Residential Segregation at Physical Neighborhood Boundaries

Kenneth Whaley
Department of Economics, University of South Florida
Tampa, Florida, USA
kenneth.whaley@usf.edu

December 15, 2023

Abstract

Physical boundaries delineate neighborhoods and are distinguishable from administrative boundaries like school districts and county lines. This paper sheds light on historic railroad placement as a predictor of contemporary segregation by employing a digitized map of Texas railroads circa 1911 to compare census block groups separated by train tracks today. Using a boundary discontinuity design, I first document an unconditional house price premium of 21% to live on the high income side of the tracks. Exploiting distinct variation in race and income demographics at railroad boundaries, I obtain hedonic estimates of the price premium for white population share and income composition. Conditional on differences in school quality and access to private consumption amenities, households are willing to pay up to 16% of home price for the race and income composition available on the high income side of the tracks.

JEL Classification: R23, J15, O18

1 Introduction

Contemporary research has given clear window into the legacy of policies that physically segregated households by race (Aaronson et al. 2021; Cook et al. 2022). Neighborhoods as we know them today are the offspring of redlining, mortgage discrimination, and Jim Crow era segregation laws. In the absence of segregation by law, theory tells us that strong preferences for segregation are not required for housing markets to end up in a segregated equilibrium (Schelling 1971; Card et al. 2008), and at least some white households are willing to pay more to live in white neighborhoods (Cutler et al. 1999).

To that end, I ask what is the willingness to pay for neighborhood race and income characteristics. Obtaining unbiased estimates for the house price capitalization of neighborhood race and income is rare because neighborhood boundaries are often poorly defined and potentially endogenous (Bischoff 2008; Keele and Titiunik 2015; Monarrez 2022). Moreover, systemic underinvestment in private and public amenities for Black and Latino neighborhoods poses an omitted variables problem as unobserved neighborhood quality will affect house prices, while being correlated with measures of racial composition (Bayer et al. 2007).

To overcome the boundary problem I take the placement of Texas railroad tracks laid in the early 1900s as distinct neighborhood boundaries today. The physical durability of railroads is remarkable, with much of the existing track mileage in place for over a century. The extent to which rail lines divide up cities into smaller neighborhoods has captured the attention of economists quantifying the effects of segregation on urban inequality, crime, and economic mobility (Ananat 2011, Cox et al. 2022, Chyn et al. 2022). Using census block group data for home values, housing quality, and race and income characteristics, I employ a hedonic model to compare neighborhoods close to, but on the opposite sides of the railroad track at a local level. Hedonic models are a common approach to estimating the price paid for specific neighborhood attributes (Rosen 1974; Witte et al. 1979; Yinger 2015).

The conceptual grounding of this paper relies on the fact that train operation creates environmental costs borne by nearby households on both sides of the boundary. If environmental costs are assumed to be indifferent to boundary side and decreasing in distance from the tracks, there are testable predictions for the house price gradient across the boundary space. Absent segregation the assignment of other neighborhood amenities is as good as random, and house prices would be lowest at the boundary and increasing in distance from the tracks. Further, the rate at which prices increase with distance from the tracks will be approximately the same on both sides of the boundary. Alternatively, the presence of systemic segregation will distort the house price gradient due to non-random assignment of households to neighborhoods according to race and income characteristics.

Mapping block group neighborhoods to railroad track segments, I employ a boundary discontinuity design that follows closely to Dell (2010). Embedding a boundary discontinuity design into a hedonic regression is empirical strategy popularized by Black (1999) and Bayer et al. (2007) to estimate the house price capitalization of school quality. The central theme of this literature is that unobserved neighborhood attributes are plausibly held constant across a small space near an administrative boundary, allowing for unbiased identification of the relationship between prices and a single neighborhood attribute like school quality. Physical boundaries, such as railroads, are less studied and predicted to influence a host of local neighborhood characteristics.

To estimate the model I use a digitized version of New Century Atlas Map that provides the location of all rail lines in Texas as of 1911 (Atack 2013). Further, geographic coordinates of each block group population center to are used to match neighborhoods to boundaries and compute distance to the tracks. The boundary side with higher median incomes is designated as the high income side of the tracks to organize the data for analysis. From the baseline model I obtain an unconditional house price premium of approximately 13-17% for the high income side of the tracks. Comparing cities to suburbs and rural areas, I find evidence of the unconditional price premium ranging from 8-19% across neighborhood types.

I next set out to understand what neighborhood attributes explain the unconditional price premium. To start I show that railroad boundaries produce discrete changes in neighborhood income composition, racial demographics, and housing quality characteristics. Neighborhoods on the high income side of the tracks are more white and have 24% higher incomes. The hedonic model predicts that increasing the share of white residents by 0.1 increases house prices by approximately 6.5-7%. Increasing neighborhood income by 10% is predicted to increase house prices by 3-4%.

The results are robust to the inclusion of a rich set of variables describing public and private amenities for each neighborhood. As a proxy for public amenities I use school level data describing test scores and per-pupil spending for the nearest elementary school, provided by Texas Education Agency. The second source of administrative data I employ are sales tax permit records provided by the Texas state comptrollers office. Address data in the tax records are used to construct counts of food and retail businesses near block group population centers as a proxy of private consumption amenities. Results from the full model show that the house price premium is explained entirely by housing quality, demographics, and amenity quality. When restricting the spatial area of analysis to mitigate omitted variable bias, the effect of racial composition becomes stronger while the income effect tapers off.

To assess the observed willingness to pay for race and income segregation I take the hedonic estimates and scale them by the estimated race and income differentials at the boundary. The results of my analysis suggest that 5-16% of observed house prices on the high income side are caused by demand for race and income segregation. The magnitude is highest for city centers, where housing supply is most inelastic, and small towns, where conservative voting dominance suggests preferences for segregation. The magnitude is smaller in suburban areas, where house prices capitalize income but not racial differences.

The results of this paper advance a literature seeking to explain house price differentials that appear correlated with neighborhood racial composition. Bajari and Kahn (2005) argue that the location choice of white households is driven by demand for housing size and neighborhoods with high human capital levels, resulting in the suburbanization of white households with the income to pay for such attributes. It follows that structural differences in degree obtainment means such neighborhoods are largely white. Yinger (2016) derives the conditions in which hedonic models represent household bids for neighborhood characteristics, revealing heterogenous preferences for Black and Latino neighborhoods. The bid rent analysis flexibly identifies a price premium for fully segregated White, Black, and Latino neighborhoods. Evidence has also shown Black and Latino households potentially pay a premium for equivalent housing (Bayer et al. 2017).

Section 2 is a description of the data used to implement the hedonic boundary discontinuity model. Section 3 provides conceptual grounding and traces out the house price gradient using raw data and a simple differences in means. Section 4 contains results of the main model estimates for the willingness to pay for neighborhood race and income characteristics. In Section 5 I conclude the paper.

2 Data

In 1911 Texas claimed the highest mileage of rail lines amongst all US States, and the New Century Atlas published that same year released maps of states and the railroads within. I use a digitized version of the 1911 railroad map provided by Atack (2013) to document physical neighborhood boundaries using ArcGIS software. I take two-mile segments of track to be considered a boundary, and the estimation sample includes boundaries with neighborhoods on both sides of the tracks. Figure 1 shows the physical layout of historic railroads across Texas.

Neighborhoods in the sample are defined by census block groups, population clusters described by contemporary characteristics and a median population of 1,480. The census provides coordinates for the block group center of population, an estimate of the central mass of residents living within block group boundaries. I use this spatial information to map neighborhood locations to the nearest two-mile railroad segment. Of the block group characteristics I am interested in median home values, household income, and racial composition.

In addition I use block group data from the census describing the median age of housing and the median number of rooms as proxies for housing quality.

The quality of local public amenities are proxied by public school data made available from the Texas Education Agency for 2321 elementary schools. For each school, TEA measures test score success by categorizing students as mastering, passing, or below passing fourth grade aptitude tests for math and reading. I use the maximum of math and reading scores for each school as a measure of test scores and a proxy for school quality. Texas Education Agency also provides campus level data for school spending, covering a smaller geographic area than school districts. For private consumption amenities I employ data from the Texas Comptrollers office providing the name, location and business type for all businesses collecting state sales taxes. Business types are categorized by North American Industry Classification System (NAICS) codes. I aggregate across several NAICS codes to create two categories, food venues and non-food retail establishments and compute totals within 1 mile of neighborhood centers.

The estimation sample includes 8,993 neighborhoods mapped to 1,215 railroad segments for an average of 7.4 neighborhoods per boundary, covering roughly 62% of Texas population (17.8M). Median income for both sides of each boundary segment is computed and neighborhoods are assigned to the *high amenity* or *low amenity* side according to which side has the higher median income. Summary stats are presented in Table 1 for the full sample and then by low and high amenity designation. Table 1 shows that house prices and incomes are 11-12% higher in high-amenity neighborhoods relative to the full sample. Descriptively, high amenity neighborhoods have a higher share of white residents and less minority residents.

Data describing local amenities are useful for my design as amenity quality is correlated with race and income composition along with local housing prices. Table 1 shows little observational differences in per-pupil spending and test score success across neighborhood types. There is also little difference in the count of food and retail consumption access across neighborhood type. Although the quantity measures do not account for differences in private amenity quality, there is variation in quantities that could be capitalized into house prices and thus I employ the restaurant and retails counts as control variables in the main regression.

3 Neighborhoods Near The Railroad Boundary

Section 3 pairs conceptual grounding with raw data analysis to motivate the econometric framework that follows in Section 4.2. I begin by describing the theoretical price gradient near railroad tracks as a function of pure disamenties caused by proximity to rail operation and failures. The central hypothesis is that deviations from the price gradient predicted purely by disamenities represent distortions caused by household sorting. I next show the presence of price distortions in the raw data, estimating spatial differences with a boundary fixed effects model. Estimates from the model illustrate descriptive differences in prices across the boundary space, and trace out heterogeneity in the price gradient for urban, suburban and remote locales.

3.1 Conceptual

50% of the census population count in Texas live within 1 mile of a railroad track, and 80% live within two miles, likely reflecting the historic benefit of proximity to economic activity between cities (Atack et al. 2010; Hodgson 2018). Nevertheless, there are a traunch of externalities posed on communities living closest to railroad operation. Trains of all type are loud (Wayson and Bowlby 1989), and there is evidence that exposure to diesel exhaust causes negative health outcomes (Garshick et al. 2004; Verma and Verter 2007). Malfunctions and stoppages are common, which can increase commute times and delay access to public services in critical times of need.[1]

[1] News articles and public opinion pieces document railroad malfunctions as a serious issue in several parts of the city. See: https://www.houstonchronicle.com/news/houston-texas/transportation/article/Cameras-data-on-blocked-train-crossings-shows-17168866.php, and https://houstonlanding.org/stopped-trains-block-houston-traffic-union-pacific/. Outlets are reporting similar stories nationwide. See:

The externalities described are characterized by spatial proximity to the physical railroad track. To elaborate on this point Figure 3 shows two curves. The first curve maps distance from the railroad track to the intensity of the disamenity, $\delta(D)$. Tntensity of the pure disamenity decreases exponentially in distance from the tracks, as noise and exhaust pollution is most poignant nearest to the tracks. The second curve in Figure 3 shows the theoretical house price gradient P_h as a function of distance to the track through the pure disamenity $\delta(D)$. Naturally, house prices are predicted by housing quality and other neighborhood characteristics X_h. The predicted house price is increasing in distance from the tracks, as pure disamenities wane across space.

Under the assumption that the pure disamenity is indifferent boundary side, I make the following testable claim. Absent systematic sorting, X_h is randomly assigned and the spatial distribution of amenities is not predicted by side of track. If amenities X_h vary smoothly across the boundary, and $\delta(D)$ is symmetric across the boundary, the price gradient P_h can be mirrored across the Y-axis of Figure 3. There will be no discrete jumps in the price, and the slope on both sides should be symmetric. Taking symmetry in the pure disamenities as given, deviations from a symmetric price gradient are caused entirely by differences in neighborhood amenities X_h. In the absence of sorting and segregation the amenities on one side of the railroad boundary should be indistiguishable from the other.

3.2 Price Differentials in the Raw Data

In this section the spatial data described in Section 2 is employed to estimate a descriptive analog to the price gradient in Figure 3, using only the raw data. Following directly from Bayer et al. (2007), spatial differences in the price of neighborhood i, at distance j from railroad boundary b can be modeled as

$$Log(P_{ijb}) = \alpha_0 + \sum_{j \neq 0}^{J} \alpha_j \mathbb{I}_j + \theta_b + \epsilon_{ijb} \tag{1}$$

https://www.washingtonpost.com/nation/interactive/2023/long-trains-block-intersections-paramedics/.

Arranging block groups into half-mile distance bins relative to the boundary, Equation 1 estimates the average price difference for each bin relative to block groups closest to the tracks ($j = 0$). Including boundary fixed effects θ_b, the model exploits within-boundary variation in neighborhood prices, measured by the census block group median home values. Taking the price outcomes as logs, the $\hat{\alpha}_j$ coefficients trace out percent differences in prices across space relative to the boundary.

To test the symmetry condition I plot the $\hat{\alpha}_j$ coefficients arranged by distance in Figure 4. The raw data for both sides of the train tracks shows prices that are lowest near the boundary but increasing in distance from the tracks. Moving from left to right across Figure 4 is akin to physically crossing from the low amenity to the high amenity side of a particular boundary, and there is an identifiable jump in prices. Comparing bin -0.5 to bin 0.5, I find that prices on the high amenity side are 20-25% higher than those at the same distance on the low amenity side. Further, I am able to reject the symmetry hypothesis and assert that neighborhood characteristics, including demographic composition, are not randomly assigned across the boundary space.

3.3 Heterogeneity in the Raw Data

Figure 4 contains estimates of the unconditional mean price differential across train tracks in locations of every type. This may mask heterogeneity in the price differentials where housing supply is inelastic, like big city centers, versus small towns where supply is less land-constrained and thus more elastic (Gyourko et al. 2008; Saiz 2010). A testable hypothesis is that inelastically supplied markets will have larger price differentials, ceteris parabus. To that end I produce the price plots for a variety of urban and rural comparison groups in Figures 5-8. Figure 5 is a broad comparison of price differentials in large and small metropolitan areas (cities and suburbs) to price differentials in small towns. Figures 6 and 7 compares city centers to suburbs, and Figure 8 is a plot of differentials in small towns.

When taken together, Figures 5-8 are descriptive evidence that price differentials exist in neighborhoods of all types. The raw price gradients in Figure 5 are indistiguishable except the high amenity price slope is steeper in small towns, and the low amenity price slope is flatter in metropolitan areas. Figures 6 and 7 show steeper price increases for the high amenity side in city centers relative to suburbs, for both big and small metropolitan areas. In the majority of locales, there is no symmetry in the price gradient when comparing neighborhoods at a particular boundary.

4 Estimation

The descriptive data visualization traces out price differentials at neighborhood boundaries across a variety of geographies. To obtain estimates for magnitude of the differentials I adopt the discontinuity model from Dell (2010), taking prices as an outcome of spatial features and neighborhood characteristics. The baseline specification is

$$Log(P_{ib}) = \alpha + \delta HighSide_i + f(Distance_{ib}) + X'_{ib}\beta + \theta_b + \epsilon_{ib}. \quad (2)$$

The boundary θ_b fixed effects allow me to compare house prices at the same boundary, conditional on local amenities X_{ib} and distance to the track, $Distance_{ib}$. Given the theoretical gradient and raw data plots predict price increases to diminish in distance, f($Distance_{ib}$) is modeled as quadratic. The flexible model yields an estimate for the price differential ($\hat{\delta}$), and hedonic estimates ($\hat{\beta}$) for the house price capitalization of each neighborhood characteristic that varies at the boundary.

The model in Equation 2 is a standard boundary discontinuity model, typically used to obtain hedonic estimates of a particular public amenity that changes across an administrative boundary (Keele and Titiunik 2015). The fundamental challenge in such a design is that neighborhood racial composition is correlated local income, and both are correlated with unobserved neighborhood amenities. The unobservables are capitalized into house prices and loaded onto coefficients for race and income. One approach is to assume unobserved characteristics captured in variation of house prices are held fixed across a small spatial area. For example, test score differences cause house prices premiums when other neighborhood characteristics are held constant across a small space divided by the school attendance boundary (Black 1999). Further, employing boundary fixed effects in such a manner yields hedonic estimates of school quality that are distinguishable from hedonic estimates of neighborhood racial composition (Bayer et al. 2007).

Railroads are physical boundaries in which many neighborhood attributes are predicted to change. Thus I am extending the boundary discontinuity design under relaxed conditions as there is not a single parameter of interest in my model. In framing the results I argue that $\hat{\delta}$ estimates the gross price premium for the high amenity side of the boundary when no additional controls are included. Subsequently adding neighborhood covariates X to the model not only yields hedonic estimates $\hat{\beta}$ for each characteristic, but $\hat{\delta}$ will trend towards zero in the full model as the gross price premium is explained by the variables added to the regression. As Figures 9-11 show, there is distinct variation in neighborhood income, racial demographics, and housing quality at the boundary. The high amenity side of the tracks has identifiably higher incomes, a larger share of white residents white residents, and higher quality housing.

Section 4.1 describes estimates for the gross price premium for the entire sample and heterogeneity across locales. Section 4.2 details the house price capitalization of neighborhood race and income, relying on variation in neighborhood composition at the boundary. In Section 4.3 I analyze the willingnes to pay for race and income composition by combining the hedonic estimates with the boundary composition differentials, to calculate the percent of house prices attributed to preferences for income and race. I last put forth a test in Section 4.4 using fictional railroad boundaries to detect results that may arise due to data construction.

4.1 The Gross Price Premium

When comparing neighborhoods at the opposite side of the same boundary, the gross price premium represents the unconditional market valuation of all differences in housing quality and neighborhood amenities that vary at the boundary. Estimates for the gross price premium are presented in Table 2, and are akin to the mean boundary effect in Figure 4. Moving from left from model 1 to 3 additional controls are added. First are boundary fixed effects, then controls for distance to the boundary are modeled as a quadratic function of miles. Columns 3-5 of Table 2 are estimates from gradually reducing the spatial area of study from a 2 mile to 0.5 mile radius from the boundary on either side. Shrinking down the spatial area will mitigate bias in the estimates as unobserved confounders are plausibly held constant across a small space, trading off data points as the sample size decreases from 6,688 to 2,435 block groups.

The estimates in Table 2 show an average house price differential of 20% for the high income side of the boundary when block groups up to 2 miles from the boundary are included in the sample. Reducing the distance bandwidth from 2 to 1 to 0.5 mile radius, the price differential monotonically decreases to 13% across the smallest spatial area. In the context of Figure 4, the results of Table 2 suggest the discrete price differential at the boundary is 13%. Since the gross price differential increases to 20% when including observations further from the boundary, I conclude tat the price gradient is steeper on the high income side of the tracks.

The results of Table 2 represent demand pricing for neighborhood characteristics on the high income side, assuming constant elasticity of housing supply. However, housing supply is on average more inelastic in large urban areas than relatively smaller urban areas, suburbs, and towns. The elasticity differences mean that demand shifts will be met with larger price increases when housing supply is inelastic. Further, if political affiliation predicts preferences for race and income integration, the polarization of voters across urban and rural areas suggest the potential for heterogeneity in the results.

Table 3 contains estimates for the gross price differential across samples based on geographic locale. Results in the upper panel are for the one mile boundary radius, and the 0.5 mile radius sample is used in the lower panel. Each model employs boundary fixed effects and the distance polynomial. The model predicts statistically significant boundary differentials across all locale types for the 1 mile radius sample. For the 0.5 mile sample the differentials are smaller and only identifiable for cities and small towns, but not in suburban areas. Considering sample size the models appear underpowered, however I explore heterogeneity across locales for all results to follow.

4.2 The House Price Capitalization of Race and Income

To explore demographic changes at the boundary, the analysis of this section first takes neighborhood income, the share of white, black, and Latino residents each as an outcome of the main model in equation 2. After tracing out the demographic changes I then return to estimating the hedonic model with prices as an outcome of neighborhood income and the share of white residents. Estimates presented for the capitalization of race and income are conditional on variation in housing quality, private consumption amenities, and public school quality. I discuss the results of each analysis across urban, suburban, and rural geographies.

Figures 12-15 contain point estimates for the discrete change in each demographic characteristic at the neighborhood boundary, for the 1 mile sample in the upper panel and the 0.5 mile sample in the lower panel. Each point is an estimate along with the 95% confidence interval for the *highside* effect in Model 2, taking each headline neighborhood characteristic as an outcome. The results in Figure 12 show an income differential of 25% for the entire 1 mile sample, ranging between 19% and 27% across locales. The 0.5 mile sample reveals stable estimates of the income effect across locales, except for a noticeably smaller income differential in big city suburban areas.

In the subsequent Figures 13-15 are estimates for discontinuities in the racial profile of neighborhoods at the railroad boundary. Pooling all locales at both the 1 mile and 0.5 mile radius, the model predicts identifiable increases in the white population share and decreases in the black and Latino population share. The estimates are more precise for white and black population differentials than Latino differentials, and the white-black composition effects are largest in small towns. In big city centers there are identifiable differentials for each of the race and ethnic groups.

With evidence of house price differentials and distinct composition changes at the railroad boundary, I next turn to estimating the capitalization of neighborhood race and income composition. To obtain the hedonic estimates I leverage the race and income differentials as identifying variation. The assumption is that historic railroad boundaries only affect contemporary house prices through variation in neighborhood amenities when conditioning on distance as a proxy for the pure disamenity. Estimates for hedonic capitalization of race and income are presented in Table 4, and each specification includes boundary fixed effects and the distance polynomial. Moving from model 1 to 3, housing quality controls and the amenity controls are added. Across columns 3-5 of Table 4 the spatial bandwidth is again reduced from 2 miles to 0.5 miles.

The estimates in Table 4 show that increasing white population share by 0.1 will increase house prices by 6.4-7.2%, and increasing neighborhood income by 10% will increase house prices by 3.2-4.4%. The statistically significant coefficient in row 1 of columns 1 - 3 imply that unobserved amenities predict 3-4% of the price premium for the high income side across the 2 mile sample space. This amenity is mitigated in the results across smaller spatial areas in columns 4 and 5. Two broad takeaways emerge from the results in Table 4. First, the race effect is larger than the income effect. Second, the race effect becomes stronger and more precise under conditions in which unobserved amenities are plausibly held constant, while the income effect weakens.

Table 5 contains estimates for the house price capitalization of race and income across different locales. The second row of Table 5 is documentation of substantial heterogeneity in the capitalization of neighborhood race across locales, as the price response in big cities and small towns is quantitatively similar and twice the magnitude of other locations. In Table 6 the sample is restricted to the 0.5 mile radius the capitalization of race remains statistically significant in city centers and small towns. In suburbs big and small the positive capitalization of white composition share is small and not identified. In all locations except big city suburbs, the effect of increasing white population share by 0.1 ($0\% < \Delta P < 9.3\%$) is larger than the price response to a 10% neighborhood income increase ($0\% < \Delta P < 4.2\%$).

4.3 Willingness To Pay

How much do households pay to live near whiter and richer neighbors in a segregated environment? To study this question the analysis of sections 4.1 and 4.2 is extended to calculate the will to pay for income and racial composition in dollars of home prices. The first step is scale the price capitalization estimate for income and white composition by the boundary differential of each variable, then apply this multiple to the mean house price at the boundary. The procedure yields an implied dollar value of housing paid solely for the observed race and income differentials. Results of the analysis are presented in Tables 7 and 8.

Estimates in Table 7 are presented separately for the capitalization of income and white composition. Each of the two upper panels take the hedonic coefficients from Table 5 and the differentials from Figures 12 - 13 to compute the income and composition multiples, which are used to convert empirical estimates to dollar values. The bottom panel of Table 7 contains the mean home value across the 1 mile radius and the implied willingness to pay for each demographic separately, then in total, for each geographic locale. There is substantial heterogeneity in mean home values and preferences for income and racial composition across locales. Home values are highest in big cities and the surrounding suburbs, and lower in smaller cities relative the surrounding suburbs. Home values are lowest in small towns, but the coefficients for income and racial composition reveal comparable preferences in small towns and big city centers. As a percentage of home values, the willingess to pay for race and income segregation is highest in big cities (16%) and small towns (14%), and lowest in suburban areas (5-8%).

The analysis is repeated for the 0.5 mile radius sample in Table 8. In all locales, except small city centers, the willingness to pay for segregation is smaller as the boundary space is reduced. As the bandwidth shrinks, omitted variables positively correlated with income and white composition are presumed identical across the boundary. The tradeoff for precision gains when reducing the spatial area is substantial data reduction and statistical power when subdividng the sample by geography. Notwidthstanding, my results show that households in small towns are willing to pay up to 12% of home value for higher income neighborhoods and a higher share of white neighbors, similar to estimates for cities (10-11.4%). To unpack whether preferences for race or income characteristics are behind the willingness to pay estimates I compare the WTP dollar values in the bottom panel of 8. I find that the income effect strictly dominates the race effect across all locales, and that the race effect is very small in suburbs and small cities. The race effect is sizeable in big cities and small towns, and dominates the income effect in such locales when inspecting the results of the 1 mile sample in Table 7.

4.4 Placebo

The results of my analysis rely on the plausibly exogenous variation in neighborhood characterstics at the railroad boundary. It is entirely possible that the arrangement of the data near railroad boundaries and the determination of the high income side may produce mechanical relationships in the data not representative of true effects. A way to test for such mechanical bias is to overlay fictional train tracks across the Texas map and assign block groups to the made up boundaries. It is relatively straightforward to shift the underlying railroad map to the north and east, then undertaking the same procedure of mapping the block groups to boundaries, identifying the high income side, and estimating the main model with boundary fixed effects. I conduct this analysis for prices and neighborhood composition variables and report the results in Table 9.

If the main results of the paper are produced mechanically through data arrangement, there will be identifiable boundary effects using the placebo train tracks. In column 1 of Table 9 I show there is no unconditional house price premium for the high income side of the placebo boundaries, as compared to the 13% premium I find in column 5 of Table 4. There is no identifiable income differential at the placebo boundaries as compared to differentials that range from 14% to 25% in the second row of Table 7. For each of the racial demographics I find no differentials at the fictional boundaries. The results of this simple robustness test at minimum rule out mechanically biased estimates due to data arrangement.

5 Conclusion

Households reveal preferences for segregation through a willingness to pay a material share of housing costs for race and income segregation. Strikingly, similar preferences for race and income segregation exist in big city centers as in small towns, although voting bases are much much different in the two locales on average. Since housing supply in big city centers is relatively inelastic, small demand changes will cause sharp changes in market prices. Considering the differences in supply elasticity, I conclude that the demand for segregation must be stronger in small towns to produce a price effect similar to that in big cities. In suburban areas the evidence is stronger for income segregation, but demand for racial segregation is less identifiable.

Economists have come to understand that railroads segregate cities, taking the divided up neighborhoods created by tracks as subunits of analysis when studying the broad implications of segregation across a housing market. The contribution of this paper is to show just how that process plays out a local level, and the direct effects are clear. When physical boundaries delineate neighborhoods, there is evidence of household sorting on race and income even when controlling for differences in neighborhood amenities. However, my analysis stops short of decoupling household expectations for future amenities based on current racial composition.

A question for future studies is whether households can be indifferent to living next to same-race neighbors, while making location decisions infer future public and private consumption amenities based on today's racial composition when making location decisions

Perhaps no other transit infrastructure in the US has a history quite like railroads. Tracks laid at industry peak near the turn of the 20th century are still utilized today, and unused tracks are still physically present all over the country. When most rail lines were initially put in place much of the country lived under de jure segregation physically restricting non-whites from neighborhoods, public spaces, and private venues. The placement of railroads served to further enforce the boundaries between race and ethnic groups. The results of this paper show that neighborhoods near railroads tracks in Texas remain segregated by income and race today.

There are qualities about Texas that can provide insights to understand segregation in many other places. The state has large cities with expansive suburbs and remote rural towns. In addition there have been large waves of migration from within and outside of US borders. As such neighborhoods have been exposed to many forces that could theoretically produce higher levels of economic integration. Yet, this study finds that segregation persists near railroads, and house prices capitalize the stark differences in neighborhood amenities. Driven by race and income characteristics, not amenities, the sticker price premium for a house on the high income side of the tracks to is near 12-20%.

Declarations

Ethical Approval

Not applicable

Funding

Not applicable

Availability of Data and Materials

This study uses publicly available data from the US Census Bureau, Texas Department of Education, Texas Department of Transportation, and the Texas Comptroller. Replication files are available at www.kdwhaley.com.

References

Aaronson, D., D. Hartley, and B. Mazumder (2021): "The effects of the 1930s HOLC "redlining" maps," *American Economic Journal: Economic Policy*, 13, 355–392.

Ananat, E. O. (2011): "The wrong side (s) of the tracks: The causal effects of racial segregation on urban poverty and inequality," *American Economic Journal: Applied Economics*, 3, 34–66.

Atack, J. (2013): "On the use of geographic information systems in economic history: The american transportation revolution revisited," *The Journal of Economic History*, 73, 313–338.

Atack, J., F. Bateman, M. Haines, and R. A. Margo (2010): "Did railroads induce or follow economic growth?: Urbanization and population growth in the American Midwest, 1850–1860," *Social Science History*, 34, 171–197.

Bajari, P. and M. E. Kahn (2005): "Estimating housing demand with an application to explaining racial segregation in cities," *Journal of business & economic statistics*, 23, 20–33.

Bayer, P., M. Casey, F. Ferreira, and R. McMillan (2017): "Racial and ethnic price differentials in the housing market," *Journal of Urban Economics*, 102, 91–105.

Bayer, P., F. Ferreira, and R. McMillan (2007): "A unified framework for measuring preferences for schools and neighborhoods," *Journal of political economy*, 115, 588–638.

Bischoff, K. (2008): "School district fragmentation and racial residential segregation: How do boundaries matter?" *Urban Affairs Review*, 44, 182–217.

Black, S. E. (1999): "Do better schools matter? Parental valuation of elementary education," *The quarterly journal of economics*, 114, 577–599.

Card, D., A. Mas, and J. Rothstein (2008): "Tipping and the Dynamics of Segregation," *The Quarterly Journal of Economics*, 123, 177–218.

Chyn, E., K. Haggag, and B. A. Stuart (2022): "The effects of racial segregation on intergenerational mobility: Evidence from historical railroad placement," Tech. rep., National Bureau of Economic Research.

Cook, L. D., M. E. Jones, T. D. Logan, and D. Rosé (2022): "The Evolution of Access to Public Accommodations in the United States," *The Quarterly Journal of Economics*.

Cox, R., J. P. Cunningham, A. Ortega, and K. Whaley (2022): "Black Lives: The High Cost of Segregation," Tech. rep., Washington Center for Equitable Growth.

Cutler, D. M., E. L. Glaeser, and J. L. Vigdor (1999): "The rise and decline of the American ghetto," *Journal of political economy*, 107, 455–506.

Dell, M. (2010): "The persistent effects of Peru's mining mita," *Econometrica*, 78, 1863–1903.

Garshick, E., F. Laden, J. E. Hart, B. Rosner, T. J. Smith, D. W. Dockery, and F. E. Speizer (2004): "Lung cancer in railroad workers exposed to diesel exhaust," *Environmental Health Perspectives*, 112, 1539–1543.

Gyourko, J., A. Saiz, and A. Summers (2008): "A new measure of the local regulatory environment for housing markets: The Wharton Residential Land Use Regulatory Index," *Urban Studies*, 45, 693–729.

Hodgson, C. (2018): "The effect of transport infrastructure on the location of economic activity: Railroads and post offices in the American West," *Journal of Urban Economics*, 104, 59–76.

Keele, L. J. and R. Titiunik (2015): "Geographic boundaries as regression discontinuities," *Political Analysis*, 23, 127–155.

Monarrez, T. E. (2022): "School attendance boundaries and the segregation of public schools in the us," *American Economic Journal: Applied Economics*.

Rosen, S. (1974): "Hedonic prices and implicit markets: product differentiation in pure competition," *Journal of political economy*, 82, 34–55.

Saiz, A. (2010): "The geographic determinants of housing supply," *The Quarterly Journal of Economics*, 125, 1253–1296.

Schelling, T. C. (1971): "Dynamic models of segregation," *Journal of mathematical sociology*, 1, 143–186.

Verma, M. and V. Verter (2007): "Railroad transportation of dangerous goods: Population exposure to airborne toxins," *Computers & operations research*, 34, 1287–1303.

Wayson, R. and W. Bowlby (1989): "Noise and air pollution of high-speed rail systems," *Journal of transportation engineering*, 115, 20–36.

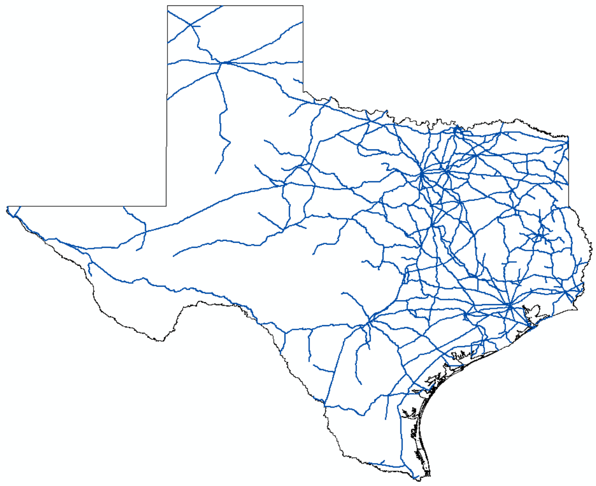
Witte, A. D., H. J. Sumka, and H. Erekson (1979): "An estimate of a structural hedonic price model of the housing market: an application of Rosen's theory of implicit markets," *Econometrica: Journal of the Econometric Society*, 1151–1173.

Yinger, J. (2015): "Hedonic markets and sorting equilibria: Bid-function envelopes for public services and neighborhood amenities," *Journal of Urban Economics*, 86, 9–25.

——— (2016): "Hedonic estimates of neighborhood ethnic preferences," *Public Finance Review*, 44, 22–51.

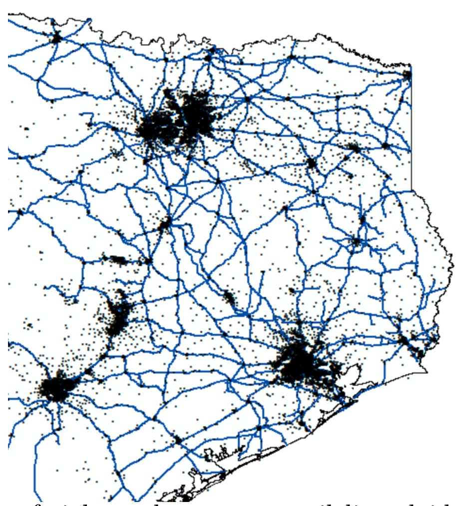
6 Figures

Figure 1: Railroads in Texas : 1911



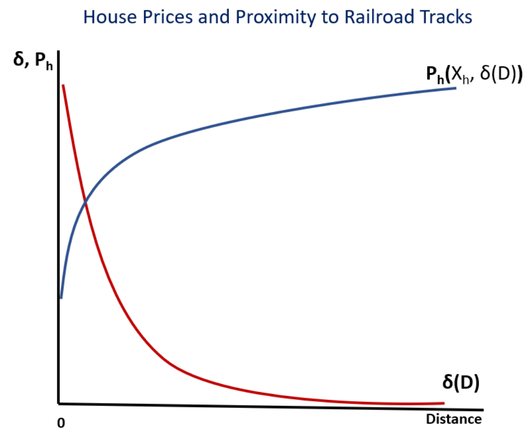
Notes: The blue lines represent freight and passenger rail lines laid in Texas as of the early 1900s. Source: 1911 New Century Map digitized by Jeremy Atack, Department of Economics, Vanderbilt University.

Figure 2: Current Texas Neighborhoods Along Historic Railroad Tracks



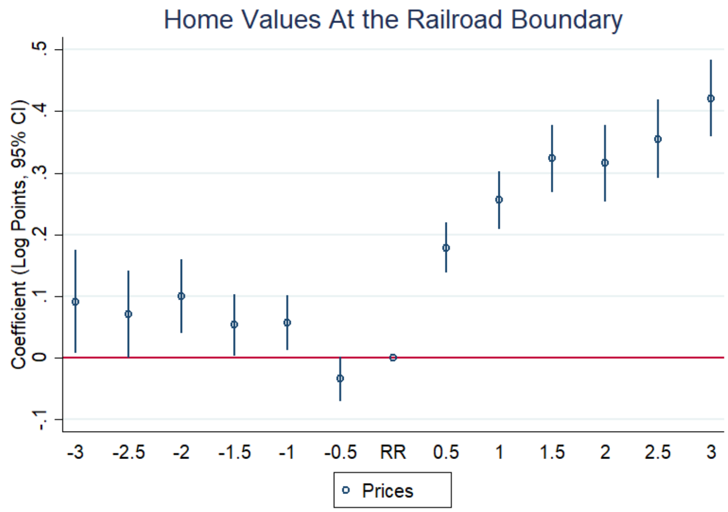
Notes: The blue lines represent freight and passenger rail lines laid prior to 1911. Each black dot represents a the population center for a census block group as of 2018. Source: 1911 New Century Atlas and Census/IPUMS National Historical GIS.

Figure 3: Price and Disamenity Curves



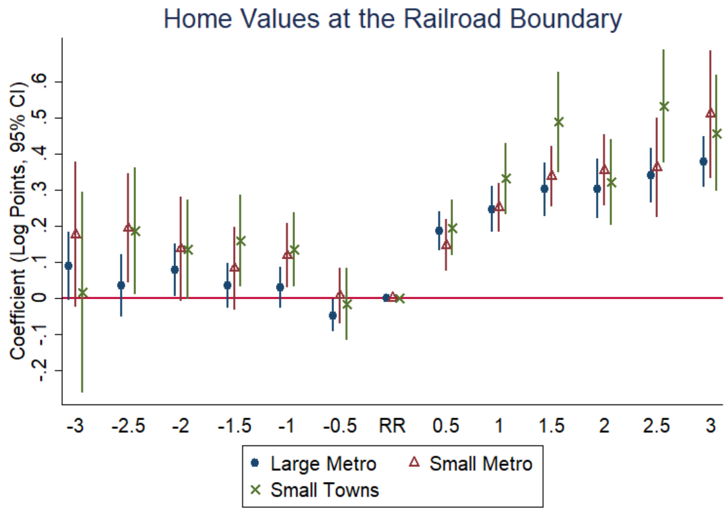
Notes: Theoretical house price gradient, where prices are a function of neighborhood attributes $\mathbf{X}$ and the pure disamenity of living near tracks $\delta(D)$. The pure disamenity is assumed to be decreasing in distance from the tracks.

Figure 4: Price and Disamenity Curves



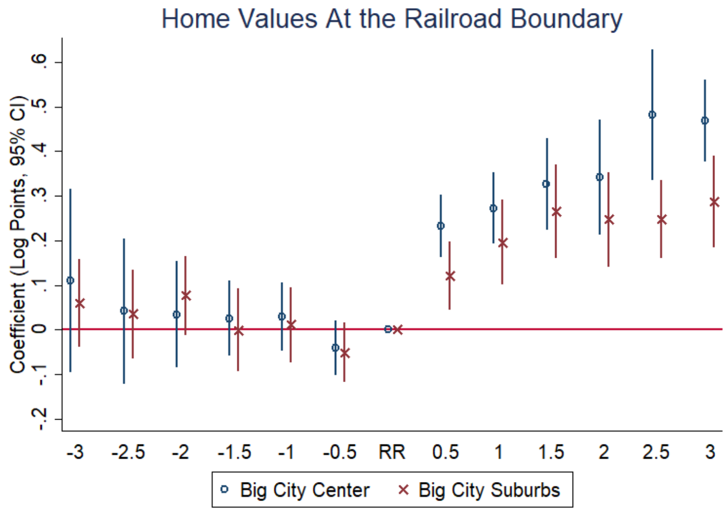
Notes: Each plot includes the $\hat{\alpha}_j$ coefficients from equation 1 for the outcome variable of interest. Neighborhoods are grouped into bins based on distance to the railroad boundary with the control group being those 0.5 miles or less from the tracks. The control group is marked by a red dot, and negative distances represent the low-amenity side of the boundary. By including boundary fixed effects, the model estimates the first-difference in mean for outcome $\mathbf{Y}$ in each bin, relative to control group neighborhoods at the same railroad segment.

Figure 5: Price and Disamenity Curves



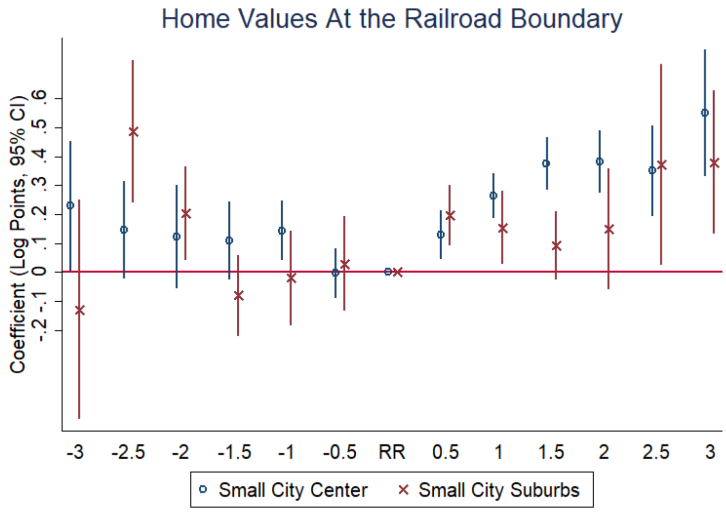
Notes: Each plot includes the $\hat{\alpha}_j$ coefficients from equation 1 for the outcome variable of interest. Neighborhoods are grouped into bins based on distance to the railroad boundary with the control group being those 0.5 miles or less from the tracks. The control group is marked by a red dot, and negative distances represent the low-amenity side of the boundary. By including boundary fixed effects, the model estimates the first-difference in mean for outcome $\mathbf{Y}$ in each bin, relative to control group neighborhoods at the same railroad segment.

Figure 6: Price and Disamenity Curves



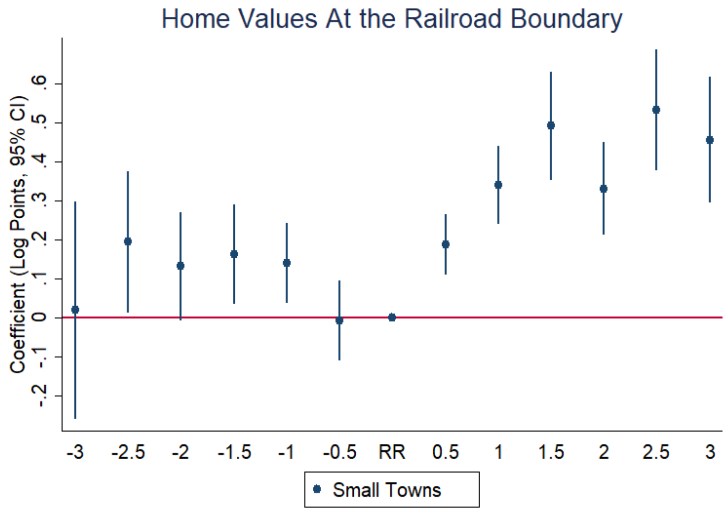
Notes: Each plot includes the $\hat{\alpha}_j$ coefficients from equation 1 for the outcome variable of interest. Neighborhoods are grouped into bins based on distance to the railroad boundary with the control group being those 0.5 miles or less from the tracks. The control group is marked by a red dot, and negative distances represent the low-amenity side of the boundary. By including boundary fixed effects, the model estimates the first-difference in mean for outcome $\mathbf{Y}$ in each bin, relative to control group neighborhoods at the same railroad segment.

Figure 7: Price and Disamenity Curves



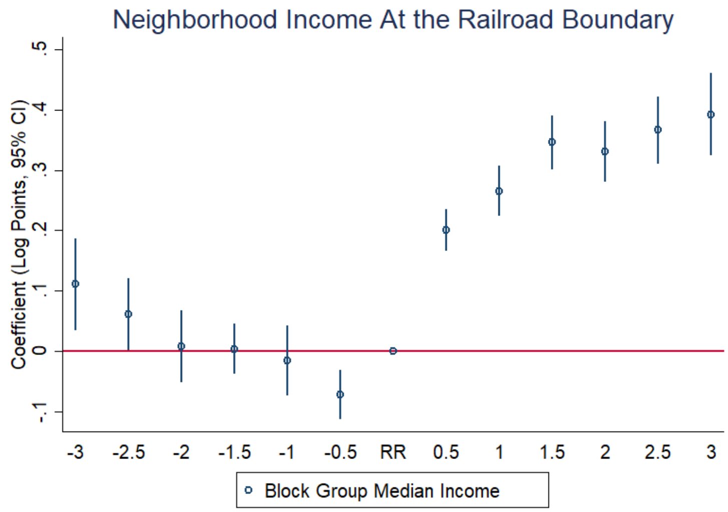
Notes: Each plot includes the $\hat{\alpha}_j$ coefficients from equation 1 for the outcome variable of interest. Neighborhoods are grouped into bins based on distance to the railroad boundary with the control group being those 0.5 miles or less from the tracks. The control group is marked by a red dot, and negative distances represent the low-amenity side of the boundary. By including boundary fixed effects, the model estimates the first-difference in mean for outcome $\mathbf{Y}$ in each bin, relative to control group neighborhoods at the same railroad segment.

Figure 8: Price and Disamenity Curves



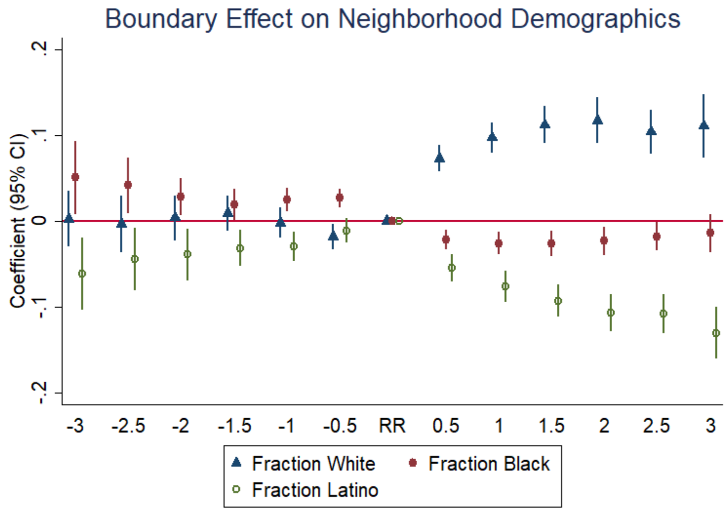
Notes: Each plot includes the $\hat{\alpha}_j$ coefficients from equation 1 for the outcome variable of interest. Neighborhoods are grouped into bins based on distance to the railroad boundary with the control group being those 0.5 miles or less from the tracks. The control group is marked by a red dot, and negative distances represent the low-amenity side of the boundary. By including boundary fixed effects, the model estimates the first-difference in mean for outcome $\mathbf{Y}$ in each bin, relative to control group neighborhoods at the same railroad segment.

Figure 9: Price and Disamenity Curves



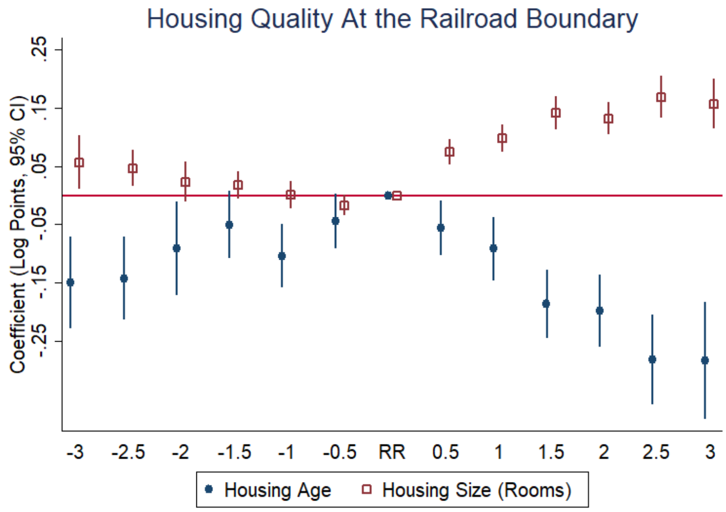
Notes: Each plot includes the $\hat{\alpha}_j$ coefficients from equation 1 for the outcome variable of interest. Neighborhoods are grouped into bins based on distance to the railroad boundary with the control group being those 0.5 miles or less from the tracks. The control group is marked by a red dot, and negative distances represent the low-amenity side of the boundary. By including boundary fixed effects, the model estimates the first-difference in mean for outcome $\mathbf{Y}$ in each bin, relative to control group neighborhoods at the same railroad segment.

Figure 10: Price and Disamenity Curves



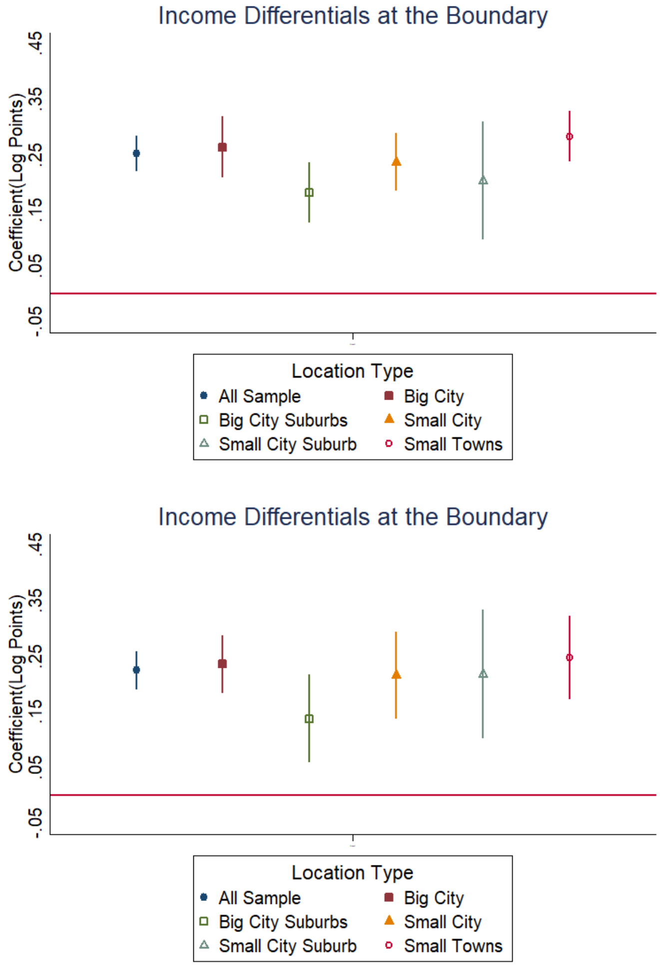
Notes: Each plot includes the $\hat{\alpha}_j$ coefficients from equation 1 for the outcome variable of interest. Neighborhoods are grouped into bins based on distance to the railroad boundary with the control group being those 0.5 miles or less from the tracks. The control group is marked by a red dot, and negative distances represent the low-amenity side of the boundary. By including boundary fixed effects, the model estimates the first-difference in mean for outcome $\mathbf{Y}$ in each bin, relative to control group neighborhoods at the same railroad segment.

Figure 11: Price and Disamenity Curves



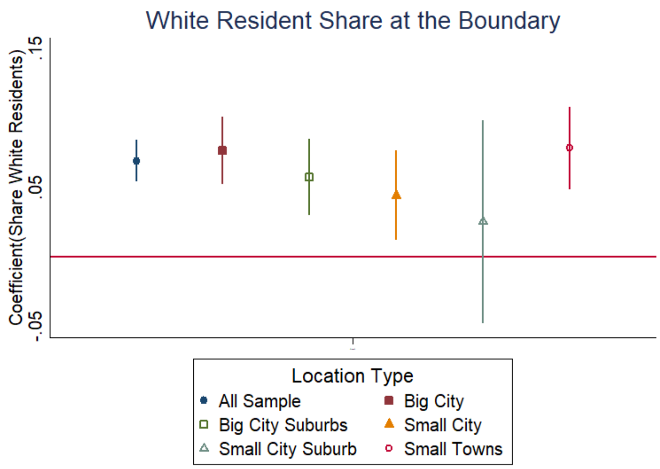
Notes: Each plot includes the $\hat{\alpha}_j$ coefficients from equation 1 for the outcome variable of interest. Neighborhoods are grouped into bins based on distance to the railroad boundary with the control group being those 0.5 miles or less from the tracks. The control group is marked by a red dot, and negative distances represent the low-amenity side of the boundary. By including boundary fixed effects, the model estimates the first-difference in mean for outcome $\mathbf{Y}$ in each bin, relative to control group neighborhoods at the same railroad segment.

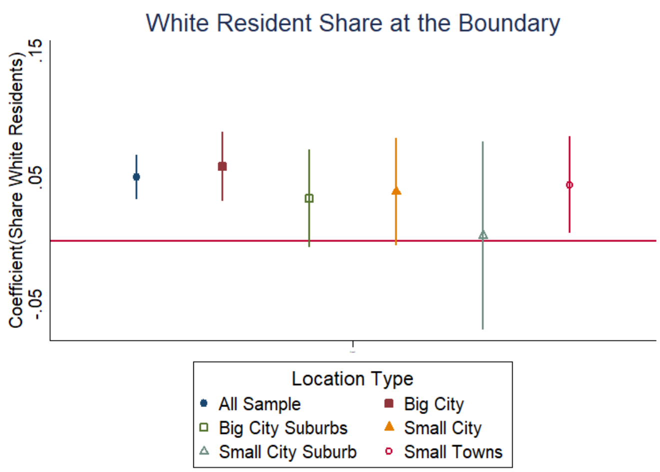
Figure 12: Spatial Variation of Select Neighborhood Characteristics



Notes: Each panel includes the first-difference estimates of the mean for various outcomes, by bin, relative to the neighborhoods nearest the railroad tracks.

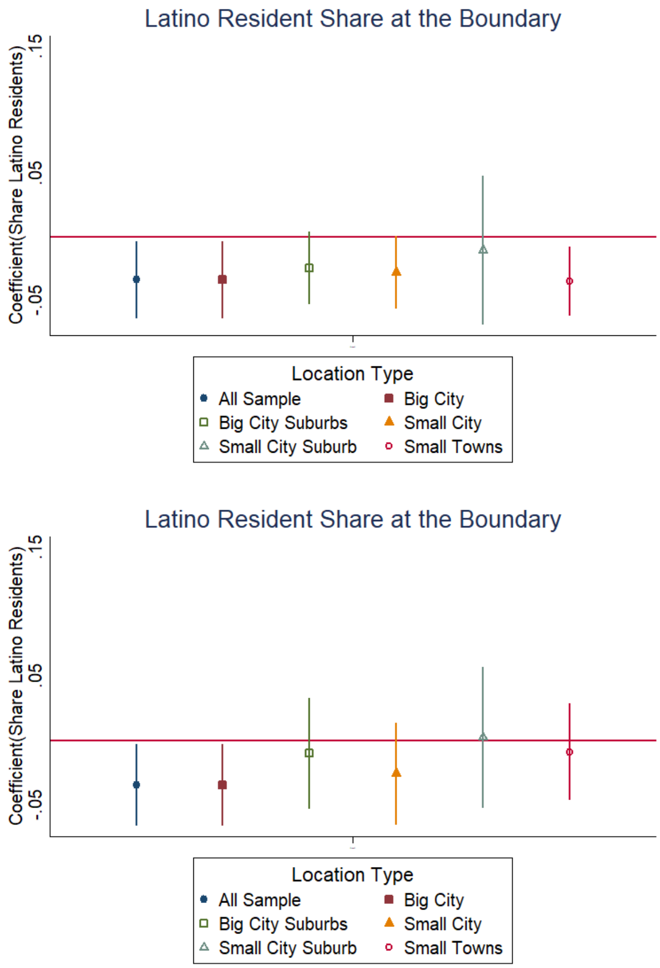
Figure 13: Spatial Variation of Select Neighborhood Characteristics





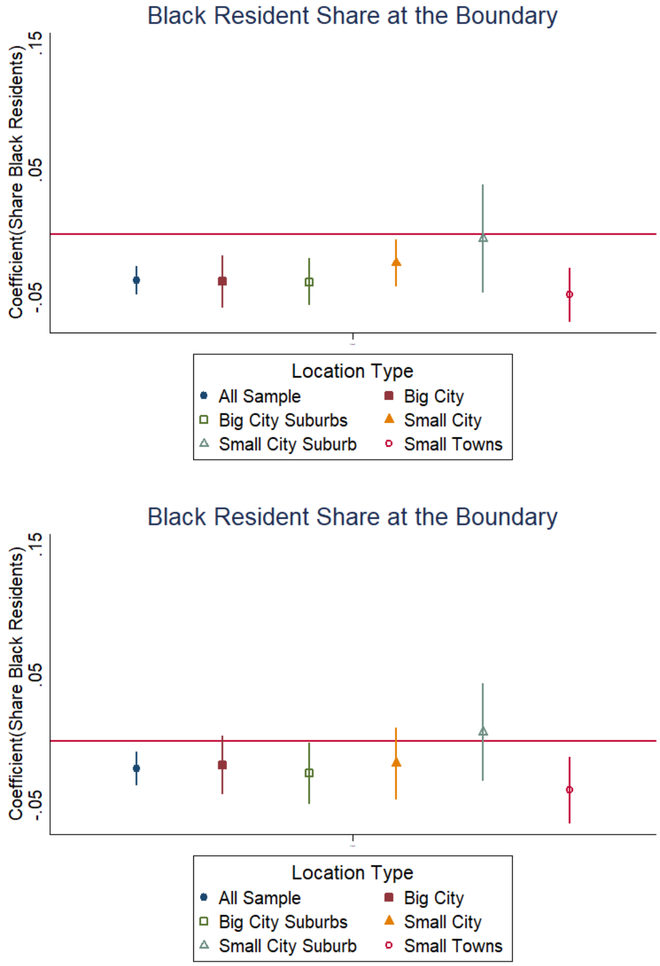
Notes: Each panel includes the first-difference estimates of the mean for various outcomes, by bin, relative to the neighborhoods nearest the railroad tracks.

Figure 14: Spatial Variation of Select Neighborhood Characteristics



Notes: Each panel includes the first-difference estimates of the mean for various outcomes, by bin, relative to the neighborhoods nearest the railroad tracks.

Figure 15: Spatial Variation of Select Neighborhood Characteristics



Notes: Each panel includes the first-difference estimates of the mean for various outcomes, by bin, relative to the neighborhoods nearest the railroad tracks.

7 Tables

Table 1: Block Group Summary Statistics

	Full Sample	Low Side	High Side
Median House Price	185,730.99	163,444.70	208,100.42
	(165865.9)	(127689.4)	(194342.9)
Median Income	66,970.32	59,123.43	74,846.49
	(37215.2)	(30689.5)	(41307.2)
Fraction White	0.40	0.36	0.43
	(0.284)	(0.275)	(0.288)
Fraction Latino	0.41	0.44	0.39
	(0.302)	(0.299)	(0.302)
Fraction Black	0.12	0.14	0.11
	(0.168)	(0.180)	(0.154)
Median Number of Rooms	5.55	5.35	5.74
	(1.256)	(1.132)	(1.342)
Median Housing Age	54.23	50.32	58.16
	(190.3)	(168.2)	(210.1)
Per-Pupil Spending	7,548.97	7,633.68	7,463.94
	(1982.1)	(2096.9)	(1856.2)
Standardized Test Mastery	0.27	0.26	0.29
	(0.124)	(0.114)	(0.132)
Food Venues / mi	9.80	10.00	9.60
	(11.58)	(12.25)	(10.87)
Retail Venues/ mi	33.38	33.42	33.34
	(29.14)	(29.63)	(28.65)
Observations	6688	3295	3393

Notes: Summary stats for block groups within 2 miles of a railroad boundary. Each unit of observation is a census block group, and the data are values from the ACS 5-year estimates. Source: IPUMS NHGIS data finder.

Table 2: Estimates of the Unconditional House Price Premium

	(1)	(2)	(3)	(4)	(5)
High Income Side	0.18***	0.20***	0.20***	0.17***	0.13***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
N	6688	6688	6688	4360	2435
Boundary FE		×	×	×	×
Distance Polynomial			×	×	×
Distance Bandwidth	2mi	2mi	2mi	1mi	0.5mi
r2	0.02	0.73	0.74	0.77	0.83

Notes: Standard errors reported and are clustered at the boundary level. Models 2-5 include boundary fixed effects, and models 3-5 include a quadratic function of distance to the boundary to control for spatial variation in the house price gradient near the boundary. The distance bandwidth describes the spatial area of analysis as a radius from the railroad boundary on either side.

Table 3: Unconditional Price Premium by Locale

	(1) Big City	(2) Big Suburb	(3) Small City	(4) Small Suburb	(5) Small Town
1 Mile Bandwidth					
High Income Side	0.19***	0.14***	0.13***	0.14*	0.12***
	(0.04)	(0.03)	(0.04)	(0.06)	(0.03)
0.5 Mile Bandwidth					
High Income Side	0.14***	0.08	0.12*	0.09	0.11*
	(0.04)	(0.05)	(0.05)	(0.07)	(0.05)
N (1mile)	1785	849	761	172	791
N (0.5mile)	991	414	421	94	512

Notes: Standard errors reported and are clustered at the boundary level. All models include boundary fixed effects and a quadratic function of distance to the boundary to control for spatial variation in the house price gradient near the boundary. The distance bandwidth describes the spatial area of analysis as a radius from the railroad boundary on either side.

Table 4: Hedonic Estimates of Neighborhood Amenities

	(1)	(2)	(3)	(4)	(5)
High Income Side	0.04***	0.04***	0.03*	0.02	0.01
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Fraction White Residents	0.64***	0.69***	0.65***	0.70***	0.72***
	(0.05)	(0.05)	(0.05)	(0.07)	(0.09)
Log(Income)	0.44***	0.38***	0.36***	0.33***	0.32***
	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)
N	6688	6688	6688	4360	2435
Boundary FE	×	×	×	×	×
Distance Polynomial	×	×	×	×	×
Housing Controls		×	×	×	×
Amenity Controls			×	×	×
Distance Bandwidth	2mi	2mi	2mi	1mi	0.5mi
r2	0.02	0.73	0.74	0.77	0.83

Notes: Standard errors reported and are clustered at the boundary level. All specifications include boundary fixed effects and control for spatial variation in the house price gradient as a quadratic function of block group distance to the boundary. Housing controls include the median age of housing and the number of rooms as a proxy for housing size. Amenity controls include two proxies for school quality and two measures of access to food and retail consumption amenities.

Table 5: Hedonic Estimates of Neighborhood Amenities : 1 Mile

	(1) Big City	(2) Big Suburb	(3) Small City	(4) Small Suburb	(5) Small Town
High Amenity Side	0.03	0.04	-0.01	0.06	-0.01
	(0.02)	(0.03)	(0.03)	(0.06)	(0.03)
Fraction White Residents	1.04***	0.20	0.55***	0.52*	0.96***
	(0.11)	(0.10)	(0.11)	(0.24)	(0.13)
Log(Income)	0.29***	0.34***	0.33***	0.18*	0.23***
	(0.04)	(0.07)	(0.06)	(0.09)	(0.06)
N	1785	849	761	172	791
FE + Full Controls	×	×	×	×	×
r2	0.86	0.86	0.80	0.84	0.84

Notes: Standard errors reported and are clustered at the boundary level. All specifications include boundary fixed effects and control for spatial variation in the house price gradient as a quadratic function of block group distance to the boundary. The full set of controls are described in Table 4.

Table 6: Hedonic Estimates of Neighborhood Amenities: 0.5 Mile

	(1) Big City	(2) Big Suburb	(3) Small City	(4) Small Suburb	(5) Small Town
High Amenity Side	0.02	-0.01	0.02	0.00	-0.01
	(0.03)	(0.05)	(0.04)	(0.06)	(0.04)
Fraction White Residents	0.88***	0.26	0.72***	0.31	0.93***
	(0.17)	(0.15)	(0.17)	(0.36)	(0.16)
Log(Income)	0.22***	0.35***	0.42***	0.06	0.26**
	(0.05)	(0.10)	(0.08)	(0.16)	(0.09)
N	991	414	421	94	512
FE + Full Controls	$\times$	$\times$	$\times$	$\times$	$\times$
r2	0.89	0.88	0.83	0.92	0.88

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Standard errors reported and are clustered at the boundary level. All specifications include boundary fixed effects and control for spatial variation in the house price gradient as a quadratic function of block group distance to the boundary. The full set of controls are described in Table 4.

Table 7: Willingess to Pay : 1 Mile

	(1) Big City	(2) Big Suburb	(3) Small City	(4) Small Suburb	(5) Small Town
Income					
Hedonic Estimate	0.29***	0.34***	0.33***	0.18*	0.23***
	(0.04)	(0.07)	(0.06)	(0.09)	(0.06)
Boundary Differential	0.27***	0.19***	0.24***	0.20***	0.28***
	(0.03)	(0.03)	(0.03)	(0.05)	(0.02)
Multiple (I)	0.078	0.065	0.079	0.036	0.064
White Composition					
Hedonic Estimate	1.03***	0.21*	0.55***	0.52*	0.97***
	(0.12)	(0.10)	(0.11)	(0.24)	(0.13)
Boundary Differential	0.08***	0.06***	0.04**	0.03	0.08***
	(0.01)	(0.01)	(0.02)	(0.04)	(0.01)
Multiple (W)	0.0832	0.012	0.022	0.0156	0.0768
Mean Home Value (P)	210,525	183,179	118,026	121,817	107,225
WTP (P×I)	16,484	11,833	9,348	4,385	6,905
WTP (P×W)	17,516	2,198	2,597	1,900	8,235
Total WTP	34,000	14,032	11,944	6,286	15,140
Percent of Home Value	16.2%	7.7%	10.1%	5.2%	14.1%

Notes: Standard errors reported and are clustered at the boundary level. In both the upper and middle panel, hedonic estimates are the coefficients presented in Table 4, and boundary differentials are the point estimates shown in Figures 12 and 13. The multiple is the product of the two estimates, and is used in the lower pane to convert the empirical estimates into dollars of home prices.

Table 8: Willingess to Pay : 0.5 Mile

	(1) Big City	(2) Big Suburb	(3) Small City	(4) Small Suburb	(5) Small Town
Income					
Hedonic Estimate	0.23***	0.35***	0.42***	0.06	0.26**
	(0.05)	(0.10)	(0.08)	(0.16)	(0.09)
Boundary Differential	0.24***	0.14***	0.22***	0.16*	0.25***
	(0.03)	(0.04)	(0.04)	(0.07)	(0.03)
Multiple (I)	0.055	0.049	0.092	0.009	0.065
White Composition					
Hedonic Estimate	0.87***	0.26	0.72***	0.31	0.93***
	(0.17)	(0.15)	(0.17)	(0.36)	(0.16)
Boundary Differential	0.06***	0.03	0.03	-0.00	0.06**
	(0.01)	(0.02)	(0.02)	(0.04)	(0.02)
Multiple (W)	0.052	0.008	0.022	0.0	0.056
Mean Home Value (P)	208,203	175,275	110,323	110,330	98,496
WTP (P×I)	11,493	8,588	10,194	1,059	6,402
WTP (P×W)	10,868	1,367	2,383	0	5,496
Total WTP	22,361	9,956	12,577	1,059	11,898
Percent of Home Value	10.7%	5.7%	11.4%	1.0%	12.1%

Notes: Standard errors reported and are clustered at the boundary level. In both the upper and middle panel, hedonic estimates are the coefficients presented in Table 5, and boundary differentials are the point estimates shown in Figures 12 and 13. The multiple is the product of the two estimates, and is used in the lower pane to convert the empirical estimates into dollars of home prices.

Table 9: Placebo Estimates of Price Differential

	(1) Log(Price)	(2) Log(Income)	(3) Fraction White	(4) Fraction Latino	(5) Fraction Black
High Income Side	0.03	0.05	-0.01	0.00	-0.00
	(0.07)	(0.06)	(0.02)	(0.03)	(0.02)
r2	0.61	0.61	0.78	0.85	0.45

Notes: Standard errors reported and are clustered at the boundary level. Placebo estimates are obtained by estimating the main model after arranging block groups on opposite sides of fictional boundaries. Column one takes prices as an outcome, and column 2-5 are estimates of demographic differentials at fictional boundaries.