, and years of planning and construction to implement. Having a solid understanding of their benefits to individuals and communities is imperative if they are to be used as policy tools.

5.1 Housing Capitalization

Throughout this paper, I have shown that higher-order transit construction is plausibly capitalized into adjacent residential communities, reflecting an expectation of positive externalities such as decreased commute times, healthier neighbourhoods, higher quality of urban amenities that are expected to be realized from the projects. There tends to be a higher level of capitalization at the outset of construction and finalization of public funds, suggesting that housing market participants anticipate and price-in future expectations into current prices. This capitalization is further adjusted at the date of metro opening, suggesting that participants readjust their expectations once the final system is completed and true benefits are realized. In any case, it appears that those individuals who live closest to the corridor tend to experience in increase in residential housing wealth, suggesting that effective transit investment can uplift communities and drive wealth creation.

These coefficients and estimates must be interpreted with caution, however. Since there is evidence that parallel trends do not hold in the difference-in-difference analysis, there could be heavy bias in the results. This could result in higher than expected coefficients, as differences in pre-trends are then included into the difference-in-differences coefficient. In Figure 4 this is especially obvious, with a clear difference in treated and control pre-trends. Additionally, spatial spillovers, resulting from control properties experiencing effects after treatment of the treated properties is highly plausible. If there is a change in migration toward properties adjacent to metro stations for example, this could put upward pressure on treated tracts and simultaneous downward pressure on control tracts. Since construction occurred during these pre-2015 periods, this migration is reasonable, as people move away from the nuisance and noise, or toward it, in anticipation of future benefits. This would create an upward bias on the treatment effect.

However, within the results, the summation of the difference-in-differences coefficients β_3 and β_1 should account for this result by adding to a value less than β_3. For 2015 in Table 1 this summation is greater than β_3, suggesting that the control groups also rose sharply in price over time. Hence, there could be extreme spillovers in the selected treated specification that is obscuring the true result. To account for this, I performed an Honest Difference-in-Differences specification which qualified the coefficient to take on a value between about -10% and 30% depending on the degree of failure of parallel trends. This extreme range of possible values however, raises some doubt into nature of the capitalization effect. Contributing to this is that in constructing the treated definitions, it is impossible to know the true boundary, if one exists, where benefits of such a project begin and end. Although I relied on the previous literature to help inform these boundaries, it is likely that they vary city by city and even street by street.

That being said, the analysis of heterogeneity, conducted in Table 11, shows that houses of different value experience the effects of transit investment differently. Houses at the median, experience the highest degree of increase of 21.8%, with properties in the 75th percentile experiencing less increase, at 15.9%. This could suggest that wealthier individuals put less value on public transit access than other socioeconomic groups, and do not reap the same benefits. This is an important consideration when trying to quantify the value of transit investment.

The 2011 results may be more reasonable. It should be noted that the dataset used for my analysis begins at 2010, so it is challenging to test balance and pre-trends. Regardless, relating to the plausibility of pre-trends, there were no confounding projects relating to the metro, but the same concerns of migration and spillovers could introduce bias into the results. The direction of the property price changes moved as expected, for the metro to be capitalized into housing prices as it was built. Especially in the 1/2 mile treated specification, the statistically significant increase of 28.0% is in the expected direction as seen in Table 2. When adding the coefficient β_3 and the coefficient β_1 on the post variable, the resulting level increase in the treated properties is about 3.4% which is a reasonable increase in the period of time and corresponds to similar findings in the literature. This strongly suggests that residential market participants anticipated the future benefits of the light rail system, and priced these benefits into their property transactions, relative to houses that are further away.

5.2 Neighbourhood Effects

It is less apparent as to the income effects of the light rail construction. Although the light rail was expected to increase employment prospects for nearby residents, the results suggested that there could be no effect. These effects may take place over much longer periods of time, as neighbourhoods adjust and jobs relocate to this corridor which has relatively more pedestrians, and better accessibility and mobility to the rest of Houston. In all models and robustness checks, no statistically significant increase in incomes were observed.

In 2015, incomes were expected to increase after the metro opening, as individuals had access to greater employment opportunities, and possibly better paying jobs. It was also hypothesized that as more clients could access the metro corridor, business opportunities could increase, which could increase the level of employment and hence average incomes of these census tracts. That being said, these benefits are not represented in the -7.0% statistically non-significant decrease in incomes. It is possible that as a result of the system, negative externalities such as crime and noise increased, causing higher income earners to leave the neighbourhoods and seek work elsewhere, decreasing average earnings in the tract. It is also possible that individuals adjacent to the metro line opted to seek different employment opportunities that offered lower pay, but closer commutes with the metro line. Thus, these individuals could earn higher amounts per hour spent working and commuting, thus driving down average income, while returning leisure time to affected residents. In this case, this would not necessarily be a negative aspect, but would be represented as a negative coefficient. In any case, it is important to note that the lack of statistical significance could result in a positive true value of the coefficient.

For the 2011 treatment year, I did not expect to find any significant increase in mean incomes. The mobility and accessibility benefits of the metro cannot be realized during the construction phase of the project, and it is likely there would be a decrease in business activity along these lines. Thus, for mean incomes to increase during this time, the likely explanation could be due to migration of higher-income individuals to the tracts in anticipation of future benefits. The results in Table 4 weakly support this with an increase of 6.1% in the preferred specification. However, the lack of statistical significance does not give confidence to the result as the standard error is larger than the coefficient. All in all, these results are more likely to suggest that there is not much of an anticipation effect in incomes.

5.3 Conclusion

The work I have accomplished in this paper provides a foundation for future research into high-stakes public transit infrastructure investment. The results suggest that generally, there appears to be a capitalization effect that is distributed to residents after the investment of significant amounts of transit dollars. However, the benefits of these projects, especially those related to socioeconomic characteristics, may present themselves over very long time periods, as transit permanence may create a path dependency which dictates the spatial pattern of future neighbourhood development, business activity and commuting patterns. Unfortunately, pinning down the causal nature of these changes over such long time frames may be challenging.

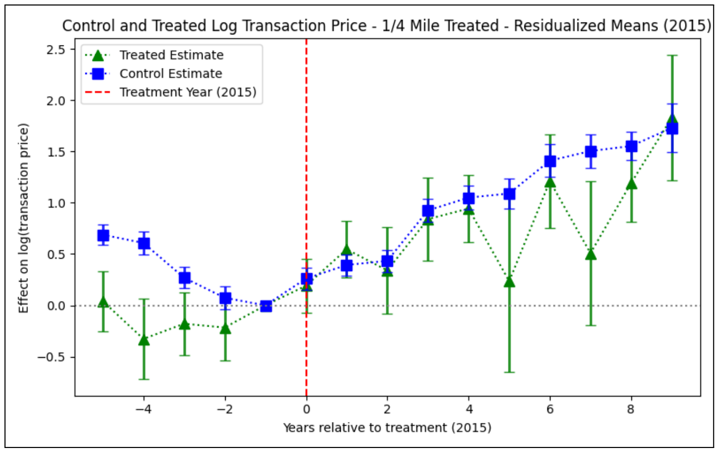
For future research, the use of more granular data, even at the administrative level, on the topics of individual employment, education and incomes, may create more feasibility in pinning down the outcome. Furthermore, the use of data relating to education, and family composition on an annual basis could form a basis for investigating whether these construction projects result in gentrification and out-migrating of existing residents. The occurrence of these forces is vital to understanding whether major public transit projects achieve the goals of local governments or simply result in implicit transfers between residents.

Therefore, the results in this paper suggest that transit may be an effective tool to raise local property values for nearby residents. However, the degree to which these benefits are very difficult to estimate, and may be heavily biased based on spatial spillovers.

A Appendices

A.1 Common Trends - 2015 Housing Characteristics

Figure 4: 2015 Treated/Control Trends



Notes: This figure plots residualized mean property values for treated and control tracts. Standard errors are clustered at the tract level. Treated properties were growing at a much lower rate in the pre-treatment periods, and after treatment, increased their rate of growth to match the control properties, less 2019 and 2021.

A.2 Balance on Observables - 2015 Housing Characteristics

Table 5: Household Characteristics Balance on Observables (2015) - 1/2 Mile Treated

Variable	Treated Mean	Treated SD	Control Mean	Control SD	t-statistic	p-value
Log Transfer Price	10.59	1.34	10.52	1.61	1.53	0.126
Bedrooms Count	2.75	1.05	2.82	1.00	-2.17	0.030
Bath Count	1.90	1.13	1.96	1.11	-1.69	0.091
Area Lot (sq. ft.)	5415.84	5022.05	6590.86	7793.94	-6.80	0.000
Area Building	1623.55	1096.16	1528.83	920.53	2.81	0.005
Decade Built	1956.43	30.31	1968.15	27.28	-12.49	0.000
Income Decile	3.40	1.72	3.72	2.46	-4.71	0.000

Notes: p-values from Welch two-sample t-test.

A.3 Balance on Observables - 2011 Housing Characteristics

Table 6: Household Characteristics Balance on Observables (2011) - 1/2 Mile Treated

Variable	Treated Mean	Treated SD	Control Mean	Control SD	t-statistic	p-value
Log Transfer Price	10.78	1.39	10.93	1.61	-2.23	0.026
Bedrooms Count	2.71	0.97	2.86	1.01	-3.13	0.002
Bath Count	1.92	1.03	2.07	1.17	-2.89	0.004
Area Lot (sq. ft.)	4987.91	2743.82	6251.57	7224.19	-7.11	0.000
Area Building	1624.90	779.87	1610.64	963.66	0.37	0.711
Decade Built	1956.76	29.91	1970.19	26.70	-9.59	0.000
Income Decile	3.53	1.80	4.14	2.66	-5.35	0.000

Notes: p-values from Welch two-sample t-test.

A.4 Balance on Observables - 2015 Neighbourhood Characteristics

Table 7: Census Characteristics Balance on Observables (2015) - 1 Mile Treated (1-3 Mile Control)

Variable	Treated Mean	Treated SD	Control Mean	Control SD	t-statistic	p-value
Log Mean Income	10.73	0.45	10.99	0.52	-2.37	0.021
Percent Bachelor's or Higher	15.16	14.51	29.27	26.59	-3.20	0.002
Unemployment Rate (%)	11.43	5.61	8.98	5.56	1.85	0.070
Mean Commute Time (min)	24.52	3.89	24.23	5.48	0.28	0.782

Notes: p-values from Welch two-sample t-test.

A.5 Balance on Observables - 2011 Neighbourhood Characteristics

Table 8: Census Characteristics Balance on Observables (2011) - 1 Mile Treated (1-2 Mile Control)

Variable	Treated Mean	Treated SD	Control Mean	Control SD	t-statistic	p-value
Log Mean Income	10.67	0.45	10.88	0.38	-1.80	0.079
Percent Bachelor's or Higher	13.75	12.91	25.83	25.66	-2.10	0.043
Unemployment Rate (%)	10.54	5.97	9.44	5.11	0.70	0.489
Mean Commute Time (min)	25.92	4.61	23.00	4.13	2.33	0.024

Notes: p-values from Welch two-sample t-test.

Table 9: Census Characteristics Balance on Observables (2011) - 1 Mile Treated (1-3 Mile Control)

Variable	Treated Mean	Treated SD	Control Mean	Control SD	t-statistic	p-value
Log Mean Income	10.73	0.45	10.99	0.52	-2.37	0.021
Percent Bachelor's or Higher	15.16	14.51	29.27	26.59	-3.20	0.002
Unemployment Rate (%)	11.43	5.61	8.98	5.56	1.85	0.070
Mean Commute Time (min)	24.52	3.89	24.23	5.48	0.28	0.782

Notes: p-values from Welch two-sample t-test.

Table 10: Census Characteristics Balance on Observables (2011) - 1.5 Mile Treated (1-3 Mile Control)

Variable	Treated Mean	Treated SD	Control Mean	Control SD	t-statistic	p-value
Log Mean Income	10.82	0.45	11.01	0.55	-1.80	0.076
Percent Bachelor's or Higher	20.46	20.91	29.84	27.02	-1.84	0.070
Unemployment Rate (%)	10.67	5.19	8.73	5.95	1.64	0.104
Mean Commute Time (min)	23.88	3.98	24.69	5.93	-0.76	0.450

Notes: p-values from Welch two-sample t-test.

A.6 Housing Capitalization - Quantile Regression

Table 11: Quantile Regression Results for Log Housing Prices (Preferred Specification)

	(1) $q = 0.25$	(2) $q = 0.5$	(3) $q = 0.75$
Intercept	25.497***	19.745***	8.564***
	(1.412)	(0.805)	(0.601)
Treatment × Post (treated_post)	0.204***	0.218***	0.159***
	(0.053)	(0.040)	(0.035)
Treated	0.551*	-0.529**	0.044
	(0.301)	(0.210)	(0.216)
Post	0.011	-0.002	0.025
	(0.062)	(0.047)	(0.042)
Bedrooms Count	0.050***	0.078***	0.060***
	(0.014)	(0.010)	(0.008)
Bath Count	0.061***	0.047***	0.030***
	(0.016)	(0.011)	(0.010)
Area Lot (1000s sq. ft.)	0.011***	0.012***	0.010***
	(0.002)	(0.001)	(0.001)
Area Building (1000s sq. ft.)	0.224***	0.273***	0.332***
	(0.018)	(0.011)	(0.008)
Year Built	-0.008***	-0.005***	0.001***
	(0.000)	(0.000)	(0.000)
Dist. to Downtown (1000 ft.)	0.115***	0.092***	0.057**
	(0.038)	(0.026)	(0.023)
Dist. to Medical Ctr. (1000 ft.)	-0.015	0.063**	0.006
	(0.043)	(0.028)	(0.024)
Dist. to Uptown (1000 ft.)	-0.059**	-0.112***	-0.035**
	(0.026)	(0.017)	(0.015)
$Pseudo\ R^2$	0.36	0.37	0.41
Obs.	12,402	12,402	12,402

Notes: This table shows the effect of the metro construction on different quantiles of housing transfer prices. There appears to be heterogeneous effects based on property value. Properties at the median of transfer price receive the most benefit of 21.8% increase, while properties at the 75th percentile experience a 15.9% increase. Robust standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.7 Housing Capitalization - Robustness - Uptown University

Table 12: Housing Price Capitalization by Treatment Distance - Uptown/Uni Control - 2015

	(1)	(2)	(3)	(4)	(5)
Intercept	4.267	10.646***	14.654***	12.100***	18.857***
	(3.877)	(3.621)	(3.154)	(4.235)	(4.465)
Treated	-0.492*	-0.830***	-0.648***	-1.061***	-0.782**
	(0.289)	(0.251)	(0.189)	(0.407)	(0.334)
Post (post)	0.056	0.392**	-0.395**	0.269*	0.343***
	(0.169)	(0.171)	(0.179)	(0.138)	(0.115)
Treatment × Post (treated_post)	0.094	0.333*	0.527***	0.317**	0.142
	(0.174)	(0.181)	(0.123)	(0.155)	(0.131)
Bedrooms Count	0.034	0.092*	0.024	-0.007	-0.019
	(0.066)	(0.054)	(0.036)	(0.048)	(0.037)
Bath Count	0.201***	0.157***	0.122***	0.173***	0.174***
	(0.057)	(0.047)	(0.039)	(0.047)	(0.046)
Area Lot (1000s sq. ft.)	-0.019	0.002	0.001	0.004	-0.003
	(0.016)	(0.014)	(0.014)	(0.012)	(0.011)
Area Building (1000s sq. ft.)	0.280***	0.272***	0.228***	0.323***	0.396***
	(0.073)	(0.055)	(0.048)	(0.053)	(0.055)
Year Built	0.003*	0.001	-0.001	0.000	-0.004
	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Distance to Downtown (1000 ft.)	-0.011	-0.099***	0.015	-0.088***	-0.060**
	(0.031)	(0.025)	(0.111)	(0.027)	(0.024)
Distance to Medical Centre (1000 ft.)	0.019	0.016	-0.033	-0.025	-0.029
	(0.034)	(0.020)	(0.125)	(0.024)	(0.023)
Distance to Uptown (1000 ft.)	-0.028	-0.058***	-0.014	-0.044*	-0.049**
	(0.029)	(0.019)	(0.138)	(0.024)	(0.022)
Year FE	Yes	No	Yes	No	No
Tract FE	No	No	Yes	No	No
R^2	0.45	0.41	0.57	0.37	0.32
Control Obs.	578	911	911	1320	1810
Treated Obs.	591	1695	1695	3087	5110
Total Obs.	1,169	2,606	2,606	4,407	6,920

Notes: The outcome of interest in this table is the log transfer price for properties. All coefficients can be interpreted as approximate percentage changes. Column (1) corresponds to the model with treated properties as 1/4 mile from the Green/Purple line, and control properties are 1/4 mile from the unbuilt Uptown/University line. Each subsequent column adds an additional quarter mile treatment/control distance. Column (3) represents the 1/2 mile treated group, including fixed effects. Clustered standard errors at the census tract level in parenthes. The coefficient on column (2) of 33.3% gives confidence to the values in the original specification, but the magnitude should be interpreted with caution. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

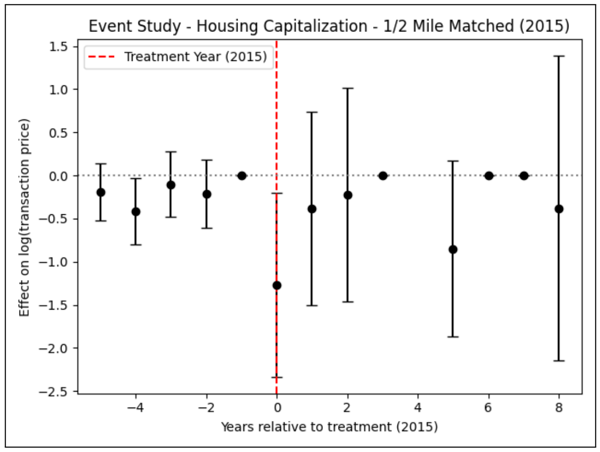
A.8 Housing Capitalization - Robustness - Matched Samples

Table 13: Pre-Treatment Balance on Property Characteristics - 1/2 Mile Treated (2015 Matched Sample)

Variable	Treated Mean	Treated SD	Control Mean	Control SD	t-statistic	p-value
Log Transfer Price	10.51	1.34	10.53	1.45	-0.33	0.739
Log Transfer Price Decile	5.88	2.59	5.89	2.60	-0.10	0.920
Bathrooms	1.74	0.94	1.76	0.94	-0.46	0.648
Bedrooms	2.67	0.76	2.69	0.75	-0.40	0.692
Decade Built	1958.41	28.53	1959.34	28.94	-0.65	0.516
Building Area (Coarse, sq ft)	1677.03	570.40	1686.39	587.58	-0.33	0.744

Notes: The coefficients in this table show that the housing characteristics between the treated and control groups are almost perfectly balanced, suggesting that the two samples are comparable on observables. P-values from Welch two-sample t-test.

Figure 5: 2015 Property Prices - Event Study - Matched Sample



Note: This figure reports event study coefficients for properties matched on observed characteristics. It appears that the true treatment effects occur after 2011, the funding date, rather than the opening date. Negative effects may be expected after the 2015 date. Standard errors are clustered at the census tract level.

Table 14: Housing Price Capitalization - Matching - 2015 Treatment Year

	(1)	(2)
Intercept	17.737***	25.303***
	(4.572)	(4.323)
Treated	-0.068	
	(0.159)	
Post (post)	0.629**	1.298
	(0.260)	(1.129)
Treatment × Post (treated_post)	-0.086	-0.656
	(0.336)	(0.542)
Bedrooms Count	-0.156**	0.040
	(0.070)	(0.071)
Bath Count	0.015	-0.071
	(0.071)	(0.074)
Area Lot (1000s sq. ft.)	-0.003	0.004
	(0.010)	(0.008)
Area Building (1000s sq. ft.)	1.099***	0.685***
	(0.118)	(0.097)
Year Built	-0.004	-0.006***
	(0.002)	(0.002)
Distance to Downtown (1000 ft.)	-0.007	0.153
	(0.018)	(0.117)
Distance to Medical Centre (1000 ft.)	0.005	0.135
	(0.018)	(0.132)
Distance to Uptown (1000 ft.)	-0.081***	-0.373***
	(0.014)	(0.112)
Year FE	No	Yes
Tract FE	No	Yes
R^2	0.28	0.54
Control Obs.	845	
Treated Obs.	788	
Total Obs.	1,633	1,633

Notes: Under this matching scheme, the treated_post coefficient becomes negative and statistically non-significant. This suggests that the metro opening does not have an effect on housing price capitalization, or even a negative effect. The summation of the post and treated_post coefficients suggest that the cause of this change is due to the increase in prices of control properties at a faster rate than treated properties. Clustered standard errors at the census tract level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.9 Neighbourhood Effects - Robustness - Matched Sample

Table 15: Impact of Treatment on Census Tract Log Mean Income - Matching - 2015

	(1)
Intercept	10.093***
	(0.078)
Treated	1.118***
	(0.127)
Post (post)	0.159***
	(0.037)
Treatment × Post (treated_post)	-0.088
	(0.060)
% Bachelors Degree	0.013***
	(0.004)
Tract FE	Yes
Year FE	No
R^2	0.98
Control Tracts	38
Treated Tracts	50
Total Obs.	88

Notes: This table presents the regression coefficients for the matched sample where treated tracts are those that are 1 mile from the metro lines. The variable of interest is log mean income. Similar to the baseline specifications in the 2015 year, the coefficient is negative and non-significant. This suggests that after the metro opening, incomes fell in treated tract. Clustered standard errors at the census tract level in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

B Works Cited

Attom Data. (2023). Property Assessed Value Data [Data set]. Attom Data. https://www.attomdata.com/data/property-valuation-data/assessor-market-value/

Attom Data (2024). Recorder Data [Data set]. Attom Data. https://www.attomdata.com/data/transactions-mortgage-data/recorder-data/.

Houston-Galveston Area Council. (July 2025). Metro LRT Stations [Data set]. H-GAC Regional GIS Data Hub. https://gishub-h-gac.hub.arcgis.com/datasets/H-GAC::metro-lrt-stations/about

Keeler, Z.T., Stephens, H.M. (2022). The capitalization of metro rail access in urban housing markets. Real Esate Economics 51(3) 533-778. https://doi.org/10.1111/1540-6229.12411

Metropolitan Transit Authority of Harris County, Texas (METRO). (August, 2003). Notice of Special Election, METRO. Metropolitan Transit Authority of Harris County, Texas. Retrieved from https://metro.resourcespace.com/pages/view.php?ref=17997&k=

Pan, Q. (2019). The impacts of light rail on residential property values in a non-zoning city: A new test on the Houston METRORail transit line. Journal of Transport and Land Use. https://doi.org/10.5198/jtlu.2019.1310

Samuel, A. (December 8, 2011). $900m awarded to extend Houston's light rail system. Rail.co. Retrieved from https://web.archive.org/web/20120606145813/http://www.rail.co/2011/12/08/900m-awarded-to-extend-houstons-light-rail-system/

U.S. Census Bureau, U.S. Department of Commerce. "Commuting Characteristics by Sex." [Data set] American Community Survey, ACS 5-Year Estimates Subject Tables, Table S0801, https://data.census.gov/table/ACSST5Y2023.S0801?t=Commuting&g=050XX00US48201$1400000.Accessedon23Jul2025.

U.S. Census Bureau, U.S. Department of Commerce. "Educational Attainment." American Community Survey, ACS 5-Year Estimates Subject Tables, Table S1501, https://data.census.gov/table/ACSST5Y2023.S1501?t=Education&g=050XX00US48201$1400000.Accessedon23Jul2025.

U.S. Census Bureau, U.S. Department of Commerce. (n.d.). Income in the Past 12 Months (in 2023 Inflation-Adjusted Dollars). American Community Survey, ACS 5-Year Estimates Subject Tables, Table S1901. Retrieved July 23, 2025, from https://data.census.gov/table/ACSST5Y2023.S1901?t=Income+and+Poverty&g=050XX00US48201$1400000.

U.S. Census Bureau. ”2020 TIGER/Line Shapefiles” [Data set]. U.S. Cenus Bureau. https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2020.html#list-tab-790442341

U.S. Census Bureau. ”2010 TIGER/Line Shapefiles” [Data set]. U.S. Cenus Bureau. https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.2010.html#list-tab-790442341

U.S. Department of Transportation (USDOT). (2010). University Corridor Fixed Guideway in Houston, Harris County, Texas Final Environmental Impact Statement and Section 4(f) Evaluation. U.S. Department of Transportation, Federal Transit Administration and Metropolitan Transit Authority of Harris County, Texas. Retrieved from: https://web.archive.org/web/20120104231632/http://ridemetro.org/CurrentProjects/pdfs/FEIS%202010/University-Corridor-FEIS-Vol-1-Pt1.pdf