Title

Beyond Density: Examining the Relationship Between Station-Area Establishments and Rail Transit Ridership

ISBN

9798252417646

Author

Hua, Michael

Publication Date

2026-06-12

UNIVERSITY OF CALIFORNIA

Los Angeles

Beyond Density:

Examining the Relationship Between Station-Area Establishments and Rail Transit Ridership

A thesis submitted in partial satisfaction of the

requirements for the degree

Master of Urban and Regional Planning

by

Michael Hua

2026

ABSTRACT OF THESIS

Beyond Density:

Examining the Relationship Between Station-Area Establishments and Rail Transit Ridership

by

Michael Hua

Master of Urban and Regional Planning

University of California, Los Angeles, 2026

Professor Michael K. Manville, Chair

Numerous studies have explored the relationship between built environment characteristics and transit ridership. However, little research has tested whether the number or size of station-adjacent commercial establishments have an independent association with rail ridership. Using negative binomial regression, I analyze the association between rail station alightings and the number of commercial establishments in three size categories: less than 700 square feet, 701–2,000 square feet, and 2,001–5,000 square feet. I also test whether these relationships vary by distance from the rail station (within a quarter-mile and a half-mile buffers) and between weekdays and weekends.

I find no statistical significance between the number of small and medium-large commercial establishments within a half-mile of a station and rail alightings. A marginal association is observed between medium establishments and rail alightings at the half-mile on weekends, though the relationship is substantively weak and does not approach the consistency or magnitude of the small establishment finding at the quarter-mile. Within a quarter-mile rail station radius, I find a modest significant positive relationship between rail alightings and the number of commercial establishments smaller than 700 square feet. The relationship between the number of commercial establishments under 700 square feet and rail alightings strengthens further on Saturday and Sunday than on weekdays. The quarter-mile proximity and weekend ridership relationship small establishments have with rail alightings highlights avenues for future research on potential causal effects and the role of small establishments in transit-oriented development planning.

The thesis of Michael Hua is approved.

Evelyn A. Blumenberg

Brian D. Taylor

Michael K. Manville, Committee Chair

University of California, Los Angeles,

2026

Dedication

This thesis is dedicated to my family and friends who have loved and supported me throughout this process.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my UCLA professors for their endless support and encouragement. To my Chair Michael Manville, thank you for providing me with confidence throughout this project, answering all my questions, and guiding me throughout the entire process of shaping and formulating my thesis. To Evelyn Blumenberg, thank you for your kind help and attention to detail with the writing and organization of this thesis, and your expertise and guidance with developing my transportation analytical thinking skills. To Brian Taylor, thank you for thoughtful feedback and writing assistance through the process. I am very grateful for your transportation expertise in helping me refine my research skills.

I am truly grateful for all the transportation, writing, and research lessons I have received from each of you throughout my time at UCLA, and I will carry them with me throughout my career.

Introduction

Land surrounding urban rail transit stations is limited, and its use is an important determinant of transit ridership (Ewing and Cervero 2010). Transit-Oriented Development (TOD) is a tool to increase transit ridership and create pedestrian-friendly environments by encouraging mixed-use and compact development in a half-mile radius surrounding major bus and rail transit stations (Cervero et al. 2002). Beyond increasing development densities and mixing land uses, might there be other ways to develop TODs to increase rail transit ridership? This thesis examines this question by analyzing whether more, smaller station area commercial establishments might increase rail transit use compared to fewer, larger ones.

In denser commercial areas, people are more likely to take transit or walk when making within-neighborhood trips (Cervero and Kockelman 1997). However, not all commercial density is equivalent; a block occupied by a single large establishment and one divided among several small establishments may differ in how they generate transit ridership. Small establishments occupy less block space than large establishments, allowing a greater number and variety of commercial destinations per block. Could a greater number and variety of commercial destinations per block make an area more attractive to pedestrians and transit riders alike, as each additional establishment increases the likelihood that a visitor finds a reason to stop?

Small establishments tend to rely on passing foot traffic rather than car trips (Sevtsuk 2020), suggesting that the profitability of small stores is linked to the pedestrian activity that rail transit can provide. A symbiotic relationship between small establishments and rail ridership warrants examination, with small establishments relying on the foot traffic rail transit provides, and rail delivering passengers to areas with a large number of destinations.

Most rail transit in Los Angeles is operated by the Los Angeles County Metropolitan Transportation Authority (LA Metro) with two subway and four light rail lines. The system currently includes 108 stations running along 121 directional miles of track. The two heavy rail station lines are the B and D lines, and the four light rail lines are the A, C, E, and K lines. These rail stations are built around a diverse mix of land uses, from low density, single-family neighborhoods surrounding the E line's Westwood/Rancho Park station to high-intensity, downtown commercial areas adjacent to the A, B, D, and E line's 7th Street/Metro Center station. These contrasts in land use can help analysts identify the association between different built environment types, specifically the location of large and small establishments and transit ridership.

Figure 1. Westwood/Rancho Park station surrounded by single-family homes (Google Maps, 2026)

Figure 2. 7th Street/Metro Center station located in downtown Los Angeles (Google Maps, 2026)

Many planners and policymakers in Los Angeles have championed the need to build at higher densities around rail station areas (Metro 2018, Cal. S.B. 79, 2025). LA Metro has adopted a Transit Oriented Communities (TOC) policy, partnering with local communities to fund and incentivize denser land use, housing, and community development activities, as a method to increase transit ridership, strengthen local communities, and capture value created by the presence of transit (Metro 2018). In 2016, Measure JJJ, a City of Los Angeles ballot measure designed to require affordable housing in development projects, established a set of incentives targeting affordable housing near public transit (City of Los Angeles, 2016). Signed into law in 2025, California Senate Bill 79 requires cities to allow higher density units within a half-mile radius of qualifying transit stations (Cal. S.B. 79, 2025).

The purpose of this study is to understand the relationship between the number of small, customer-serving organizations or commercial establishments near rail stations and rail ridership. I estimate a station-level regression model using both point of interest (POI) data on commercial establishments and LA Metro ridership data to examine the relationship between small establishments located within a half and quarter-mile radius of rail stations and rail alightings[1].

[1] Alightings refers to the number of people disembarking a transit vehicle at a given station

Project Significance

Despite Los Angeles County spending billions of dollars on public transit capital and operating costs every year, transit ridership per capita has continued to fall since its 1985 peak, even as the rail network expands (Brozen et al. 2018). During the COVID-19 pandemic, transit ridership dipped further and has been slow to recover (Ziedan et al. 2023). Understanding the relationship between the size of station-area commercial establishments and transit use can better inform researchers and planners working to address low ridership.

Evolving retail trends post-pandemic may point to a rise in small-format stores. Post-pandemic, Los Angeles has experienced higher retail vacancy rates compared to pre-pandemic conditions (Hutson and Orlando 2023). Big box retailers are increasingly shifting toward smaller stores, which can produce higher revenues per square foot (Williams 2024). With a rise in online shopping, e-commerce, and evolving retail preferences, retailers are responding by opening more small-format stores over traditional big-box businesses (King 2023). The owners of smaller stores can focus their operations on more specific merchandise and consumers, allowing them to access more targeted demographics (Roman 2024). Although these stores do not reflect the 700 square footage scale level discussed in this thesis, the growing business trends related to store size make understanding store size's relationship to transit ridership increasingly important.

Outside of transit ridership, small commercial spaces can reduce business expenses and increase commercial occupancy rates (King 2023). In terms of total square footage, smaller commercial spaces have lower rents than larger ones, increasing cost savings[2].

The presence of commercial establishments can help improve local economic activity and street vitality. Smaller stores can have more reverberating effects on the local economy

[2] Lower total rents may increase the tenant demand for small commercial spaces, reducing total business vacancy rates.

than big-box stores, as small store revenue can be used to subcontract from local providers, hire local employees at higher wages, and improve public infrastructure around stores (Sevtsuk 2020). Higher building density, intersection density[3], transit stops, and accessible commercial facilities can promote recreational walking (Zeng et al. 2023). TOD station areas tend to have higher numbers of retail jobs than non-TOD station areas (Ganning and Miller 2020).

If small commercial spaces in a typical TOD station-area radius show a positive relationship with ridership, further investigation should explore aspects of that relationship beyond simply development density, such as whether increasing the supply of small commercial establishments might support transit ridership. A better understanding of the relationship between the number and size of station-area commercial establishments and rail transit use helps inform planners and public officials on how best to increase the effectiveness of TOD designs. The findings also may have implications for the types of incentives necessary to encourage developers to invest in the types of stores and restaurants that generate the most patronage and transit use.

[3] Intersection density is the number of street intersections within a defined area

Literature Review

This review of literature proceeds from the general to the specific: it begins with the broad determinants of transit ridership, narrows to the relationship between land use and transportation, and then focuses on how station-area built environment characteristics relate to rail ridership. Across the literature, commercial activity is typically measured at the aggregate level, e.g. a station area’s total number of jobs or building area, rather than by the number or size of the establishments within station areas. POI data now make establishment-level measurement feasible, yet few studies have used it to ask whether the number and size of establishments are associated with ridership. This thesis seeks to address this gap by analyzing the relationship between the number of small establishments and rail ridership within station catchment areas.

Factors Determining Transit Ridership

A series of socioeconomic characteristics, policy, and built environment factors can affect transit use. Population size, economic vitality, and share of people with low levels of car access are associated with transit ridership (Taylor et al. 2003). Transit service frequency and fare price are also important determinants of transit ridership (Taylor et al. 2009). Priced parking and availability can affect automobile usage, increasing the likelihood of taking transit (Shoup 2005). Built environment factors have inelastic effects on traveling, but when combined could produce changes in travel behavior (Ewing and Cervero 2010).

Land Use and Transportation

There is an extensive body of scholarship on the relationship between the built environment and transportation (Ewing and Cervero 2010). Scholars have highlighted the effects of five aspects of the built environment (5Ds): density, diversity, design, destination accessibility and distance to transit (Ewing and Cervero 2001). Two other (D) factors, demographics and demand management, have significant effects on travel behavior as well (Ewing and Cervero 2010). Studies that include density measures find that higher population, employment, and commercial densities are positively associated with transit use (Cervero and Kockelman 1997, Sung and Oh 2011); however, these studies typically just consider overall building density by land use type, and do not consider the number of station-adjacent establishments as a determining factor in travel behavior.

While dense development without the other D factors has a small effect on travel behavior, it is associated with less driving (Ewing and Cervero 2017). Employment density is consistently associated with reduced likelihood of auto use and higher transit ridership (Chatman 2003, Pan et al. 2017). People who live closer to commercial buildings have an increased likelihood of using modes other than the automobile to make shopping trips (Jiao et al. 2011).

Commercial development near transit stations/stops can influence mode choice, but with nuanced results. Studies show a positive relationship between living in a transit-oriented development and transit ridership; however, this relationship is partly due to transit-riding residents who self-select into transit-oriented neighborhoods (Cervero 2007). This study contributes to the literature by examining whether dense small commercial establishments can play a role as a destination in attracting transit riders.

The Role of Station-Area Characteristics

Several station-level studies have examined the relationship between built environment factors in catchment areas and rail ridership. Kuby et al. (2004) concluded that total employment (jobs), distance to the central business district (CBD), residential population near stations, renter status, bus connections, park-and-ride parking spaces, and presence of transfer and terminal stations influence ridership. Commercial and office floor area are significantly associated with station boardings (Sohn and Shim 2010). Residential building area and number of educational buildings, entertainment venues, and shopping centers are significantly associated with rail transit ridership (Zhao et al. 2013). Average building area is also positively associated with transit ridership (Sung and Oh 2011). Loo et al. (2010) found that place-specific factors play significant roles in determining station level ridership. While zero-vehicle households are generally associated with transit use (Taylor et al. 2009), Loo et al. (2010) found a positive correlation between car ownership and rail ridership in New York City and Hong Kong, attributed to pick-up, drop-off, and park-and-ride rider behavior. Studies in different cities may generate different patterns based on local ridership behavior and culture.

Different catchment sizes show that employment density is a stronger predictor of ridership in a quarter-mile catchment area, whereas population density is more influential at the half-mile level (Guerra et al. 2012). Other characteristics like bus connectivity and mixed land use can also affect ridership across varied size spatial catchments (Jun et al. 2015).

While station area analyses have examined the effects of various land use factors, the literature on the relationship between establishment size and transit ridership is limited. Total commercial floor area is positively associated with transit use (Sohn and Shim 2010, Zhao et al. 2013). My study goes beyond aggregate commercial floor area by using POI data to further divide commercial floor area into specific uses to determine if the composition and size of commercial space are associated with transit ridership. My study isolates the relationship between small establishments and transit ridership by controlling for many of the above factors in my analysis.

Conceptual Model

I test the hypothesis that the number of small commercial establishments is associated with rail transit ridership, independent of other factors, so as total commercial square footage.

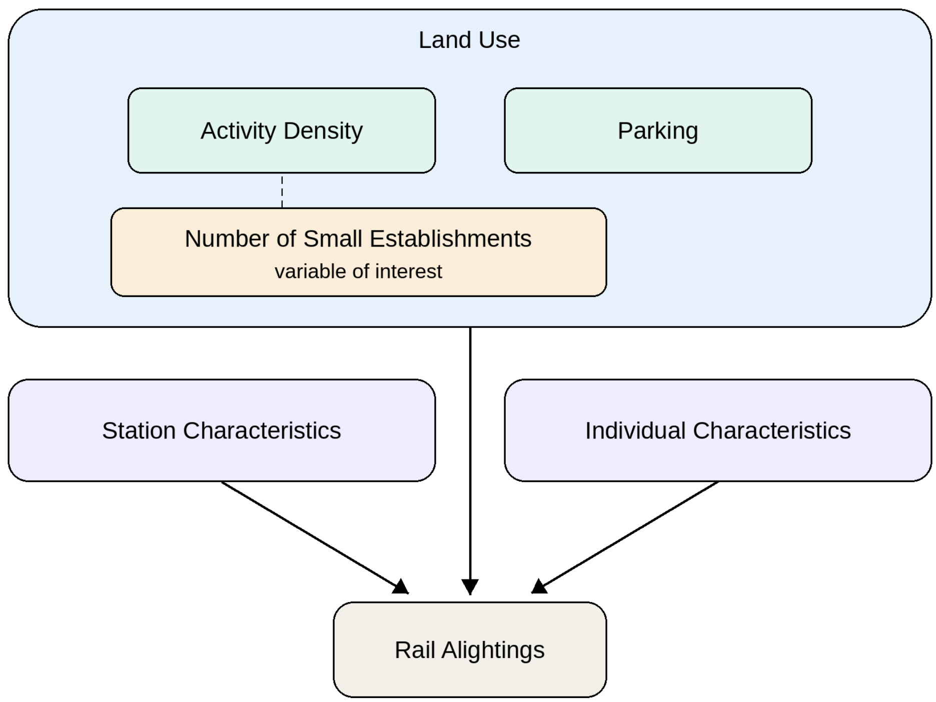

Figure 3. Conceptual model for statistical models

Drawing from the existing research, the schematic in Figure 3 depicts the factors associated with rail ridership, including my variable of interest, the number of small commercial establishments within either a half-mile or quarter-mile radius of a rail station. The dependent variable is average weekday alightings per rail station. The models also control for: parking, individual characteristics, land use, activity density, and station characteristics. I describe the variables in detail below.

Rail Alightings: I use average daily station alightings in 2025 as my dependent variable because I am interested in the role of small establishments in attracting trips to a destination, rather than capturing both boardings and alightings. While boardings are more directly measured, it is reasonable to assume most commercial trips are round trips, leading to boardings and alightings at a given station being roughly equivalent. I prefer to use alightings as a measurement because they can directly represent destination-side attraction.

Number of small commercial establishments: This is my variable of interest. I hypothesize that the number of small commercial establishments will be positively associated with ridership. As the number of small commercial establishments in the area increases, rail station arrivals may grow as well because riders are attracted to the more diverse opportunities to work and/or shop.

Land Use: As the research shows, land uses surrounding rail stations can affect transit ridership. Land use around each of the rail stations will vary with respect to the number and size of housing, businesses, and street intersections, each with differential effects on rail ridership. For example, if a station area includes high housing densities, naturally there will be more people who can ride transit compared to station areas with lower residential densities.

Parking: The amount of parking in rail station areas can affect ridership. Plentiful un- or under-priced parking in the area means a greater number of residents with automobiles and, typically, lower transit usage. The amount of surface parking near stores strongly influences whether people drive to them (Jiao et al. 2011). Parking requirements in Los Angeles for commercial buildings can vary, ranging from one parking space per 100 square feet to one parking space per 250 square feet (Chester et al. 2015).

Individual Characteristics: Individual characteristics such as socioeconomic status, race/ethnicity, and vehicle ownership play important roles in rail ridership. Lower-income people are more likely to use public transit than wealthier people, and people who do not own cars are even more likely to use public transit. Rail riders, who tend to be wealthier than bus riders, are more likely to own automobiles (Schouten et al. 2021).

Activity Density: Activity density is the concentration of both jobs and residents in an area. Areas with higher activity densities may have greater transit ridership than other areas as more people enter the area to live or work (Chatman 2008).

Station Characteristics: Station characteristics include aspects of each station’s physical location relative to the Los Angeles Central Business District (7th St. and Metro) and distance from one another. They also include whether the station is a terminal or transfer station and whether the service has operated on the station for a full year.

Data and Measurement

Point of Interest (POI) data

I use Safegraph's Global Places and Geometry dataset to approximate commercial establishment square footage (Safegraph 2026). The dataset consists of POIs and their respective polygons for non-residential buildings. I define a POI as a customer-facing commercial establishment. I remove non-commercial land uses like parks, schools, and industrial buildings from the dataset because they fall outside of the definition of commercial establishment. Safegraph tracks each unique POI's location, longitude, latitude, NAICS[4] code, whether a location is open or not, and polygon structure by scraping publicly available API[5] data, open store locators, and licensed third party data. Safegraph generates its polygons through a mixture of hand drawing and programming on satellite imagery. Safegraph updates their data monthly and has been available since 2019. I use a March 2026 snapshot dataset to identify each active establishment's location, count, and square footage. I use these data to count the number of establishments under 700 square feet, calculate the total unique polygon square feet, and develop a hospital dummy variable to account for stations near hospitals, where high concentrations of medical employment may increase ridership. I discuss these limitations in the data limitations section below.

Ridership Data

From LA Metro, I use 2025 monthly light and heavy rail station level ridership data. Aggregated by month, the dataset contains LA Metro's A, C, E, K, B and D line's individual station's average daily "on"s and "off"s divided by the weekday, Saturday, and Sunday. In total

[4] The North American Industry Classification System (NAICS) classifies various business establishments into different industries. It assigns each business a numerical code that identifies its industry type.
[5] An Application Programming Interface (API) is a set of rules and protocols that allows different software applications to communicate with each other.

there are 108 rail stations–92 light rail stations and 16 heavy rail stations. I aggregate the data from each month by averaging and weighting the number of days in each month.

Several stops on Metro's A and D lines were not open for the full 2025 calendar year due to construction or mid-year station openings. The Metro A line extension opened in September 2025, adding four new stations: Pomona North, San Dimas, Glendora, and La Verne / Fairplex. I did not include these four stations in the regression due to the limited ridership data (4 months). I include Metro's C line AMC / LAX station which opened in June 2025, allowing for seven months of ridership data. Eight of Metro's D line stations were closed from January to March 2025 due to construction. However, I include all D line stations in the dataset, in this case allowing for nine months of ridership data. Since I primarily use average daily alightings in my model, the analysis should not be seriously affected by the partial-year stations. Still, I account for the partial-year stations by including a "full-year" dummy variable in the model.

Figure 4. Map of Los Angeles Metro Rail and Bus Rapid Transit System (LA Metro, 2026)

Jobs Data

The jobs numbers are from the 2023 employment data from Longitudinal Employer-Household Dynamics (LEHD) Origin-Destination Employment Statistics (LODES). These data are collected by the United States Census Bureau and are organized by census block. The data are separated into three categories, origin-destination, residence area characteristics, and workplace area characteristics. My study uses the workplace area data, which provides a total count of jobs located in each census block.

Demographic Data

I use the American Community Survey (ACS) 5-Year 2023 dataset to assemble demographic data as my control variables. The ACS data provide median income, race/ethnicity, and household vehicle ownership at the census block group level. Two rail stations, El Segundo and Mariposa, do not have a complete set of census block group data because they are located near Los Angeles International Airport (LAX). These stations' half-mile buffers cover a set of partial census block groups, while the quarter-mile buffer does not touch any residential census block groups. Accordingly, I exclude the El Segundo and Mariposa stations from the quarter-mile analysis but retain them in the half-mile analysis.

Street Network Data

I use the United States Environmental Protection Agency's (EPA's) Smart Location 2018 database to quantify street intersection density. I use a measure of street intersection density weighted to emphasize pedestrian and bicycle connectivity over car-oriented intersections. The EPA calculated this variable by counting intersection legs and weighting links based on average automobile travel speeds (EPA 2018). The EPA then assigns each census tract an intersection density value accordingly.

Parking Data

Chester et al. (2015) created a Los Angeles County parking growth model from 1990 - 2010 using roadway and building stock estimates in conjunction with minimum off-street parking requirements. They provide estimates of parking lot spaces at the 2010 census tract level. I use their parking estimates rather than other public parking data sources, because the other data sources contain only metered parking spots or government-owned parking lots. I cannot measure whether a specific small establishment has a parking lot, but I use the total number of station-area parking lots as an aggregate measure of parking availability.

Station Characteristics Data

I gathered station characteristics data (transfer stations, terminal stations, LA Metro bus transfers) from public information available from LA Metro. I also used geographic information systems to calculate the Euclidean distance from each station to the 7th/Metro Center station in downtown Los Angeles as a rough proxy of location (urban, suburban) in the metro area and location (center, fringe) in the rail transit network.

Data Limitations

Shared Polygons

The Safegraph data have two polygon classes: owned polygon and shared polygon. Safegraph defines an owned polygon as an independent polygon with one POI attached to it. Safegraph defines a shared polygon as a polygon with at least two POIs attached to it. The ratio of shared polygons to owned polygons is 3:1 in the half-mile radius and 4:1 in the quarter-mile radius. When two businesses share the same polygon, I do not know the proportion of the polygon occupied by each of two businesses. Instead, I evenly estimate the average polygon size by dividing the polygon's total square footage by its number of commercial tenants. By dividing the square footage, I average out discrepancies between buildings that could hold large and small businesses. For example, a 1,000 sq. ft. building with an 800 sq. ft. and a 200 sq. ft. tenant is calculated in my data as a 1,000 sq. ft. building with two 500 sq. ft. tenants.

Multilevel Buildings

Safegraph also does not properly account for multilevel buildings' polygons. Since Safegraph uses satellite imagery to generate its polygons, a ten-story building will have the same polygon as a one-story building with the same base. Taller buildings tend to have more tenants, and I am unable to identify a building's number of floors. When I estimate the shared polygon square footage, I may overestimate the number of small establishments as the shared polygon's total square footage does not reflect the true multilevel building square footage. I acknowledge this by running separate regressions with both the original polygon square footage and estimated polygon square footage (a polygon's total square footage divided by the number of commercial tenants). The original polygon square footage will underestimate the number of small establishments, and the estimated polygon square footage will overestimate the number of small establishments.

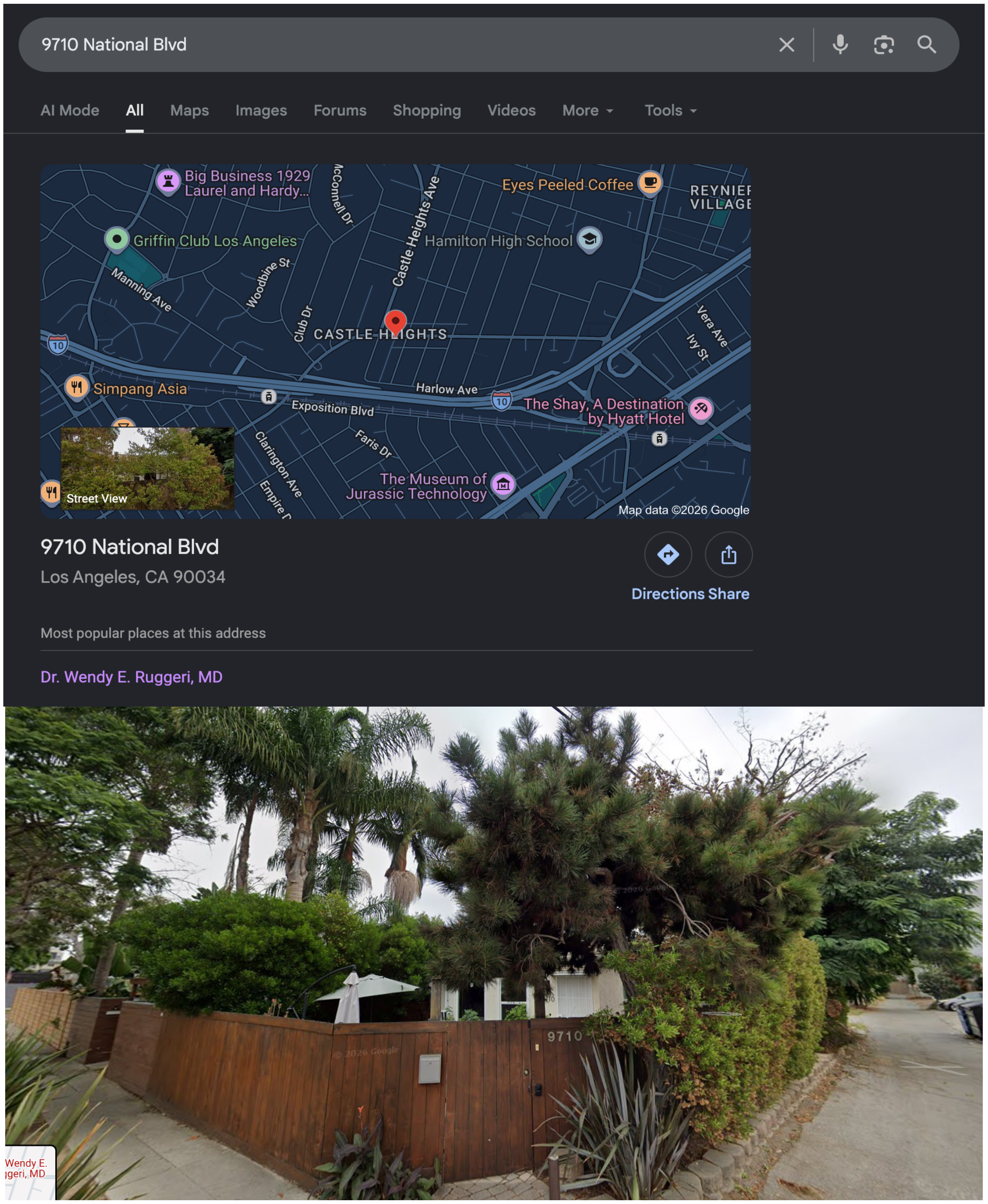

Figure 5. Example of non-existing business address (Google Search, Google Maps, 2026)

Phantom Buildings

While Safegraph scrapes online databases for address information, its original data source can have errors. Figure 5 shows an example of phantom businesses, in which a business's address may be located online but their actual physical location is different. For example, online queries indicate that a physician’s office is located at 9710 National Boulevard, when Google Maps shows a single-family home. It also is possible that information on some small establishments is not publicly available online, reducing the total number of small businesses in the database. While the extent of these errors is not easily quantified, the data should be understood as imperfect and results interpreted accordingly.

Census Track/Block Group Approximations

To analyze the half-mile and quarter-mile rail station buffers, I calculate the percentage of each census tract or block group that falls within the buffer radius, then multiply that percentage by the tract's total demographic counts. I sum the apportioned census counts across the intersecting tracts to estimate the total count for the tract/block group areas within the radii. This method assumes that population and demographic characteristics are uniformly distributed within each tract, which may introduce measurement errors in tracts with uneven spatial distributions of jobs and residents.

Sample Size and Characteristics

Table 1 below includes descriptive statistics for the areas and ridership within a half-mile radius of the stations in my sample. Table 2 similarly shows descriptive statistics for the areas and ridership within a quarter-mile radius of stations. The station characteristics for quarter-mile stations and half-mile stations are the same and, therefore, not repeated in Table 2. Notably, the estimated means and standard deviations of establishments under 700 sq. ft. vary substantially from the original establishment area means and standard deviations. This difference suggests that the estimated establishment count variables are measuring more establishments than the original establishment count, with higher levels of variation across stations.

Table 1. Half-Mile Radius Descriptive Statistics

Variable	Mean	Std Dev	N	Source
Dependent Variables				
Avg. Weekday Alightings	2,006	2,829	104	LA Metro (2025)
Avg. Saturday Alightings	1,602	2,262	104	LA Metro (2025)
Avg. Sunday Alightings	1,369	1,932	104	LA Metro (2025)
Independent Variables				
½-Mile Orig — Establishment Count (< 700 sq. ft.)	11.7	21.2	104	Safegraph (2026)
½-Mile Estimated — Establishment Count (< 700 sq. ft.)	419.8	633.6	104	Safegraph (2026)
Control Variables				
Logged ½-Mile Total Unique Polygon sq. ft.	15.4	1.0	104	Safegraph (2026)
½-Mile Hospital Within Buffer (dummy)	0.59	0.50	104	Safegraph (2026)
½-Mile Total Population	10,879	7,105	104	ACS 5 Year (2023)

½-Mile Median Household Income ($)	81,324	33,604	104	ACS 5 Year (2023)
½-Mile % Black (%)	12.6	13.6	104	ACS 5 Year (2023)
½-Mile % Hispanic (%)	48.6	24.9	104	ACS 5 Year (2023)
½-Mile Total Jobs	16,214	31,242	104	ACS 5 Year (2023)
½ Mile Zero Vehicle Household (%)	15.6	10.5	104	ACS 5 Year (2023)
½-Mile Intersection Density	166.4	78.7	104	EPA (2018)
½-Mile Parking Count	25,715	24,764	104	Chester et al. (2015)
Bus Transfer Count	4.0	1.5	104	LA Metro (2025)
Distance to 7th St/Metro Center (mi)	8.0	5.5	104	GIS Calculation
Terminal Station (dummy)	0.10	0.30	104	LA Metro (2025)
Transfer Station (dummy)	0.07	0.26	104	LA Metro (2025)
Distance to Nearest Station (mi)	0.82	0.48	104	GIS Calculation
Full-Year Data (dummy)	0.93	0.24	108	LA Metro (2025)

Table 2. Quarter-Mile Radius Descriptive Statistics

Variable	Mean	Std Dev	N	Source
Dependent Variables				
Avg. Weekday Alightings	2,039	2,847	102	LA Metro (2025)
Avg. Saturday Alightings	1,630	2,275	102	LA Metro (2025)
Avg. Sunday Alightings	1,393	1,943	102	LA Metro (2025)
Independent Variables				
¼-Mile Orig — Establishment Count (< 700 sq. ft.)	3.3	6.8	102	Safegraph (2026)
¼-Mile Estimated — Establishment Count (< 700 sq. ft.)	144.6	310.1	102	Safegraph (2026)
Control Variables				

Logged ¼-Mile Total Unique Polygon SQ FT	14.1	1.2	102	Safegraph (2026)
¼-Mile Hospital Within Buffer (dummy)	0.24	0.43	102	Safegraph (2026)
¼-Mile Total Population	2,581	1,788	102	ACS 5 Year (2023)
¼-Mile Median Household Income ($)	78,841	31,930	102	ACS 5 Year (2023)
¼-Mile % Black (%)	13.0	14.2	102	ACS 5 Year (2023)
¼-Mile % Hispanic (%)	48.1	26.5	102	ACS 5 Year (2023)
¼-Mile Total Jobs	5,477	11,143	102	ACS 5 Year (2023)
¼-Mile Zero Vehicle Household (%)	16.1	12.0	102	ACS 5 Year (2023)
¼-Mile Intersection Density	171.3	92.8	102	EPA (2018)
¼-Mile Parking Count	7,011	7,439	102	Chester et al. (2015)

Methodology

Previous research relies primarily on ordinary least squares (OLS) (Kuby et al., 2004, Sohn and Shim, 2010, Loo et al., 2010). OLS assumes a linear relationship between independent and dependent variables. However, my dependent variable, average daily rail transit alightings, is an overdispersed non-negative count variable. OLS, therefore, is not suitable for count variables because it can count negative values for alightings, which is impossible. A Poisson (Choi et al. 2012) regression model also would not be appropriate because it assumes that the sample mean is equal to the variance. However, my dependent variable, rail transit alightings, is over-dispersed with a variance-to-mean ratio of 3,966:1 across the 102-station sample.

Scholars have used Poisson Pseudo Maximum Likelihood (PPML) and negative binomial regressions in ridership studies because of their ability to handle over-dispersed count variables (Manville et al. 2024, Thompson et al. 2012). PPML estimates variance as a linear function of the mean while negative binomial estimates variance as a quadratic function of the mean. Figure 6 displays the relationship between the mean and variance of each regression's binned residuals. I conduct a mean-variance bin test to identify the relationship between the variance and the mean. I find that a negative binomial model has a R^2 of 0.98 and PPML has a R^2 of 0.89. While both R^2 values are high, I conclude that a negative binomial regression is the appropriate model to use for this analysis.

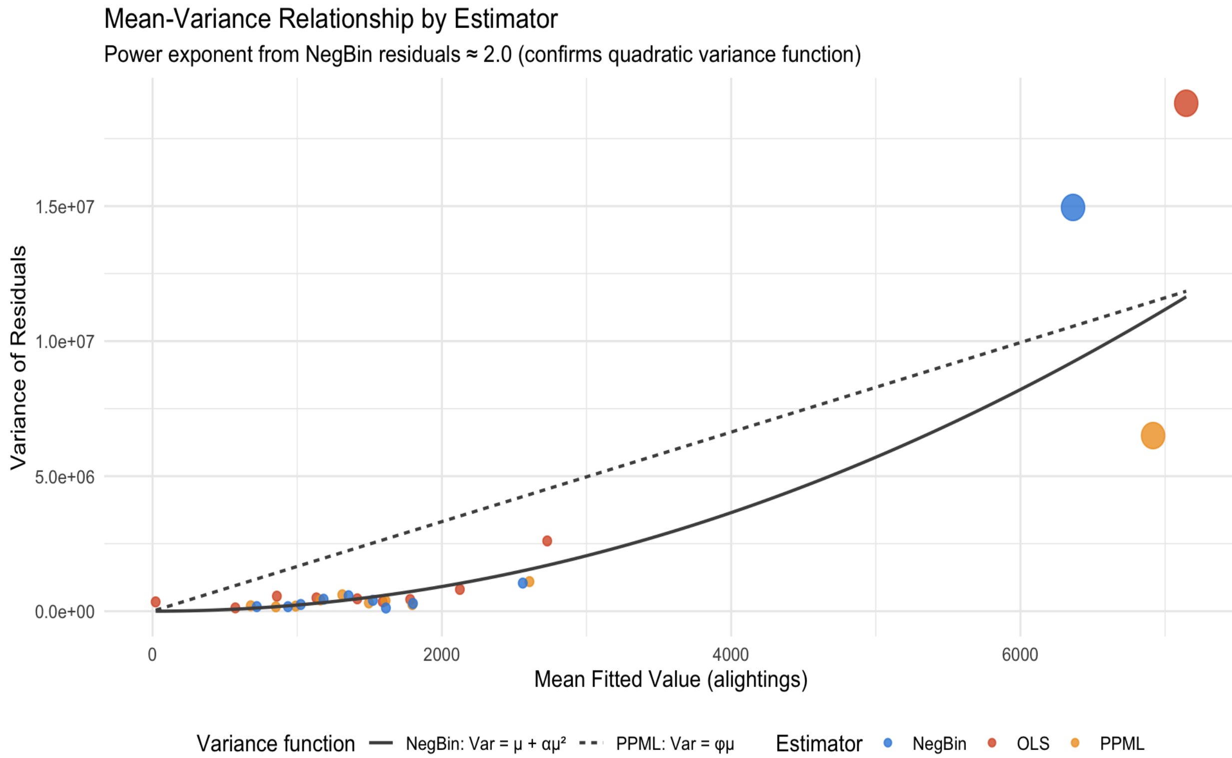

Figure 6. Mean-variance relationship by estimator type

I conduct six station-level negative binomial regression models. In Models 1 and 2, I test whether there is a relationship between small business establishments and rail ridership using the original and estimated number of small establishments. I define 'small' as commercial establishments under 700 square feet. I define a commercial establishment as a non-residential customer-facing business or organization. I then replicate these models (Model 3 and Model 4) to examine whether the findings differ between weekdays and weekends. Finally, in Models 5 and 6, I examine whether there is a relationship between rail alightings and the number of medium and medium-large commercial establishments.

1. Model 1 and 2: The relationship between establishments under 700 sq. ft. in a quarter and half-mile radius and rail alightings using original and estimated data.

2. Model 3 and 4: The relationship between establishments under 700 sq. ft. in a quarter and half-mile radius and rail alightings on the weekday versus weekend (Saturday and Sunday).
3. Model 5 and 6: The relationship between establishments under 701-2,000 and 2,001-5,000 sq. ft. in a quarter and half-mile radius and rail alightings on the weekday versus weekend (Saturday and Sunday).

The negative binomial regression takes the following form:

$$\text{Alightings} = \exp(\beta_0 + \beta_1\text{Establishments} + \beta_2\text{Total SQFT} + \beta_n\text{NControls})$$

Where:

Alightings are the average daily weekday, Saturday, or Sunday alighting per rail station;

Establishments are the number of commercial establishments under 700, 701-2000, or 2,001-5,000 square feet;

Total SQFT is the logged total square footage of commercial establishments in the quarter or half-mile radius; and

Controls are the additional variables under the Control Variables header in Table 1.

Results

My first models test the predictive power of the original square footage and estimated square footage count of establishments under 700 sq. ft. Each model uses the full set of control variables listed in Table 1. Table 3 includes the results of my variables of interest with the p-value and a p-value corrected with heteroskedasticity-robust standard errors. The Appendix, Table A shows the complete regression results.

Table 3. Original versus estimated square footage as predictors of weekday rail alightings

		Negative Binomial				
		Coefficient	Standard Error	P-value	Robust Standard Error	Robust P-value
Model 1: Original Square Footage Count (Weekday)	¼ Mile 700	0.024	0.012	0.03 **	0.013	0.06 *.
	½ Mile 700	0.006	0.004	0.12	0.007	0.40
Model 2: Estimated Square Footage Count (Weekday)	¼ Mile 700	-0.0001	0.0001	0.81	0.0003	0.81
	½ Mile 700	0.0001	0.0001	0.36	0.0002	0.33

Note: * = P <.10. ** = P <.05. *** - P < .01.

The original number of establishments under 700 sq. ft. within a quarter-mile is the only value that is statistically significant at the 95% level. When factoring in heteroskedasticity, the level of significance drops out of the 95% level and the relationship between the number of small establishments in the quarter-mile radius and rail alightings is no longer significant. Comparing the original square footage counts and estimated square footage counts, both the estimated square footage counts are statistically insignificant with smaller coefficients.

Table 4. Weekday rail alightings model fit comparison of original and estimated square footage using AIC and BIC

Statistics	Original Square Feet Count	Estimated Square Feet Count
AIC	-1.44	1.90
BIC	1.18	4.53

Table 4 shows the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) statistics for the original and estimated square footage count of establishments under 700 square feet. The original square footage count has lower AIC and BIC values than the estimated square footage count, indicating a better model fit. Given these results, I use the original square footage count in favor of the estimated square footage count in all my subsequent models.

Table 5. Comparison of small establishments' relationships to rail alightings across weekday, Saturday, and Sunday

		Negative Binomial				
		Coefficient	Standard Error	P-value	Robust Standard Error	Robust P-value
Model 1: Weekday Average Daily Alightings	¼ Mile 700 sq. ft.	0.024	0.012	0.03 **	0.013	0.06 *.
	½ Mile 700 sq. ft.	0.006	0.004	0.12	0.007	0.40
Model 3: Saturday	¼ Mile 700 sq. ft.	0.030	0.011	0.01 ***	0.015	0.05 **.

Average Daily Alightings	½ Mile 700 sq. ft.	0.007	0.004	0.05*	0.008	0.34
Model 4: Sunday Average Daily Alightings	¼ Mile 700 sq. ft.	0.031	0.011	0.01 ***	0.015	0.05 **.
	½ Mile 700 sq. ft.	0.008	0.004	0.05*	0.008	0.33

Note: * = P <.10. ** = P <.05. *** - P < .01.

Table 5 shows the associations between small establishments in a half- and quarter-mile on rail alightings based on the day of the week. At a half-mile radius, the robust p-value shows no change in statistical significance, ranging from 0.40 on the weekday to 0.34 on Saturday and 0.33 on Sunday. However, at a quarter-mile radius the weekday p-value decreases from 0.06 to 0.05 on Saturday and then to 0.05 on Sunday, becoming modestly statistically significant. The coefficient increases from 0.024 on the weekday to 0.030 on Saturday and 0.029 on Sunday. Within a quarter-mile radius, the IRR[6] for Saturday is 1.030, indicating a 2.98 percent increase in alightings per additional small establishment, while the IRR for Sunday is 1.031, indicating a 3.12 percent increase. The number of establishments under 700 square feet has a more pronounced effect on weekend ridership than on weekday ridership. In both weekday and the weekend alightings, the half-mile radius is not statistically significant, while the quarter-mile radius results are modestly significant with p-values just below 0.05. The consistent significance suggests that the number of small establishments close to a rail transit station has a positive relationship with transit alightings, while those further from a rail station do not.

Table 6. Comparison of medium establishments as predictors of rail alightings across weekday, Saturday, and Sunday

[6] Incidence Rate Ratio (IRR) is the exponentiated regression coefficient (eβ) measuring the multiplicative change in expected count.

			Negative Binomial				
	Buffer	Day	Coefficient	Standard Error	P-value	Robust Standard Error	Robust P-value
Model 5: Medium - Original Square Footage Count (701–2,000 sq. ft.)	¼ Mile	Weekday	-0.002	0.003	0.514	0.003	0.489
		Saturday	-0.001	0.003	0.929	0.003	0.924
		Sunday	-0.001	0.003	0.914	0.003	0.911
	½ Mile	Weekday	0.001	0.001	0.226	0.001	0.198
		Saturday	0.002	0.001	0.043 **	0.001	0.076 *.
		Sunday	0.002	0.001	0.062 *.	0.001	0.093 *.

Note: * = P <.10. ** = P <.05. *** - P < .01.

Table 7. Comparison of medium-large establishments as predictors of rail alightings across weekday, Saturday, and Sunday

			Negative Binomial				
	Buffer	Day	Coefficient	Standard Error	P-value	Robust Standard Error	Robust P-value
Model 6: Medium-Large - Original Square Footage Count (2,001–5,000 sq. ft.)	¼ Mile	Weekday	0.001	0.001	0.797	0.001	0.811
		Saturday	0.001	0.001	0.562	0.001	0.563
		Sunday	0.001	0.001	0.578	0.001	0.59
	½ Mile	Weekday	0.001	0.001	0.506	0.001	0.502
		Saturday	0.001	0.001	0.159	0.001	0.168
		Sunday	0.001	0.001	0.19	0.001	0.201

Note: * = P <.10. ** = P <.05. *** - P < .01.

Tables 6 and 7 show the differences in the size of the counts of establishments in a half- and quarter-mile radius. Higher numbers of medium to medium/large establishments in station areas are not significantly associated with rail alightings. Notably on the half-mile level, medium establishments on the weekend robust p-values are between 0.05 and 0.10, indicating a marginally significant relationship between medium establishments and rail alightings. However, the coefficient is relatively small, with both Saturday's and Sunday's values being 0.002. Their IRR of 1.002 indicates a 0.2 percent increase in rail alightings per every medium establishment. The results show a marginal association between rail alightings and medium-size establishments within a half-mile, but the relationship is weaker and less consistent than small establishments within a quarter-mile.

Table 8. Moran's I Test for spatial autocorrelation for weekday, Saturday, and Sunday

	Moran's I	P-value
Weekday	0.05	0.25
Saturday	0.06	0.15
Sunday	0.06	0.18

I conduct a Moran's I Test to determine if the establishments under 700 square feet within a quarter-mile of a station association with rail alightings results suffer from spatial autocorrelation.[7] Table 8 shows that the weekday, Saturday, and Sunday residual results indicate no significant spatial autocorrelation in the negative binomial regression.

I also compute variation inflation factors (VIF) to test for multicollinearity. On the quarter-mile radius regressions, no variables had VIF scores above five. However, when switching to the half-mile radius regressions, two variables had VIF scores above or approaching five, which indicates potentially problematic multicollinearity: parking count's VIF increased from 4.77 in the quarter-mile regression to 7.02, and percent zero vehicle households' VIF increased from 3.60 to 4.93. VIF values above five suggest multicollinearity among the variables, inflating standard errors and reducing the power of hypothesis tests. I then tested models excluding percent zero vehicle households and parking count in turn. Changes in the coefficients and significance of my variable of interest were negligible. I thus decided to keep these control variables in the half-mile regression because the availability of parking and car access are important conceptual determinants of mode choice.

On the weekends, total average rail alightings decrease from an average of 2,039 to 1,630 alightings, as fewer people work on weekends, and a disproportionate share of rail trips are work trips (Wang and Renne 2023). Yet the relationship between alightings and small

[7] Spatial Autocorrelation is the degree of correlation among nearby observations. The association of a variable with itself through space is measured by spatial autocorrelation and it can be negative or positive (Verma 2021).

establishments strengthens on the weekend. As trip purpose proportionally shifts from work to leisure, there are more rail riders at stations with a greater number of small establishments. Fewer work trips and more leisure trips have more riders going to stations with small establishments to eat, drink, and shop. Logically this makes sense, as weekend rail riders may be drawn to areas with significant commercial diversity.

Similar ridership studies have used threshold effects to examine the relationship between land use variables and transit ridership (Shao et al. 2020, Zhu et al. 2024). Threshold models can help identify non-linear patterns between the built environment and transit ridership. Built environment variables like population and employment density have threshold effects on rail ridership (Ding et al. 2019). I test for threshold effects in the relationship between the number of small establishments and alightings to assess whether a certain number of establishments is needed before any changes in the relationship with rail ridership occur.

Zhu et al. (2024) uses Hansen's (2000) panel threshold model to determine whether transit use increases when development intensity and land use mix exceed certain thresholds. Hansen's (2000) panel threshold model estimates the threshold by testing all candidate values to find the threshold value that has the lowest sum of squared residuals, or in the case of the negative binomial regression, the maximum log-likelihood value. After estimating the threshold, I employ a bootstrap method[8] for 500 iterations to estimate the threshold confidence intervals.

[8] A bootstrap method is a resampling technique that forms small random samples from a larger dataset, allowing for inference of population properties based on the resampled sets.

Table 9. Panel threshold model across weekday, Saturday, and Sunday

Day type	Optimal Threshold	N below Threshold	N above Threshold	Bootstrap p-value
Weekday	6	87	15	0.524
Saturday	6	87	15	0.660
Sunday	6	87	15	0.648

The threshold model predicted six small establishments as the optimal threshold, with 87 stations below and 15 stations above this threshold. Bootstrap p-values across the weekdays, Saturday, and Sunday were not statistically significant, meaning that I cannot reject the null hypothesis that the relationship between small establishments and rail alightings is the same across the full range of establishments.

The stronger effect that small establishments have in the quarter-mile radius aligns with existing literature, with Guerra et al. (2012) finding that the number of jobs in a quarter-mile radius are a stronger predictor of ridership than the half-mile radius. Jun et al. (2015) also found that the proportion of commercial land and mixed land use had a positive relationship on rail ridership in smaller catchment areas than larger ones. The quarter-mile radius tends to have a stronger relationship with commercial land uses and development than the half-mile radius. Gutiérrez et al. (2011) found that population and employment located closer to rail stations generate more boardings than those farther away.

The results reflect that small establishments exhibit a similar proximity relationship. Small stores, which rely more heavily on nearby foot traffic (Sevtsuk 2020), could benefit from higher concentrations of pedestrian foot traffic closer to station areas, as foot traffic density near transit stations is the highest at the quarter-mile radius level (Huang et al. 2017).

Overall, however, I find that the number of small establishments in a half-mile radius of rail stations is generally not statistically significantly associated with rail alightings. The number of small establishments in a quarter-mile radius is modestly significant, indicating an exploratory role for proximity given the suggestive relationship with rail ridership.

Conclusion

This study estimates station-level negative binomial regression models to analyze the relationship between the number of commercial establishments under 700 square feet and Los Angeles Metro rail alightings. The study also examines the changes between the relationship comparing weekday and weekend ridership, larger establishments between 701-2000 and 2001-5000 square feet, and establishment threshold effects. The results suggest that small commercial establishments in a half-mile radius are not associated with rail alightings. The number of small commercial establishments in a quarter-mile radius has a modestly statistically significant relationship with rail alightings, with a stronger relationship appearing on weekends than weekdays. Medium-sized establishments (701-2000 sq. ft.) in a half-mile radius on weekends are marginally significant, but with a negligible coefficient, and large establishments (2001-5000 sq. ft.) are not significant on any level. No threshold effect between small establishments and rail alightings was observed to be statistically significant.

Results from this study demonstrate that there is a modestly significant relationship between size of establishment and rail alightings. Smaller establishments closer to rail stations have a stronger relationship to alightings than medium and medium-large establishments, and closer small establishments have a stronger relationship than small establishments further away from rail stations. Future research could look towards identifying any causal relationships between the number of small establishments and rail alightings, given data on ridership and business growth over a period of time.

The modest results may stem partly from data and methodological limitations that future research can address. Research of this sort would be improved by increasing data accuracy on point of interest locations. Gathering a specific square footage for each active commercial business regardless of building height will help reduce errors that could overestimate or underestimate commercial establishment size. Aligning the variable of interest, business establishment square footage, with more accurate data sources could help. More recent studies have employed different regression models, such as geographic weighted regression (GWR) (Li et al. 2020) and machine learning algorithms (Yang et al. 2023) to study local interactions and non-linear relationships between transit ridership and the built environment. Other regression models can find non-linear or spatial relationships between small establishments and rail ridership that this study does not identify. Loo et al. (2010) found that place-specific factors can contribute to differences in transit ridership suggesting the importance of replicating this study in other metropolitan areas.

These equivocal findings regarding commercial establishment density and rail transit use do not necessarily argue against the virtues of smaller commercial establishments. Outside of transit ridership, small commercial establishments can prove useful by developing on underused small lots. Development on small lots encourages smaller buildings and incremental growth (Brozen et al. 2018), which can increase commercial and employment density by developing on previously under-utilized land parcels. In pedestrian environments that are safe, comfortable, and equitable, small-format stores can produce a stimulating and diverse environment (Sevtsuk, 2020). While the relationship between small commercial establishments and rail transit ridership is not definitive, small stores may still play an important role in the future of commercial development. Planners should continue to look and identify various trends and relationships in regard to changes in commercial establishment size to find strategies that may increase transit ridership.

References

Brozen, M., Hartzell, M., Manville, M., Monkkonen, P., & Vallianatos, M. (2018). Transit oriented Los Angeles: Envisioning an equitable and thriving future. UCLA Lewis Center for Regional Policy Studies. https://escholarship.org/uc/item/4r514641

Cal. S.B. 79, 2025–2026 Reg. Sess. (2025). Retrieved from https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=202520260SB79

Cervero, R. (2007). Transit-Oriented Development's Ridership Bonus: A Product of Self-Selection and Public Policies. *Environment and Planning A: Economy and Space*, *39*(9), 2068–2085. https://doi.org/10.1068/a38377

Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, *2*(3), 199–219. https://doi.org/10.1016/S1361-9209(97)00009-6

Chatman, D. G. (2003). How Density and Mixed Uses at the Workplace Affect Personal Commercial Travel and Commute Mode Choice. *Transportation Research Record: Journal of the Transportation Research Board*, *1831*(1), 193–201. https://doi.org/10.3141/1831-22

Chatman, D. G. (2008). Deconstructing development density: Quality, quantity and price effects on household non-work travel. *Transportation Research Part A: Policy and Practice*, *42*(7), 1008–1030. https://doi.org/10.1016/j.tra.2008.02.003

Chester, M., Fraser, A., Matute, J., Flower, C., Pendyala, R., & Metis Center For Infrastructure And Sustainable Engineering. (2024). Los Angeles County

Parking Space Infrastructure Inventory (Version 1.0, pp. 1014120, 2852). ASU Library Research Data Repository. https://doi.org/10.48349/ASU/F8EQ0Y

Chester, M., Fraser, A., Matute, J., Flower, C., & Pendyala, R. (2015). Parking Infrastructure: A Constraint on or Opportunity for Urban Redevelopment? A Study of Los Angeles County Parking Supply and Growth. *Journal of the American Planning Association*, *81*(4), 268–286. https://doi.org/10.1080/01944363.2015.1092879

Choi, J., Lee, Y. J., Kim, T., & Sohn, K. (2012). An analysis of Metro ridership at the station-to-station level in Seoul. *Transportation*, *39*(3), 705–722. https://doi.org/10.1007/s11116-011-9368-3

Ding, C., Cao, X., & Liu, C. (2019). How does the station-area built environment influence Metrorail ridership? Using gradient boosting decision trees to identify non-linear thresholds. *Journal of Transport Geography*, *77*, 70–78. https://doi.org/10.1016/j.jtrangeo.2019.04.011

Ewing, R., & Cervero, R. (2010). Travel and the Built Environment: A Meta-Analysis. *Journal of the American Planning Association*, *76*(3), 265–294. https://doi.org/10.1080/01944361003766766

Ewing, R., & Cervero, R. (2017). “Does Compact Development Make People Drive Less?” The Answer Is Yes. *Journal of the American Planning Association*, *83*(1), 19–25. https://doi.org/10.1080/01944363.2016.1245112

Ganning, J., & Miller, M. M. (2020). Transit oriented development and retail: Is variation in success explained by a gap between theory and practice?

Transportation Research Part D: Transport and Environment, *85*, 102357. https://doi.org/10.1016/j.trd.2020.102357

Google. (2026). [Google Maps image of Los Angeles Metro Stations]. Google Maps. https://www.google.com/maps

Guerra, E., Cervero, R., & Tischler, D. (2012). Half-Mile Circle: Does It Best Represent Transit Station Catchments? *Transportation Research Record: Journal of the Transportation Research Board*, *2276*(1), 101–109. https://doi.org/10.3141/2276-12

Gutiérrez, J., Cardozo, O. D., & García-Palomares, J. C. (2011). Transit ridership forecasting at station level: An approach based on distance-decay weighted regression. *Journal of Transport Geography*, *19*(6), 1081–1092. https://doi.org/10.1016/j.jtrangeo.2011.05.004

Handy, S. L., & Clifton, K. J. (2001). Local shopping as a strategy for reducing automobile travel. *Transportation*, *28*(4), 317–346. https://doi.org/10.1023/A:1011850618753

Huang, R., Moudon, A. V., Zhou, C., Stewart, O. T., & Saelens, B. E. (2017). Light rail leads to more walking around station areas. *Journal of Transport & Health*, *6*, 201–208. https://doi.org/10.1016/j.jth.2017.02.002

Hutson, N. M., & Orlando, A. W. (2023). The effects of COVID-19 on downtown land use: Evidence from four U.S. cities. *Journal of Urban Affairs*, 1–19. https://doi.org/10.1080/07352166.2023.2195665

Jiao, J., Moudon, A. V., & Drewnowski, A. (2011). Grocery Shopping: How Individuals and Built Environments Influence Choice of Travel Mode.

Transportation Research Record: Journal of the Transportation Research Board, *2230*(1), 85–95. https://doi.org/10.3141/2230-10

Jun, M.-J., Choi, K., Jeong, J.-E., Kwon, K.-H., & Kim, H.-J. (2015). Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul. *Journal of Transport Geography*, *48*, 30–40. https://doi.org/10.1016/j.jtrangeo.2015.08.002

King, K. (2023, December 12). The American store is shrinking. *The Wall Street Journal*. https://www.wsj.com/real-estate/commercial/the-american-store-is-shrinking-7603e979

Kuby, M., Barranda, A., & Upchurch, C. (2004). Factors influencing light-rail station boardings in the United States. *Transportation Research Part A: Policy and Practice*, *38*(3), 223–247. https://doi.org/10.1016/j.tra.2003.10.006

Li, S., Lyu, D., Huang, G., Zhang, X., Gao, F., Chen, Y., & Liu, X. (2020). Spatially varying impacts of built environment factors on rail transit ridership at station level: A case study in Guangzhou, China. *Journal of Transport Geography*, *82*, 102631. https://doi.org/10.1016/j.jtrangeo.2019.102631

Loo, B. P. Y., Chen, C., & Chan, E. T. H. (2010). Rail-based transit-oriented development: Lessons from New York City and Hong Kong. *Landscape and Urban Planning*, *97*(3), 202–212. https://doi.org/10.1016/j.landurbplan.2010.06.002

Los Angeles County Metropolitan Transportation Authority. (n.d.). *Schedules*. LA Metro. https://www.metro.net/riding/schedules-2/

Los Angeles County Metropolitan Transportation Authority. (2025). Rail station ridership data [Data set]. LA Metro. https://www.metro.net/

Manville, M., King, H., Matute, J., & Lau, T. (2024). Neighborhood change and transit ridership: Evidence from Los Angeles and Orange Counties. *Journal of Transport Geography*, *121*, 104048. https://doi.org/10.1016/j.jtrangeo.2024.104048

Pan, H., Li, J., Shen, Q., & Shi, C. (2017). What determines rail transit passenger volume? Implications for transit oriented development planning. *Transportation Research Part D: Transport and Environment*, *57*, 52–63. https://doi.org/10.1016/j.trd.2017.09.016

Peng, J., Fu, X., Wu, C., Dai, Q., & Yang, H. (2025). Comparative analysis of nonlinear impacts on the built environment within station areas with different metro ridership segments. *Travel Behaviour and Society*, *38*, 100898. https://doi.org/10.1016/j.tbs.2024.100898

Roman, E. (2024, September 20). Why big box retailers are experimenting with small-format stores. *Yahoo Finance*. https://finance.yahoo.com/news/why-big-box-retailers-are-experimenting-with-small-format-stores-132814528.html

SafeGraph. (2022). Global Places (POI) & Geometry [Dataset]. Dewey Data. https://doi.org/10.82551/SMXB-1K04

Schouten, A., Blumenberg, E., & Taylor, B. D. (2021). Rating the composition: Deconstructing the demand-side effects on transit use changes in California. *Travel Behaviour and Society*, *25*, 18–26. https://doi.org/10.1016/j.tbs.2021.05.007

Schuetz, J. (2015). Do rail transit stations encourage neighbourhood retail activity? *Urban Studies*, *52*(14), 2699–2723. https://doi.org/10.1177/0042098014549128

Sevtsuk, A. (2020). Street commerce: Creating vibrant urban sidewalks. University of Pennsylvania Press.

Shao, Q., Zhang, W., Cao, X., Yang, J., & Yin, J. (2020). Threshold and moderating effects of land use on metro ridership in Shenzhen: Implications for TOD planning. Journal of Transport Geography, 89, 102878. https://doi.org/10.1016/j.jtrangeo.2020.102878

Sharma, V., Sharma, M., Pandita, S., Kour, J., & Sharma, N. (2021). Application of geographic information system and remote sensing in heavy metal assessment. In *Heavy Metals in the Environment* (pp. 191–204). Elsevier. https://doi.org/10.1016/B978-0-12-821656-9.00011-0

Shoup, D. (2005). The high cost of free parking. American Planning Association, Planners Press.

Sohn, K., & Shim, H. (2010). Factors generating boardings at Metro stations in the Seoul metropolitan area. *Cities*, *27*(5), 358–368. https://doi.org/10.1016/j.cities.2010.05.001

Sung, H., & Oh, J.-T. (2011). Transit-oriented development in a high-density city: Identifying its association with transit ridership in Seoul, Korea. *Cities*, *28*(1), 70–82. https://doi.org/10.1016/j.cities.2010.09.004

Taylor, B. D., Miller, D., Iseki, H., & Fink, C. (2009). Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas.

Transportation Research Part A: Policy and Practice, *43*(1), 60–77. https://doi.org/10.1016/j.tra.2008.06.007

Thompson, G., Brown, J., & Bhattacharya, T. (2012). What Really Matters for Increasing Transit Ridership: Understanding the Determinants of Transit Ridership Demand in Broward County, Florida. *Urban Studies*, *49*(15), 3327–3345. https://doi.org/10.1177/0042098012443864

U.S. Census Bureau. (2023). American Community Survey 5-year estimates [Data set]. https://data.census.gov/

U.S. Census Bureau, Longitudinal-Employer Household Dynamics Program. (2024). LEHD Origin-Destination Employment Statistics (LODES). https://lehd.ces.census.gov/data/

U.S. Environmental Protection Agency. (2021). Smart Location Database technical documentation and user guide (Version 3.0). https://www.epa.gov/system/files/documents/2023-10/epa_sld_3.0_technicaldocumentationuserguide_may2021_0.pdf

Wang, X., & Renne, J. L. (2023). Socioeconomics of urban travel in the U.S.: Evidence from the 2017 NHTS. *Transportation Research Part D: Transport and Environment*, *116*, 103622. https://doi.org/10.1016/j.trd.2023.103622

Williams, K. (2024, November 30). For big-box retailers, one of the longest-running experiments is shrinking. *CNBC*. https://www.cnbc.com/2024/11/30/for-big-box-retail-one-of-longest-running-experiments-is-shrinking.html

Zeng, Q., Bao, X., & Dewancker, B. J. (2023). Association between built environment on transport and recreational walking in Japan: The case of

Kitakyushu. *City and Built Environment*, *1*(1), 10. https://doi.org/10.1007/s44213-023-00011-2

Zhao, J., Deng, W., Song, Y., & Zhu, Y. (2013). What influences Metro station ridership in China? Insights from Nanjing. *Cities*, *35*, 114–124. https://doi.org/10.1016/j.cities.2013.07.002

Zhu, H., Peng, J., Dai, Q., & Yang, H. (2024). Exploring the long-term threshold effects of density and diversity on metro ridership. *Transportation Research Part D: Transport and Environment*, *128*, 104101. https://doi.org/10.1016/j.trd.2024.104101

Ziedan, A., Brakewood, C., & Watkins, K. (2023). Will transit recover? A retrospective study of nationwide ridership in the United States during the COVID-19 pandemic. *Journal of Public Transportation*, *25*, 100046. https://doi.org/10.1016/j.jpubtr.2023.100046

Appendix

Table A. Model's Full Regression Results with all Variables

Variable	Coefficient	SE	P-Value
(Intercept)	7.499366	0.047095	0.0000
Number of Stores in a Half-Mile Radius under 700 sq. ft.	0.004358	0.004279	0.2408
Log of Total Polygon Square Feet in Half-Mile Radius	-0.047260	0.058053	0.4156
Half-Mile Hospital Dummy	-0.044630	0.121398	0.7131
Half-Mile Total Population	0.000020	0.000011	0.0591
Half-Mile Median Income	-0.000012	0.000003	0.0000
Half-Mile Percent Black	-0.000158	0.006596	0.9809
Half-Mile Percent Hispanic	0.007859	0.005437	0.1483
Half-Mile Total Jobs	0.000005	0.000003	0.1044
Number of Bus Transfers in 300ft	0.042781	0.038592	0.2676
Distance to 7th St/Metro	-0.014838	0.011924	0.2133
Half-Mile Intersection Density	-0.000026	0.000904	0.9771
Half-Mile Parking Count	0.000000	0.000006	0.9376

Terminal Station Dummy	0.433112	0.176692	0.0142
Transfer Station Dummy	1.547793	0.226480	0.0000
Mile Distance between Stations	0.419662	0.128336	0.0011
Full Year of Operation Dummy	-0.181223	0.359709	0.6144
Half-Mile Percent Zero Vehicle Households	-0.005444	0.010186	0.5930