Local Segregation of Neighborhoods and Amenities

Kenneth Whaley*

March 16, 2023

Abstract

This paper sheds light on historic railroad placement as a predictor of contemporary segregation. Employing a digitized map of Texas railroads circa 1911 to compare census block groups separated by tracks in 2018, I first document discrete changes in house prices, income, and racial composition at the railroad boundary. I then use spatial difference-in-differences to estimate an unconditional house price premium of 21% to live on the high amenity side of the tracks. Hedonic estimates of the model predict the house price premium is more likely explained by differences in income, racial demographics and test scores; and less likely driven by differences in private consumption amenities such as restaurants and bars. To mitigate the effects of unobserved neighborhood quality attributes, I estimate the model on samples progressively close to the railroad boundary on either side. In doing so I find new evidence that neighborhood racial demographics are a stronger predictor of the house price premium than income, school quality, and access to private consumption amenities.

JEL Classification: R23, J15, O18

*Department of Economics, University of South Florida.

1 Introduction

The United States has a long, persistent history of residential segregation that continues to draw the attention of news outlets and public forums.[1] Economists and sociologists have put forth models of segregated housing markets as an equilibrium outcome within cities, resulting in the stratification of households by race and income (Schelling 1971, Cutler et al. 1999, Logan and Parman 2017). A central claim is that decentralized segregation (or decentralized sorting) occurs in the absence of laws prohibiting non-white individuals from predominantly white spaces. However, an abundance of such laws created the racial segregation that existed under the Jim Crow Era doctrine of *separate but equal*. One such case is *Plessy v. Ferguson* in 1896, in which the US Supreme Court upheld a Louisiana law requiring separate railcar accomodations for Black and White passengers (Kelley 2010). The decade to follow would represent the industry peak for railroad traffic (Gallamore and Meyer 2014).

Railroads are arguably more infamous for segregating residential neighborhoods, as opposed to the physical separation of riders in a given passenger car. This phenomenon stands today even as railroad utilization has been in decline for more than half of a century (Gallamore and Meyer 2014). In this paper I use the historic placement of railroad tracks to analyze contemporary neighborhood segregation at a micro level. The physical durability of train tracks is remarkable, and I test whether railroads laid in the early 1900s create distinct neighborhood boundaries today. Variation in 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). This paper contributes to the literature by credibly identifying the effects of this relationship through a study of neighborhoods on opposite sides of the track.

I start with a digitized version of New Century Atlas Map that provides the location of all rail lines in Texas as of 1911 (Atack 2013). I enrich the railroad map by overlaying a map

[1] Feature stories have been written in The Economist, Wall Street Journal, Washington Post, the New York Times to name a few.

of Texas census block groups provided by the US Census. Block groups are relatively small land areas with a median population of 1,437 residents; and the Census provides geographic coordinates of the block group population center for additional spatial granularity. The coordinates are used to map block groups to a side of the nearest railroad tracks, creating a cross section of neighborhoods augmented with Census data for house prices, income, racial demographics, and housing quality at the block group level. I categorize each side based on which has the higher median income, from here referred to as the high amenity side.

After categorizing the neighborhoods, a boundary discontinuity model is estimated by regressing outcome variables describing neighborhood characteristics on a series of indicators that arrange block groups into bins based on distance to the tracks. The boundary discontinuity design is a method popularized in the urban/neighborhood choice literature to study house price capitalization of public good quality based and proximity of a single-family home to an administrative boundary (Black 1999, Bayer et al. 2007, Keele and Titiunik 2015). My econometric strategy relies on the simple assumption that pure disamenities of living near a railroad (noise and smoke pollution, congestion, outages) have symmetric effects on the residents of each side and diminish with distance from the track. Under this assumption I posit that absent segregation there should be no discrete changes in observable neighborhood attributes at the railroad boundary, and house prices should increase symmetrically with distance from either side of the boundary. With data describing 8,993 block group neighborhoods near Texas railroads, I am able to test if households and amenities are organized in a way that reflects the nature of residential segregation.

From the discontinuity regressions I plot the coefficients of each distance bin and show sharp increases in house prices and neighborhood incomes at the railroad boundary. Further, both house prices and incomes increase asymmetrically with distance from the tracks. The magnitude of the price and income effects are similar, and when taken together under my hypothesis display clear patterns of local segregation. Moving beyond income and prices, I carry out this procedure for neighborhood housing quality variables, racial demographics, school spending/test scores, and counts of food and retail establishments near each block group population center. The results trace out unique effects of segregation on neighborhood demographics and amenities near railroad boundaries.

Next, I estimate the difference in mean house prices for each distance bin on the high amenity side, relative to equidistant neighborhoods on the opposite side of the same train tracks. To do so I implement a spatial difference in differences model with boundary fixed effects. The unconditional or gross price premium is obtained from a version of this model with no additional neighborhood controls and is estimated to be near 21% ($40,290) for a home on the high amenity side. Including the robust set of neighborhood covariates in the model I estimate the net or conditional house price premium to be near 5% ($9,600) ceteris parabus.

To interpret the extent to which each neighborhood characteristic explains the price premium, the endogeneity concern is that unobserved neighborhood quality attributes are correlated with observed characteristics and housing prices. I argue that railroad segregation creates distinct variation in neighborhood attributes, allowing me to credibly recover hedonic estimates for the effect of each characteristic on the neighborhood housing prices. The average age of housing decreases symmetrically from the boundary while the average size of housing jumps discretely at the boundary and increases with distance on the high amenity side. The variation in racial demographics appear orthogonal to the housing quality measures, and variation in public amenities (proxied by school quality) can be distinguished from variation in private amenities (proxied by counts of restaraunts and retail shops).

In reporting results of my hedonic estimation strategy I follow a standard approach in the urban literature that involves comparing coefficients of interest across specifications as more covariates are added and the spatial area of analysis is reduced.[2] The goal is to uncover the nature of bias created by unobserved neighborhood attributes with the idea that the potential for bias increases across space. I document consistent evidence that income

[2]The tradeoff for precise estimates in a small geographic area is a decrease in the sample size.

effects on local house prices shrink as the model becomes more restrictive and the effects of neighborhood racial composition grow stronger. As compared to the estimates from the least restrictive model specification, income effects decrease from $\beta = 0.93$ to $\beta = 0.31$, while the coefficient on the share of white residents increases from $\beta = 0.44$ to $\beta = 0.72$. The preferred model predicts that house prices decrease in the age of housing, increase in housing size, and positively capitalize incremental increases in elementary school test scores. I find no effect of private consumption amenities in a full model that conditions on income and school quality.

A clear takeaway from the exercise of this paper is the role train tracks play in enforcing segregationist boundaries on neighborhoods. The precision and magnitude of the results paint a consistent picture of how neighborhoods change when traversing from one side of a railroad track to another. Further, the effect is not contained to one specific margin of interest. There are identifiable changes in the physical quality of housing, the income and racial composition of residents, and the quality of schools. Less identifiable are changes in private consumption amenities, though this paper only attempts to detect quantity and not quality changes in food and retail shopping. The differences add up to generate a 21% observable price premium for living on the good side of the tracks.

2 Railroad Placement and Segregation

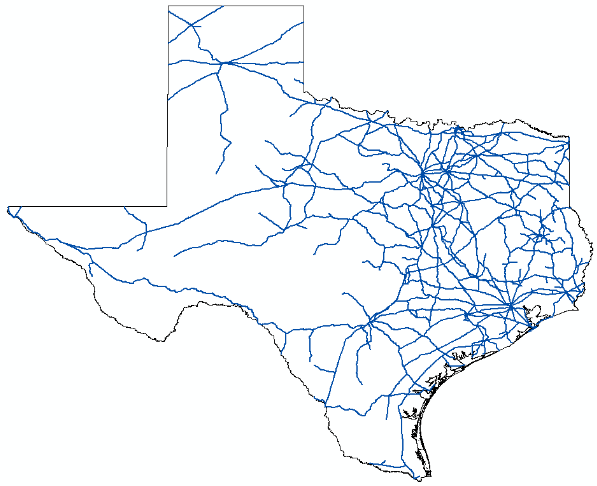
It is important to characterize the cultural landscape when railroad tracks were laid, and what about US society connects contemporary neighborhoods to an era over 100+ years ago. Much of the rail lines in existence today were installed prior to 1920. At its height circa the turn of the 20th century, the bustling industry moved freight and passengers across the territorial US at speeds not obtained by other modes of transportation. While it is likely that railroads were initially laid to foster economic growth and facilitate transit, de jure segregation prevailed and open acts of violence were carried out against Black residents (Cook et al.

2018, Williams 2022). The hostile environment for non-whites is further evidenced by the way Black riders were treated on passenger cars. During the Jim Crow era laws in Texas and other southern states dictated where Black passengers could sit and demanded deference to White passengers (Roback 1986, Osborn 2002). Thus, the racial animus towards non-whites was a pre-existing condition only exacerbated by the growing presence of railroads in the periods following reconstruction. It follows that household-level decisions are more responsible for the increase in segregation from 1880-1940 (Logan and Parman 2017). Combined with the racial stratification of wealth, property values and income, Blacks experienced restricted access to high quality public and private amenities (Cook et al. (2022)).

Although utilization has steadily declined there have been little effort to remove substantial portions of track. The more likely scenario is abandonment or discontinued maintenance, but the physical presence remains in place. Track durability is a subtle but pivotal feature that translates historical segregation into decentralized sorting we observe today. Since neighborhood boundaries created by railroads remained in place for over a century, segregation is less likely to dissipate through periods of distinct macro-migration waves like The Great Migration, white flight, Hispanic immigration, and the more recent urban revival in American cities (Derenoncourt 2022, Boustan 2010, Boustan and Margo 2013, Caetano and Maheshri 2023, Couture and Handbury 2020). So long as in-migrating individuals hold preferences for living near same race neighbors, port-of-entry theory suggests preexisting segregation will be slow to change. In addition it is posited that white households will pay more to segregate themselves amongst neighborhoods with concentrated white populations (Cutler et al. 1999).

Figure 1 shows the far-reaching extent of rail lines in Texas as of 1911. The history of the railroad industry in Texas is centered in a complicated web of land use and petroleum politics, peak rail traffic in the state occured only slightly later than the majority of US States (DeGolyer 1970). Jim Crow policies also dominated the passenger train arrangements like other southern US states (Osborn 2002). What makes Texas unique is the vast landscape comprised of urban, suburban and rural areas all affected by track placement.

Figure 1: Railroads in Texas : 1911



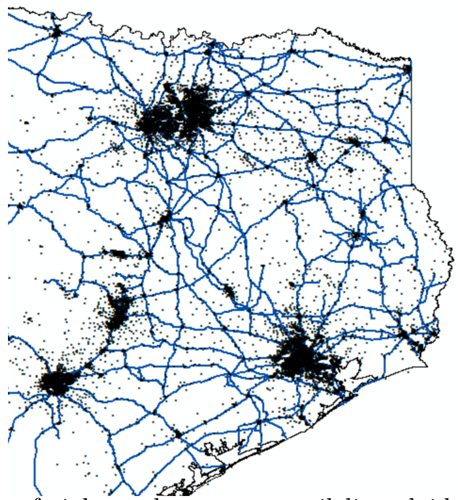
Notes: The blue lines represent freight and passenger rail lines laid in Texas as of the early 1900s. Source: 1911 New Century Map digitized by Jeremy Atack, Department of Economics, Vanderbilt University.

3 Data

The late 1920s are considered the peak era of railroad transit in Texas. 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 create the a shapefile of historic railroad segments using ArcGIS software. Each two-mile segment of track is considered a boundary, and the estimation sample includes only boundaries with neighborhoods on geographically opposite sides of the boundary.

Neighborhoods in the sample are defined by census block groups, small clusters of population with contemporary characteristics defined in 2018. The median block group population is 1,480 and the sample covers roughly 62% of Texas population (17.8M). The census provides coordinates for the block group center of population, an estimate of the central mass of residents living within block group boundaries. This spatial information is advantage over block group centroids, and I use the coordinates to map neighborhood locations to the nearest two-mile railroad segment.

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.

Neighborhood descriptors included in the census data include median home values, median age of housing, and the median number of rooms as a proxy for home sizes. Residents are described by median income, racial demographic shares and age profiles. The estimation sample includes 8,993 neighborhoods mapped to 1,215 railroad segments for an average of 7.4 neighborhoods per boundary. Neighborhood summary stats are located in table 1.

Median income for both sides of each boundary segments are computed and neighborhoods are assigned to the *high-amenity* or *low-amenity* side according to which side has the higher median income. Table 1 shows that house prices and incomes are 11-12% higher in high-amenity neighborhoods relative to the full sample. Demographically, high amenity neighborhoods have a higher share of white residents and less minority residents.

Public school data is 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 disticts. This is fruitful as my research design requires high frequency variation in small spacial geographic areas. The median elementary school spends $ 7,461 per-pupil in 2018, with little observational differences for schools in high-amenity areas. The last piece of data come 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.

Table 1: Block Group Summary Statistics, 2018

	All Block Groups	Full Sample	Low Amenity Side	High Amenity Side
House Prices	184,274.27	191,877.45	168,296.40	214,732.42
	(144657.8)	(159119.0)	(122305.4)	(185213.5)
Median Income	69,820.91	70,741.09	62,476.90	78,750.81
	(36752.1)	(38411.3)	(32095.8)	(42153.9)
Fraction White	0.43	0.40	0.36	0.45
	(0.289)	(0.284)	(0.276)	(0.285)
Fraction Hispanic	0.39	0.40	0.43	0.37
	(0.295)	(0.299)	(0.296)	(0.300)
Fraction Black	0.11	0.12	0.14	0.10
	(0.157)	(0.167)	(0.182)	(0.150)
Spending Per-Pupil	7,354.34	7,459.78	7,583.08	7,340.29
	(1939.4)	(1934.9)	(2041.1)	(1818.4)
Mastery	0.28	0.28	0.26	0.30
	(0.125)	(0.128)	(0.115)	(0.137)
Median Rooms	5.71	5.71	5.49	5.93
	(1.272)	(1.344)	(1.209)	(1.430)
Housing Age	45.06	46.52	45.15	47.85
	(158.4)	(162.2)	(150.2)	(173.1)
Food Venues/Sq. Mi	7.38	8.67	8.92	8.42
	(9.909)	(10.44)	(11.02)	(9.844)
Other Retail/Sq. Mi	25.83	30.36	30.67	30.06
	(25.95)	(26.96)	(27.43)	(26.50)
Observations	13872	8996	4388	4608

4 Discontinuities at Railroad Boundaries

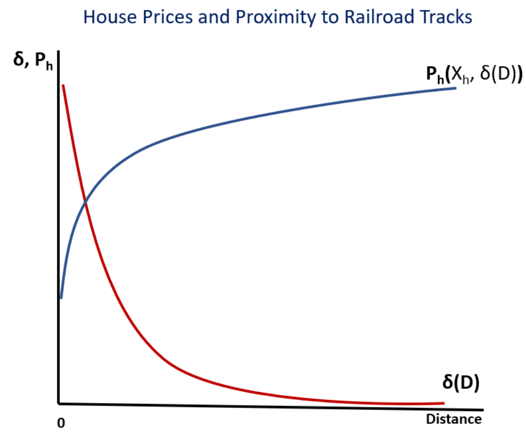
A useful starting point in this analysis is to consider what neighborhoods on either side of railroads would look like if demographic and amenity segregation was unrelated to the presence of railroad boundaries. The pure disamenity of living near a railroad are the spillover costs imposed on residents on either side of the tracks. Malfunctions, outtages, congestion along with noise and air pollution are less than ideal but the effects decrease in distance from the tracks.

Absent segregation house prices would be lowest close to the tracks and rise with distance as the pure railroad disamenities wane. Figure 3 shows the theoretical house price gradient increasing in distance from the tracks as a function of pure disaminities δ and neighborhood characteristics $\mathbf{X}$. Absent segregation $\mathbf{X}$ is randomly assigned, and the spatial distribution of amenities is not predicted by side of track. In that case, one could mirror the price gradient across the Y-axis of figure 3 and the slope on both sides should be symmetric. The working assumption is that δ is perfectly symmetric across the Y-axis, and how δ decreases in distance is indifferent to each side. It follows that deviations from the symmetry condition for prices can be viewed as changes to the house price gradient caused entirely by differences in neighborhood amenities $\mathbf{X}$. Recall that neighborhood amenities would be indistinguishable in the absence of segregation.

The refined spatial nature of the data allow me to estimate the house price gradient shown in Figure 3. To do so I exploit the arrangement of census block group population centers on either side of railroad boundaries and the distance of each block group from the track. The goal is to identify discrete changes in prices and amenities at the boundary along with patterns in observable neighborhood traits moving farther from the tracks.

I first document the observable changes in neighborhood composition near and across the tracks by arranging the block groups into bins based on distance. Each bin contains block groups within half-mile bands moving away from the railroad in either direction, exploiting the refined location data of the population centers to get an accurate projection of neigh-

Figure 3: Price and Disamenity Curves



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

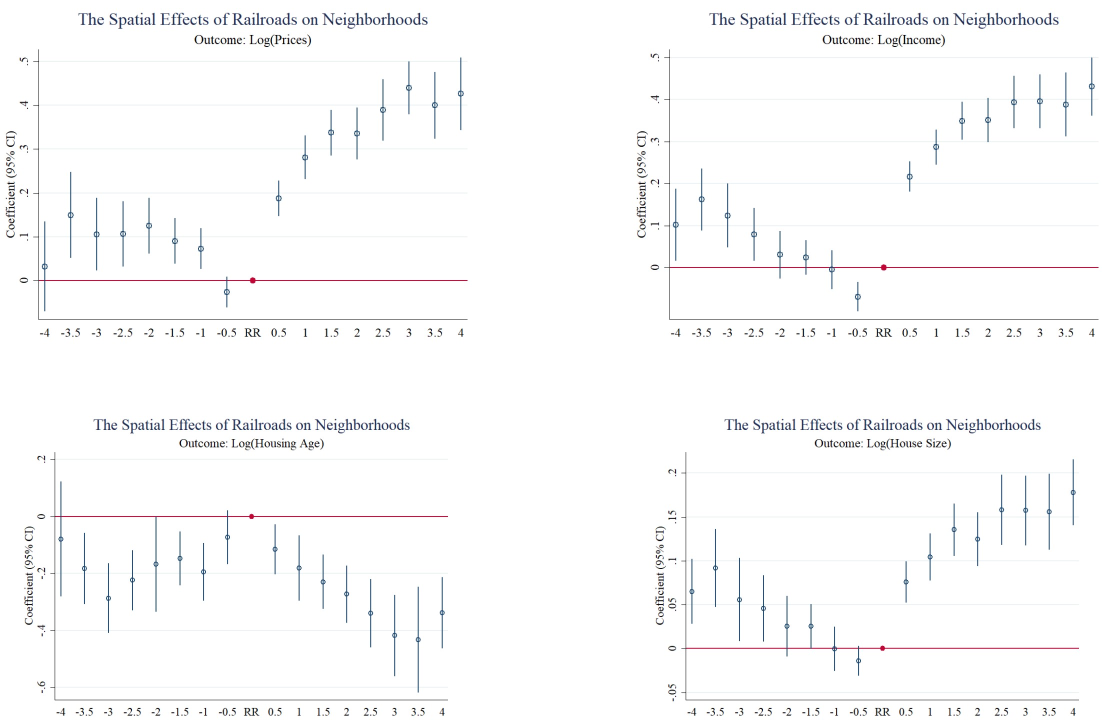
borhoods where the majority of residents live. Taking every two-mile railroad segment as a boundary, I estimate the model

$$Y_{ijk} = \alpha_0 + \sum_{j \neq 0}^{J} \alpha_j \mathbb{I}_j + \theta_k + \epsilon_{ijk} \tag{1}$$

for each outcome of interest describing neighborhood **i**, in distance bin **j**, near railroad boundary **k**. Bins on the low-amenity side are given negative values, and neighborhoods within 0.5 miles from the boundary are in bin 0 (the omitted category). By including boundary fixed effects θ_k, the estimates $\hat{\alpha}_j$ represent differences in each outcome across space, relative to the mean for neighborhoods in bin 0 at the same boundary. Using this procedure I put forth descriptive evidence of discrete changes in neighborhood attributes crossing from the low to high amenity side.

Absent segregation the slope of the house price gradient away from the track should be symmetric. To test this hypothesis I plot the $\hat{\alpha}_j$ coefficients arranged by distance bin, with control group neighborhoods being those closest to the tracks. The first two outcomes of interest are the natural log of house prices and natural log of neighborhood income, as shown in the upper half of figure 4 where two striking patterns emerge. First there is a discrete jump in both prices and incomes when reading the graph from left to right. Second, both house prices and income rise when moving away from the railroads on either side. The bottom half of figure 4 contains estimates for housing quality measures such as median housing age and the median number of rooms as a proxy for home size. On average homes nearest to the tracks are older, indicative of lower real-estate investment and construction activity. House size increases in distance from the boundary on both sides, but shows a similar discontinuity as the price and income estimates.

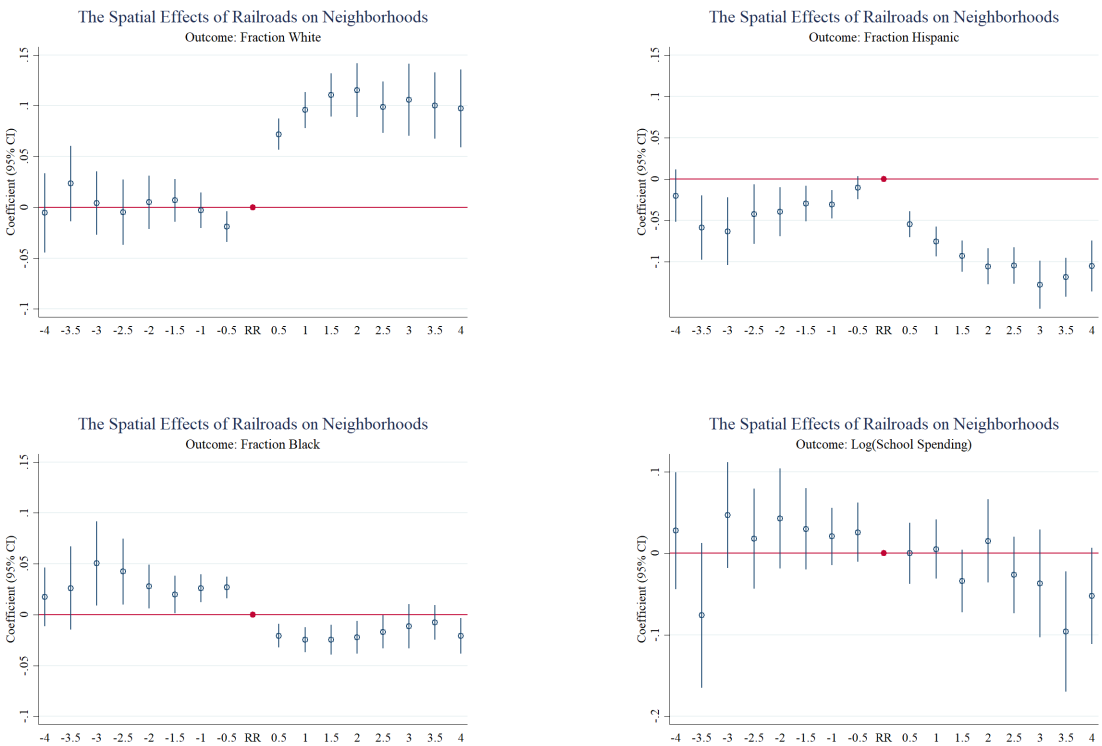
Figure 4: Discontinuities in Prices, Incomes, and Housing Quality



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 **Y** in each bin, relative to control group neighborhoods at the same railroad segment.

The next set of neighborhood outcomes I explore are the racial characteristics of residents. In figure 5 I plot the distance-bin estimates for regressions taking the population share of each race or ethnic group as an outcome. Figure 5a shows the fraction of white households jumps when crossing to the high amenity side, while figure 5b and figure 5c show respective decreases in the fraction of both Hispanic and Black residents. The results of shown for all race groups appear symmetric in distance to the railroad.

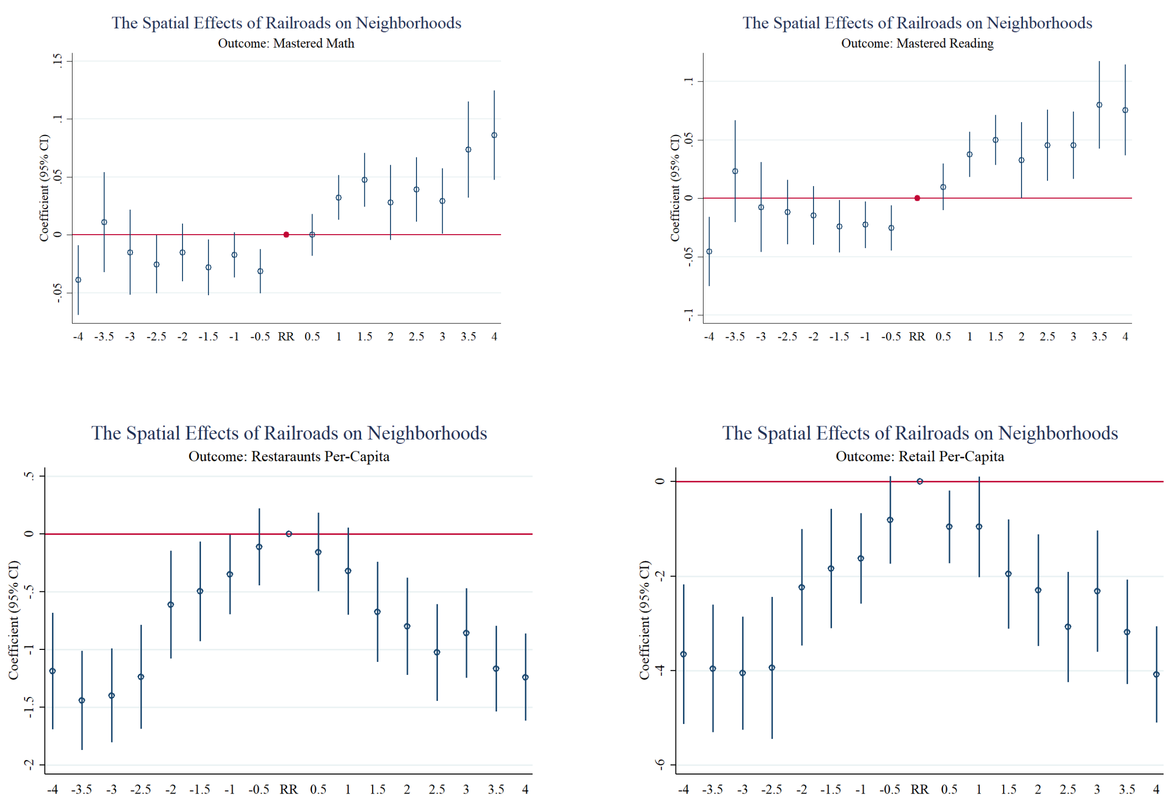
Figure 5: Discontinuities in Neighborhood Race and Ethnic Composition



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.

The analysis continues with results detailing the nature of public amenities (ie. schooling) and private consumption amenities (food venues and retail shops). In figure 5d I document weak evidence for higher school spending on the low amenity side of the tracks, but the estimates are imprecise and not statistically significant. In the upper panel of figure 6 are estimates that show math and reading scores that increase smoothly across the railroad boundary to the high amenity side. In the bottom panel of figure 6, the model estimates the density of both food and non-food retail establishments to be highest closest to railroad boundaries. On the surface the patterns reflect Hotelling spatial competition for sellers of homogenous consumption goods, where firms clustering at near the track boundary are accessible to households on both sides (Graitson 1982, Brown 1993).

Figure 6: Discontinuities in Public and Private Amenities



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.

5 Estimation

Upon observing remarkable traits of neighborhoods segregated by railroad tracks, I turn to estimating the magnitude difference of house prices on the high amenity side of the tracks.

Given the prevalence of white household sorting into neighborhoods on the high amenity side, the house price premium can be interpreted as the willingness to pay for proximity to white neighbors and high income amenities. To put bounds on the price premium I first estimate the gross price premium, defined as the difference in mean house prices relative to other neighborhoods at the same railroad boundary. I then estimate the net price premium, defined as the difference in conditional mean house prices relative to other neighborhoods at the same railroad boundary, controlling for observable differences in neighborhood characteristics. The approach of this section is to specify a spacial difference-in-differences model that can flexibly produce estimates of the house price premium and hedonic estimates for the house price capitalization of local neighborhood attributes.

5.1 The House Price Premium

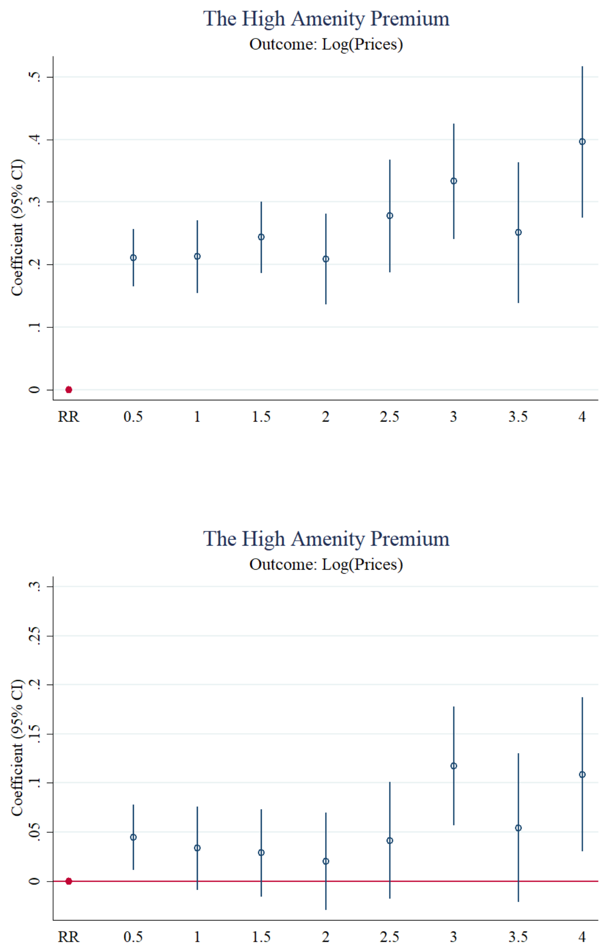
The spatial difference-in-difference specification takes the form

$$Log(P_{ijk}) = \sum_{j \neq 0}^{J} \pi_j \mathbb{I}_j \times \mathbb{H}_j + X_{ijk}^{'}\beta + \theta_j + \theta_k + \epsilon_{ijk}. \tag{2}$$

P_{ijk} is median house price in census block group **i** in distance bin **j** at railroad segment **k**. The interpretation of the house price premium is as follows. Including both bin θ_j and boundary θ_k fixed effects allows me to compare mean neighborhood house prices at similar distances from the boundary but opposite sides of the track. The model is flexible and can identify both discrete changes at the boundary and heterogeneity in slopes as distance from the railroads increase on each side. To that end I interact each distance-bin dummy with an indicator for the high amenity side of the railroad segment. The $\hat{\pi}_j$ coefficients in equation 2 are estimates of the price premium for being on the high amenity side, interpreted as the difference in the conditional mean house price relative to the inner bin (j=0) and the distance bin pair on the low amenity side.

The gross or unconditional house price premium can be shown by plotting the $\hat{\pi}_j$ coefficients from a version of equation 2 that omits the vector of neighborhood covariates X_{ijk}. In the top panel of figure 7 I show estimates of a 21% gross premium at the boundary, that increases steadily for three miles on the high income side. An easy way to think of the gross price estimate is through differencing the estimates in each distance bin pair of figure 4a.

Figure 7: Difference-In-Difference Estimates



Notes: Each plot includes the $\hat{\pi}_j$ coefficients from the hedonic regression in equation 2. 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, as marked by a red dot. By including boundary and distance bin fixed effects the model produces difference-in-difference estimates for the house price premium on the high-amenity side, relative to the control group neighborhoods and neighborhoods in the same bin on the low-amenity side. The estimates in the bottom figure include a rich set of neighborhood amenity controls that explain the gross premium showed in the upper figure.

The net price premium is found by again plotting the $\hat{\pi}_j$ coefficients, this time from a specification of equation 2 that includes neighborhood characteristics $\mathbf{X}$. In the lower panel of figure 7 I document a net house price premium of 5%, most precisely estimated at the boundary and increasing in distance from the tracks.

5.2 The House Price Capitalization of Amenities

To analyze which neighborhood features drive the price premium I estimate versions of equation 2 that vary by which covariates are included and the comparison group as determined by spatial distance. I first provide the OLS estimates of $\hat{\beta}$, the house price capitalization of each X_{ijk} in table 2. Moving across specifications to the right, I add more controls and continually shrink down the spatial area of comparison to deal with endogeneity of unobserved neighborhoods characteristics correlated with $\mathbf{P}$ and $\mathbf{X}$. In principle an OLS model with a rich set of controls should suffer from minimal omitted variables bias in a small enough space. The trade-off is a material reduction in sample size.

The first model with minimal neighborhood controls and spatial restrictions shows a near unit elastic relationship between income and house prices, $\beta = 0.93$, implying that a 1% increase in income increases prices by 0.93%. House prices are also increasing in the share of white residents in the neighborhood with a coefficient of 0.44. Taken as is this implies that a 0.1 increase in the fraction of white residents increases house prices by 0.044 or roughly 4.4%. In terms of physical housing quality, the effect of housing age is small, but there is a strong negative coefficient on house size. With no spatial constraints on comparison groups this represents house prices being higher in the urban core where lot sizes are smaller.

In column two I add test scores to the model and find positive capitalization of this school quality measure. Here test scores are measured as the fraction of students that achieved a *mastery* rating for either reading, math or both. Considering the fraction of students at the mean school to master the material is close to 0.28, an material increase of 0.1 would raise house prices by 8.6%. Converslely I find very little evidence for the house price capitalization of food and retail shopping amenities. In column three I add boundary fixed effects to the model to make an apples to apples comparison of neighborhoods in the same geographic area.

In column three I include boundary fixed effects, which is akin to comparing variation in the **X's** across all neighborhoods on either side of the railroad segment. When comparing only neighborhoods across a tighter space, the effect of income and test scores test scores diminishes while the effect of racial composition grows stronger. The argument for validity of including boundary fixed effects is through the estimate for house size. When comparing neighborhoods at the same spatial area, those with large houses have higher house prices. Over a larger spatial area (columns 1 and 2) house sizes are more reflective of proximity to the urban core instead of size as a measure of quality (columns 3-5) .

In columns four and five I add even stronger restrictions on the sample based on distance from the boundary. The first restriction is to neighborhoods $\leq$ 1.5 miles from the boundary and the second shrinks the distance band down to 0.5 miles on either side. The assumption is that smaller distance bands diminish the possibility for omitted variable bias as comparison neighborhoods are more likely to be more similar in unobserved amenities. Ideally the tradeoff for reduced sample size are more precise estimates for the relationship of interest. As I analyze smaller and smaller spatial areas in columns four and five of table 2, the house price capitalization of income and school quality becomes smaller while the capitalization of race becomes stronger.

Table 2: OLS Estimates of Neighborhood Amenities

Outcome: Log(House Prices)	(1)	(2)	(3)	(4)	(5)
Log(Income)	0.93***	0.72***	0.41***	0.35***	0.31***
	(0.03)	(0.03)	(0.02)	(0.02)	(0.04)
Fraction White Residents	0.44***	0.43***	0.63***	0.69***	0.77***
	(0.03)	(0.03)	(0.04)	(0.05)	(0.09)
Log(Housing Age)	-0.03*	-0.08***	-0.11***	-0.06***	-0.04
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Log(Housing Size)	-0.43***	-0.14*	0.12*	0.20**	0.23*
	(0.06)	(0.06)	(0.05)	(0.07)	(0.11)
Test Scores		0.86***	0.60***	0.60***	0.51***
		(0.07)	(0.06)	(0.09)	(0.12)
Food/Sq. Mi		0.01**	0.00	0.00	0.01*
		(0.00)	(0.00)	(0.00)	(0.00)
Retail/Sq.Mi		0.00***	0.00	-0.00	-0.00
		(0.00)	(0.00)	(0.00)	(0.00)
Boundary FE			×	×	×
N	8991	8991	8991	5524	1848
r2	0.61	0.67	0.82	0.86	0.86

Standard errors in parentheses are clustered at the boundary level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Next I report the estimates from the preferred specification of equation 2 that includes both boundary and bin fixed effects. Moving from column 1 to 3 in table 3 again involves the implementation of tighter and tighter spatial restricitions on the sample. I interpret these estimates as explaining the gross-price premium for each bin in figure 7. Recall that including the full set of neighborhood covariates in the regression is the difference between the gross (21%) and net (5%) price premium of being on the high amenity side. Overall the magnitudes of the estimates largely are largley similar to the OLS results in column 5 of table 2 with more precise estimates of housing quality covariates house age and size.

To put the magnitude of the estimates into context consider descriptive means of income, fraction of white households and test scores between the low and high amenity side. The log-log coefficient of $\beta = 0.31$ implies a 10% income increase prices by 3.1%. The 26%

mean difference in incomes between the high and low amenity side increases house prices roughly 8.1%. The fraction of white residents increases by roughly 0.1 on the high amenity side, predicted by the model to increase house prices by 7.2%. The fraction of students that master either the reading or math portion of the elementary school aptitude test increases by 0.04 on the high amenity side, which would increase house prices by a small but statistically significant 2.24%.

Table 3: Hedonic Estimates of Neighborhood Amenities

Outcome: Log(House Prices)	(1)	(2)	(3)
Log(Income)	0.39***	0.33***	0.31***
	(0.02)	(0.02)	(0.03)
Fraction White Residents	0.60***	0.66***	0.72***
	(0.04)	(0.05)	(0.06)
Log(Housing Age)	-0.10***	-0.06***	-0.06***
	(0.02)	(0.02)	(0.02)
Log(Housing Size)	0.12*	0.20**	0.24**
	(0.05)	(0.07)	(0.09)
Test Scores	0.57***	0.58***	0.56***
	(0.06)	(0.08)	(0.08)
Food/Sq. Mi	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)
Retail/Sq.Mi	0.00	-0.00	-0.00
	(0.00)	(0.00)	(0.00)
Boundary Fixed Effects	×	×	×
Bin Fixed Effects	×	×	×
Sample	All	$\leq 3mi$	$\leq 1mi$
N	8991	5524	3993
r2	0.82	0.83	0.84

Standard errors in parentheses are clustered at the boundary level.

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

6 Conclusion

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. Further analysis is required to understand the mechanisms in which drive the main results, and to rule out potential unobserved confounders. The main concern would be that administrative boundaries like school districts, county lines, and congressional districts may coincide with railroad tracks. This would suggest that variation in house prices at the boundary is driven by differences in tax rates or local public goods across administrative boundaries. In addition, right of way clearance created by historic railroad tracks could be the precursor for highway development, which if unobserved may indicate that segregation is not drive solely by railroad boundaries but a combination of railroads and highways.

Discrimination in real estate markets also may bias my results. When banks participated in redlining, maps were drawn with boundaries around neighborhoods set by the racial composition of residents. The neighborhoods were then assigned a grade based on bank-determined attractiveness of lending to potential homebuyers there. If redlining coincided with railroad boundaries, my results could be capturing the legacy of real estate discrimination instead. That does not, however, invalidate this study because the implication is that track position caused the redlined maps to be drawn a certain way which is indeed part of the legacy of railroad induced segregation.

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. I estimate the sticker price premium for a house on the high amenity side of the tracks to be near 21%. This premium is explained by the desire to live near higher income neighbors, in whiter neighborhoods that have better housing quality and schools outcomes. My model predicts the strongest determinants of the house price premium to be the racial demographics of ones neighbors on the high amenity side.

References

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.

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.

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

Boustan, L. P. (2010): “Was postwar suburbanization “white flight”? Evidence from the black migration,” *The Quarterly Journal of Economics*, 125, 417–443.

Boustan, L. P. and R. A. Margo (2013): “A silver lining to white flight? White suburbanization and African–American homeownership, 1940–1980,” *Journal of Urban Economics*, 78, 71–80.

Brown, S. (1993): “Retail location theory: evolution and evaluation,” *International Review of Retail, Distribution and Consumer Research*, 3, 185–229.

Caetano, G. and V. Maheshri (2023): “Explaining Recent Trends in US School Segregation,” *Journal of Labor Economics*, 41, 175–203.

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*.

Cook, L. D., T. D. Logan, and J. M. Parman (2018): “Racial segregation and southern lynching,” *Social Science History*, 42, 635–675.

Couture, V. and J. Handbury (2020): “Urban revival in America,” *Journal of Urban Economics*, 119, 103267.

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.

DeGolyer, E. L. (1970): “Texas Railroads: The End of an Era,” *The Southwestern Historical Quarterly*, 73, 356–380.

Derenoncourt, E. (2022): “Can you move to opportunity? Evidence from the Great Migration,” *American Economic Review*, 112, 369–408.

Gallamore, R. E. and J. R. Meyer (2014): *American railroads: Decline and renaissance in the twentieth century*, Harvard University Press.

Graitson, D. (1982): “Spatial competition a la Hotelling: A selective survey,” *The Journal of Industrial Economics*, 11–25.

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

Kelley, B. M. (2010): *Right to ride: Streetcar boycotts and African American citizenship in the era of Plessy v. Ferguson*, Univ of North Carolina Press.

Logan, T. D. and J. M. Parman (2017): "The national rise in residential segregation," *The Journal of Economic History*, 77, 127–170.

Osborn, W. S. (2002): "Curtains for Jim Crow: Law, race, and the Texas railroads," *Southwestern Historical Quarterly*, 105, 393–427.

Roback, J. (1986): "The political economy of segregation: The case of segregated streetcars," *The Journal of Economic History*, 46, 893–917.

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

Williams, J. (2022): "Historical lynchings and the contemporary voting behavior of blacks," *American Economic Journal: Applied Economics*, 14, 224–53.