How New Transit Infrastructure Shapes Housing Markets: A Case Study of Berryessa BART Station

Aiden Y.

Abstract

This study examines the impact of the Berryessa BART station on housing prices in San Jose amid the city's growing affordability crisis. Utilizing housing market data from Redfin and Zillow, I will use a Two-Way Fixed Effects (TWFE) regression model to compare housing price trends within a 3-mile radius of the station (treatment group) to similar unaffected areas (control group). This model will be created and account for how factors like distance to transit station, property size, number of bedrooms, and bathrooms influence housing values. While existing research shows that new transit stations generally increase property values and demand, they can also displace lower-income residents and widen socioeconomic gaps. However, the specific short-term and long-term effects of the Berryessa BART station within San Jose's Berryessa neighborhood's environment remain underexplored. This study aims to fill this gap by analyzing how the Berryessa BART station, introduced in June 2020, affects housing prices, availability, and displacement patterns in San Jose. The findings will support urban planning strategies that balance transit development with housing equity and affordability, promoting sustainable and inclusive growth in the region.

Definitions and Assumptions

Based on prior research that finds that effects tend to diminish after 2-3 miles, The study assumes that houses beyond 3 miles from the Berryessa BART station have their housing prices unaffected (Mathur, 2020) and that broader economic conditions, such as inflation and employment, remain relatively stable throughout the study period.

Two-Way Fixed Effect Regression Model

A Two‑Way Fixed Effects (TWFE) regression model controls for unobserved factors both across units (e.g., neighborhoods, firms, individuals) and over time (e.g., years, quarters). In this study, the model can be written as $ln(Price) = \beta_0 + \beta_1 ln(distance) + \beta_2 Beds + \beta_3 Baths + \beta_4 Sqft + \beta_5 Treated$

Literature Review

California Transit Village Movement

In his 1996 study, Cervero illustrates how mixed‑use neighborhoods clustered around BART, Caltrain, and light‑rail stops in the Bay Area are greatly affected by transit stations. Between 1985 and 1994, over 6,500 housing units were built within a quarter‑mile of stations at densities of 20-60 housing units per acre, supporting strong ridership while attracting mostly young working-class men who rely on rail for commuting and leisure (Cervero, 1996). Evenpark‑and‑ride‑lot "transit villages" such as the 300‑unit Villages of La Mesa adjacent to San Diego Trolley's Amaya Station leased nearly 100% upon opening, suggesting a strong demand for transit‑oriented living in locations (Cervero, 1996).

Charlotte's Light-Rail System

In a study on Charlotte's public transportation effects on housing prices, Billings found that single-family homes within one mile of newly opened LYNX Blue Line stations increased by 4.0% relative to control neighborhoods, while condominiums increased by 11.3% over seven years (Billings, 2011). These station-area price increases continued even after accounting for broader market trends and other potential factors that affect the price. Commercial properties, however, did not exhibit statistically significant gains, suggesting that in Charlotte's relatively lower-density environment, light rail functioned more as a neighborhood commodity rather than the direct reason that prices are changing (Billings, 2011).

Comparing Housing Prices in Other Cities

In his 2020 study, Mathur applied quantile regressions on single-family home sales within five miles of the Warm Springs BART Extension in South Fremont. He observed that house prices began rising more than a decade before the extension's completion, showing strong anticipation among buyers and developers (Mathur, 2020). Notably, lower-priced homes experienced the largest relative increases in value, suggesting that early speculation and investor bundling of "transit-adjacent" housing can disproportionately affect the most affordable market segments (Mathur, 2020). While some homeowners capitalized on increased equity, others faced heightened cost burdens and the risk of displacement as local affordability stopped happening (Mathur, 2020).

In Athens in 2013, Efthymiou and Antoniou examined a period of network expansion that included metro, tram, and suburban-rail stations. By finding listings for more than 16,000 properties and using various regression models, they found that proximity to metro and tram stops was positively correlated with both sale and rental prices, whereas proximity to the older ISAP urban rail line, the airport, or major port facilities created a negative effect due to noise and other negatives (Efthymiou & Antoniou, 2013). Their analysis shows that not all rail modes uniformly enhance property values. Modern, high-frequency systems like a metro tend to boost housing demand, while noisier or lower-capacity corridors like freight-oriented lines can decrease adjacent property prices (Efthymiou & Antoniou, 2013).

Bart Ridership and Housing

In 2023, Wasserman and Taylor found that BART ridership is very much concentrated at the peak hours because of Oakland and San Francisco's larger job market, which strains the resources of the local community and creates overcrowding (Wasserman and Taylor 2023). This growth has also been rapidly increasing, increasing upwards of 25-30% in the 2010-2020 decade, targeting mostly high-income riders. Meanwhile, the off-peak commuting has not changed quite as much, with there being a lack of growth in the last decade. This trend, for the lower-income riders who often relied on BART for affordable transportation, was no longer able to afford the nearby housing and were displaced into neighborhoods to the east, which increased commute times for those residents. In some parts, the housing units adjacent to the BART Stations

Thematic Synthesis

Across these diverse settings, several themes show up: first, single-family homes within close proximity to new rail stations tend to have increases in prices, from about 4% to 11% (Billings, 2011; Mathur, 2020; Efthymiou & Antoniou, 2013). Second, modern, high-capacity rail systems (metro, light rail) tend to positively impact nearby property values, while legacy rail lines or noisy corridors (e.g., freight or older commuter rail) can lower values due to noise and safety concerns (Efthymiou & Antoniou, 2013). Third, in some markets, the price increase begins years before station openings, which is because of marketing of the housing before the opening and buyer speculation (Mathur, 2020). Finally, improved transit access often means higher rents that displace lower-income households, as seen in both Athens and the Bay Area (Efthymiou & Antoniou, 2013; Wasserman & Taylor, 2023). In summary, there are many complex ways that new transit influences nearby housing prices, and as a result, policymakers must carefully balance the advantages, such as higher ridership, reduced car reliance, and increased property-tax revenue, against potential downsides like noise, traffic congestion, and the risk of displacing existing residents.

Gap in Research

Although substantial research has examined how public transport affects housing prices in various cities, there do not exist studies that examine the specific impacts of the Berryessa BART station on San Jose’s real estate market. Most existing work considers general transit-related trends across multiple cities but doesn’t cover the distinctive features of San Jose, including a tech-driven economy, severe housing shortages, and rapidly rising living costs. Furthermore, the Berryessa station opened during changing patterns, growing remote work due to COVID, widening income gaps, and changing travel behaviors, which complicate the situation. As a result, investigations on how this new transit station influences housing values are useful for policymakers and developers and help us understand trends that may affect other locations. Such insights are crucial for crafting urban policies that promote transit development while allowing affordability and equity.

Methodologies

Data collection was done using a set of databases, including Redfin, Zillow, and Google Maps, after selecting 2 similar neighborhoods within 3 miles of the newly introduced BART station and randomly selecting 25 similar houses from each neighborhood. A similar method was used for the control group, by using Zillow to identify 2 neighborhoods outside of the 3-mile radius that had similar characteristics to the first 2 selected, such as housing prices, number of bedrooms and bathrooms, and property size. Google Maps was then used to calculate the driving distance to the nearest BART station from each property before and after the introduction of the Berryessa BART station, being the Warm Springs/South Fremont station before and the Berryessa station after. Redfin was then used to find the housing prices of the selected houses during March 2020, March 2022, and March 2024. Statistical analysis was then performed to create a basic data summary of the variables and was used to create a regression model with the TWFE method. A significance level test was then used to then determine the significance of the distance to the nearest BART station. One limitation of this was the low sample size, since only 100 houses were selected, making it not representative of the whole city.

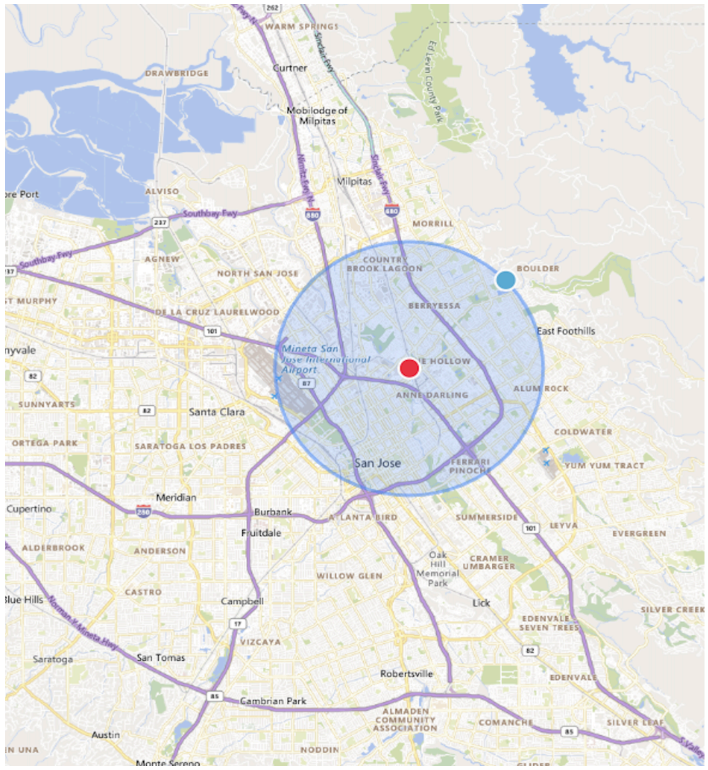

Figure 1: Map of a 3-mile radius around the Berryessa BART station

This circle identifies treatment groups and control groups. Treatment group subjects are located within the blue circle (3 miles of the station), and the control groups, which are not affected by the new introduction of the station, are outside of the circle.

Data

Basic summary statistics for all observations, including all observations from March 2020 (-3 months), before the new Bart station, all observations from March 2022 (+21 months), and March 2024 (+45 months). The column under (1) means the columns in the treatment group (within 3 miles of the new station), the column under (2) is the control group, a similar neighborhood outside of the 3-mile radius. Column (3) combines both sets of data.

Table 1: Summary Statistics for Treated and Control Groups

	Group 1		Group 2		All Samples	
	Mean	SD	Mean	SD	Mean	SD
Price (thousands)	1314.08	254.84	1365.05	305.59	1339.56	282.05
Distance (miles)	7.70	4.28	20.61	4.59	14.15	7.84
Beds	3.58	0.61	3.36	0.48	3.47	0.56
Baths	1.88	0.55	2.02	0.32	1.95	0.46
Sq ft	1333.26	292.82	1479.42	226.93	1406.34	271.57
Observations	150		150		300	

Basic summary statistics of the observations from March 2020 (-3 months). The column under (1) means the columns in the treatment group (within 3 miles of the new station), the column under (2) is the control group, a similar neighborhood outside of the 3-mile radius. Column (3) combines both sets of data.

Table 2: Summary Statistics for 2020

	Group 1		Group 2		All Samples	
	Mean	SD	Mean	SD	Mean	SD
Price (thousands)	1293.80	274.77	1341.90	302.00	1317.85	288.26
Distance (miles)	10.70	0.51	23.57	1.65	17.136	6.58
Beds	3.58	0.61	3.36	0.48	3.47	0.56
Baths	1.88	0.56	2.02	0.32	1.95	0.46
Sq ft	1333.26	294.80	1479.42	228.47	1406.34	272.48
Observations	50		50		100	

Basic summary statistics of the observations from March 2022 (+21 months). The column under (1) refers to the treatment group (within a 3-mile radius of the new station), while the column under (2) represents the control group, a similar neighborhood located outside the 3-mile radius. Column (3) combines both sets of data.

Table 3: Summary Statistics for 2022

	Group 1		Group 2		All Samples	
	Mean	SD	Mean	SD	Mean	SD
Price (thousands)	1319.70	242.49	1375.02	295.64	1347.36	270.44
Distance (miles)	1.70	0.40	14.68	2.19	8.19	6.71
Beds	3.58	0.61	3.36	0.48	3.47	0.56
Baths	1.88	0.56	2.02	0.32	1.95	0.46
Sq ft	1333.26	294.80	1479.42	228.48	1406.34	272.48
Observations	50		50		100	

Basic summary statistics of the observations from March 2024 (+45 months). The column under (1) refers to the treatment group (within a 3-mile radius of the new station), while the column under (2) represents the control group, a similar neighborhood located outside the 3-mile radius. Column (3) combines both sets of data.

Table 4: Summary Statistics for 2024

	Group 1		Group 2		All Samples	
	Mean	SD	Mean	SD	Mean	SD
Price (thousands)	1328.74	250.01	1378.22	323.32	1353.48	288.61
Distance (miles)	1.70	0.40	14.68	2.19	8.19	6.71
Beds	3.58	0.61	3.36	0.48	3.47	0.56
Baths	1.88	0.56	2.02	0.32	1.95	0.46
Sq ft	1333.26	294.80	1479.42	228.47	1406.34	272.48
Observations	50		50		100	

Empirical Strategy and Results

Data Groups

For the periods of which data was collected, there were 100 observations from March 2020 (3 months before the opening of the Berryessa BART station), March 2022 (21 months after the opening of the Berryessa BART station), and March 2024 (45 months after the opening of the Berryessa BART station).

There are two types of groups in this study: the treatment group, which consists of single-family homes located within a 3-mile radius of the new Berryessa BART Station, and the control group: a set of neighborhoods beyond that 3-mile radius whos prices remain "unaffected" by the new station.

TWFE Regression

We stack every home-sale observation at each estimate window (March 2020, March 2022, March 2024) into one pooled panel. We model it with this equation:

$$ln(Price) = \beta_0 + \beta_1 ln(distance) + \beta_2 Beds + \beta_3 Baths + \beta_4 Sqft + \beta_5 Treated$$

ln_price	Coefficient	Std. err.	t	P>|t|
ln_distance	.0224868	.0185803	1.21	0.227
beds	.0349798	.0239312	1.46	0.145
baths	.0032131	.032307	0.10	0.921
sqft	.0002459	.0000566	4.34	0.000
treated	.0251553	.0352324	0.71	0.476
_cons	6.637005	.1059252	62.66	0.000

Table 5: Regression Results of TWFE

Discussion

In this scenario, the only coefficient that reaches conventional levels of statistical significance is the effect of square footage, with a coefficient of 0.0002459 and a p-value of < 0.001. In other words, each additional square foot of living area is associated with a roughly 0.0246% increase in sale price. Over several hundred square feet, this incremental effect compounds into a meaningful price increase for larger homes.

By contrast, the coefficient on the treatment indicator of 0.0251553, including a p value of 0.476, which captures whether a home lies within three miles of the new Berryessa BART station, does not differ significantly from zero in this pooled snapshot. In plain terms, once we control for distance (with the natural log), bedroom count, bathroom count, and square footage, the "treated" homes exhibit only a 2.5% higher log price than "control" homes, and the standard error is too large to rule out a null effect.

The distance coefficient of 0.0224868 with a p-value of 0.227 also fails to get statistical significance. A positive sign on that term would normally imply that homes farther from the station mean higher prices, an economically counterintuitive result, yet the magnitude is small enough (about a 2.2% change in price per 1% change in distance) and so imprecise that we cannot conclude.

Conclusion

This study set out to evaluate how the opening of Berryessa BART in June 2020 has influenced nearby single-family home prices in San Jose's Berryessa neighborhood. By comparing 100 properties within a 3-mile radius of the new station ("treated" group) to 100 similar homes just beyond that boundary ("control" group) across three snapshots, March 2020 (pre-opening), March 2022 (21 months post-opening), and March 2024 (45 months post-opening), and tested whether proximity to the station translated into statistically significant price gains when controlling for dwelling size, bedroom and bathroom counts, and distance to transit.

In this regression covering March 2020, March 2022, and March 2024, the "Treated" coefficient of 0.0251553 (SE = 0.0352324, p = 0.476) is not statistically different from zero. Likewise, the ln(Distance) coefficient of 0.0224868 (s.e. = 0.0185803, p = 0.227) and the Bedrooms coefficient of 0.0349798 (s.e. = 0.0239312, p = 0.145) fail to reach significance, as does the Bathrooms coefficient of 0.0032131 (s.e. = 0.032307, p = 0.921). The only variable that remains highly significant is square footage, with a coefficient of

0.0002459 per sq ft (s.e. = 0.0000566, $p < 0.001$). This indicates that, once we control for size, bedroom/bathroom counts, and proximity, there is no significant BART‑related price increase in this period. In other words, within this study, variations in house size explain nearly all of the price differences, and the three‑mile "treated" designation shows no statistically significant effect on ln price.

In conclusion, this analysis concludes that Berryessa BART has not yet conferred a statistically significant increase on nearby single-family home prices once house size and other basic features are controlled and accounted for. Policymakers also need to account for the widespread use of driving that many residents in the San Jose region use, as that could discourage them from using public transport, and as a result, outweigh the benefits of having new stations. By doing so, San Jose can maximize the benefits of Berryessa BART for riders, residents, and the region's long-term sustainability.

References

Billings, S. B. (2011). Estimating the value of a new transit option. *Regional Science and Urban Economics*, *41*(6), 525-536. https://doi.org/10.1016/j.regsciurbeco.2011.03.013

This research explores the impact of light-rail transit (LRT) on property values in Charlotte, North Carolina. It argues that both price gradients to transit stations and overall neighborhood trends must be considered to assess LRT's effects. The study found that LRT increased single-family property values by 4% and condominium values by 11.3% within one mile of stations. However, it had no significant effect on commercial properties. The results suggest that LRT serves more as a tool for economic development in specific neighborhoods rather than just for transportation. The author does research on public policy with the University of North Charlotte. The research seems to be objective, and the target audience is to people of Charlotte or to economists. The source was only a little helpful, as it doesn't have too much to do with Bay Area, but I can use it as inspiration

Mathur, Shishir. "Impact of transit stations on house prices across entire price spectrum: A quantile regression approach." *Land Use Policy* 99 (2020): 104828 https://doi.org/10.1016/j.landusepol.2020.104828

This study investigated the effects of the Warm Springs Bart Station on house prices in Fremont, using data from 2000 to 2018. The research found that house prices increased quite a bit within 8 kilometers of the station. However, price increases began about a decade before the station's opening. This suggests that public transport impacts housing prices in the area. The author is a professor and researcher in urban and regional planning at SJSU, and the target audience is probably just people in the Bay region. The content is objective and the text is readable. It will definitely be useful, since the Warm Springs station is the closest station to the ones I am trying to research.

Wasserman, J. L., & Taylor, B. D. (2023). State of the BART: Analyzing the determinants of Bay Area Rapid Transit Use in the 2010s. *Transportation Research Part A: Policy and Practice*, *172*, 103663. https://doi.org/10.1016/j.tra.2023.103663

This article discusses the increasing peaking problem in the Bart system during the 2010s, where the ridership was heavily concentrated within peak hours. Off-peak trips were very empty. Many factors influenced this such as station crowdedness, driving traffic, etc. This study highlights the financial strain on the US public transport system. The authors are all transportation experts, which provides credibility. The target audience appears to be policymakers, transit planners, and residents. This source is reliable, as it uses data to be objective and the focus on BART is relevant to my studies

Cervero, R. (1996). California's transit village movement. *Journal of Public Transportation*, *1*(1), 103-130. https://doi.org/10.5038/2375-0901.1.1.6

This artcle was mostly talking about transit villages. Transit villages are like dense mixed communities around a rail station which could potentially reduce car use. California has tried this in places like El Ceritto, and Pleasant Hill. It also used examples from the Bay Area and also compared this to Europe. The author is a professor in transportation at Berkeley. The information seems to be objective and the target audience appears to be urban planners or developers, so it's not quite targeted towards me. The source was somewhat helpful to me, as there is some data that I could use.

Efthymiou, D., & Antoniou, C. (2013). How do transport infrastructure and policies affect house prices and rents? Evidence from Athens, Greece. *Transportation Research Part A: Policy and Practice*, *52*, 1-22. https://doi.org/10.1016/j.tra.2013.04.002

This paper uses hedonic price models and spatial econometrics to explore transport infrastructure's effects on real estate values in Athens. The study finds that proximity to metro, tram, suburban railway, and bus stations increases prices, while older rail systems and airports tend to lower them due to noise and such. The results provide insights for urban planning and policy design. This is relevant to the research being done, as it includes possible factors to consider.

Peng, Y., Tian, C., & Wen, H. (2021). How does school district adjustment affect housing prices: An empirical investigation from Hangzhou, China. *China Economic Review*, *69*, 101683 .https://doi.org/10.1016/j.chieco.2021.101683

This study investigated the effects of educational quality on housing prices, particularly following a school district adjustment in Hangzhou in 2012. Using a difference-in-differences (DID) approach combined with quantile regression, the authors reveal significant price differences for houses assigned to higher-quality schools, particularly for low-priced and small houses. The paper addresses issues like endogeneity and highlights the effects across housing markets. This research is relevant for making policies regarding education equity and housing market regulation. This paper is relevant to my studies because it provides another example of how I can conduct my research.

Levkovich, O., Rouwendal, J. & van Marwijk, R. The effects of highway development on housing prices. *Transportation* 43, 379–405 (2016). https://doi.org/10.1007/s11116-015-9580-7

This study investigated how highway development projects in the Netherlands affected housing prices. The authors analyze the impact of two newly developed highways using repeat sales and a Difference-in-Difference approach, focusing on three transportation-related effects: accessibility, traffic, and noise pollution. They used a large housing transaction dataset from the Dutch Land Registry, and while the findings found that improved accessibility increases housing prices, the extra traffic and noise led to price decreases.

De Chaisemartin, C., & D'Haultfœuille, X. (2023). Two-way fixed effects and differences-in-differences estimators with several treatments. *Journal of Econometrics*, *236*(2), 105480. https://doi.org/10.1016/j.jeconom.2023.105480

TWFE regressions with multiple treatments generally produce each coefficient as a weighted sum of that treatment's own effects plus a "contamination" term from the other treatments whenever their effects are heterogeneous, leading to potentially large bias. The authors propose a new DID estimator that matches groups switching on a given treatment to controls whose other treatments remain constant, thereby eliminating contamination but often reducing sample size and precision. However, in an application to daycare regulations, the robust estimator differs significantly from the TWFE estimates.