The Color of Activity Space: Redlining and the Racial Segregation of Everyday Places

Abstract

Racial segregation is conventionally located where people sleep, yet groups meet or miss one another in the places they pass through by day: the stores, clinics, churches, and restaurants that make up their activity spaces. We argue that this segregation in activity spaces is not merely a shadow of residential sorting but a lasting imprint of exclusionary institutions that organized American space by race, and trace it to a paradigmatic case: 1930s "redlining," in which federal appraisers graded neighborhoods in a color-coded hierarchy of mortgage risk steeped in racial assumptions. Using the inferred racial composition of visitors to four million establishments, we compare locations across historical redlining boundaries with undrawn counterfactual boundaries dividing neighborhoods with similar preexisting differences. Nearly a century later, lower-graded neighborhoods host activity spaces more racially segregated from their metros, and redlining explains about 6% of the standard deviation of place-level segregation at these borders. The effect works mainly through neighborhood racial composition: redlining left neighborhoods more Black and less white than their metros today, and visitors—even those from elsewhere—mirror that skew, suggesting racialized preferences over destination choices. A bureaucratic classification long abandoned still organizes whom Americans encounter each day.

Keywords: redlining; activity spaces; racial inequality; urban mobility; institutional persistence

Introduction

Where people live has long defined how we measure racial segregation (Massey and Denton, 1993), yet the neighborhoods where people sleep are only part of the story: groups also meet or stay apart in the places they move through by day: the grocery store, the clinic, the church, the coffee shop. Over the past decade, researchers using large-scale mobility data have converged on a robust empirical finding: these "activity spaces"[1] are less racially segregated than residential neighborhoods, yet remain substantially sorted, with cross-group exposure in daily destinations falling well short of what metropolitan populations would predict (Candipan et al., 2021; Moro et al., 2021). While commercial and civic places, such as restaurants and public parks, exhibit markedly lower levels of racial segregation, local amenities, such as churches and schools, which have historically been among the most

[1] We measure segregation *in* specific locations, based on the racial composition of their visitors, rather than segregation *of* individuals' activity spaces, following a similar design choice by Baldenius et al. (2025). In the typology of Müürisepp et al. (2022), ours is therefore a place-based rather than person-based measure; we explain this distinction and our rationale below. When reporting our findings, we use "activity space segregation" or "experienced segregation" as shorthand for this place-based measure. When discussing prior studies, however, these terms may refer to person- or flow-based measures, depending on the study's design.

segregated institutions (Gordon, 1964), remain highly segregated by race among their visitors (Athey et al., 2021).

Whether these locations are shared or sorted influences how opportunity and resources are exchanged or hoarded across the color line, carrying important implications on how urban life is *lived* racially. For Black and Latino residents of disinvested neighborhoods, an integrated activity space is a matter of *access*: to the banks, clinics, and stores their own blocks lack, and to the cross-race ties along which economic information and opportunity may flow (Chetty et al., 2022; Small, 2009). For white and affluent residents, it is one of *encounter*: the cross-group contact that residential privilege otherwise forecloses, through which prejudice softens and shared institutions take root (Allport, 1954). For the metropolis, it becomes a question of *cohesion*: whether the daily round of movement dissolves residential segregation or merely recasts it in a more fluid form carries important implications for generalized social trust, cohesion, and prosociality crossing the color line (Rieger et al., 2026).

As mobility data and measures of experienced segregation have become increasingly available to social scientists, a parallel scholarship has increasingly explored *how* this sorting arises, tracing it to a handful of proximate, *present-day* sources. Notably, scholars have linked racial segregation in activity spaces to macro-level features of the urban landscape, such as city density, public transit quality, and the distribution of amenities (e.g., stores, banks, and clinics) that themselves attract different sets of visitors (Athey et al., 2021); to micro-level individual characteristics, such as income, which is strongly associated with race and shapes which destinations people can reach and afford (Davis et al., 2019; Moro et al., 2021); and to meso-level neighborhood racial composition and existing group boundaries, whereby people gravitate toward places they perceive as their own (Fiel, 2021). Each accounts for part of the pattern, yet the account remains, by design, largely *descriptive*: beyond a range of factors associated with experienced segregation, whether contemporary patterns bear the lasting causal imprint of urban institutions and policies—especially *historical* exclusionary practices that profoundly shaped the racial geography from which contemporary American cities emerged—remains largely unclear (for a few exceptions of causal or quasi-experimental studies of experienced segregation, see Aaronson et al., 2025; Baldenius et al., 2025; Heine et al., 2025; Loo et al., 2024).

In this paper, we trace modern-day experienced racial segregation in activity spaces, measured through large-scale mobility flow data in about 4 million establishment, to historical urban exclusionary institutions and study the long-term causal imprinting of one policy in particular: the grading of American neighborhoods by the Home Owners' Loan Corporation (HOLC) in the 1930s. Created during the New Deal to refinance distressed mortgages, HOLC produced "Residential Security" maps that sorted urban neighborhoods in 239 cities between 1935 and 1940 into four grades, from the "best" A (colored green on the maps) to "declining" C (yellow) and "hazardous" D (red). These maps were drawn in consultation with local lenders and realtors and were steeped in the racial assumptions of the era (Hillier, 2005). Banks and lenders accessed maps in which neighborhoods with notable presence of Black or immigrant residents were marked as poor lending risks, or constructed their own maps of geographic lending patterns that were heavily influenced by the HOLC maps (Woods, 2012, for debates among historians about the degree to which lending was influenced by the HOLC map, also see Hillier, 2003). Whether or not the maps themselves directly shaped credit allocation, the appraisal regime they codified channeled mortgage lending toward higher-graded, overwhelmingly white neighborhoods while restricting conventional credit in lower-graded areas where Black and immigrant residents were concentrated, leaving households there dependent on costlier and less secure financing and allowing the resulting disparities in homeownership, property values, and neighborhood investment to compound over the decades that followed (Aaronson et al., 2021b; Faber, 2020; Winling and Michney, 2021).

We extend the burgeoning discussion of the long-term imprinting effect of HOLC grading from outcomes immediately linked to the policy itself—such as home values (Appel, 2016) and residential segregation (Faber, 2020)—to racial segregation in *everyday mobility*, an outcome more broadly linked with new forms of urban segregation and inequality in contemporary America. As redlined neighborhoods may already have differed systematically from non-redlined counterparts before HOLC grading, we adapt the difference-in-discontinuities approach used in the recent work of Aaronson et al. (2021b) by comparing activity space segregation near *actual* redlining borders (grades B vs. C and D) with *placebo* borders separating neighborhoods that received the same grade but, after reweighting, exhibit similar discontinuities in pre-redlining characteristics. An alternative identification that focuses on actual boundaries where the pre-map cross-border gap in the characteristics HOLC weighed was near zero and that were drawn largely to complete grading polygons (Aaronson et al., 2021b) shows similar results.

We highlight three new findings. First, relative to the metropolitan racial distribution benchmark, places in redlined neighborhoods within a quarter mile of the border exhibit excess experienced racial segregation. The effect is estimated within establishment categories and changes little when the neighborhood's present-day income is further conditioned on. Second, the observed effect operates predominantly through changes in neighborhood racial composition resulting from redlining. Being rated as "declining" or "hazardous" increases a neighborhood's Black share relative to its metropolitan area and decreases its white share, which may shape the racial composition of visitors through racialized preferences over neighborhoods (Besbris et al., 2015). We find redlining barely shifts the mix of establishment types that could itself sort visitors (Athey et al., 2021) across the boundary and accounts for essentially none of the estimate. Finally, the causal effect of redlining on experienced racial segregation varies by place type and the average travel distance of typical visitors. Religious; retail and grocery; and recreation, cultural, and dining places, where visitors may be more likely to hold habitual and racialized preferences, show the largest long-term effects of redlining. The effect is further larger at places whose visitors travel shorter distances and are therefore more likely to be aware of the local neighborhood's racial context.

We make three contributions to the literature on mobility-based experienced segregation, the enduring consequences of redlining, and the long-term effects of bureaucratic classification. First, we advance research on mobility-based segregation from descriptive analysis towards identifying its gradient in historical exclusionary practices: we are among the first to provide design-based evidence on its plausibly causal determinants and, to our knowledge, the first to show that the HOLC grading system—a federal housing-finance classification regime created in the 1930s—shaped racial segregation in the everyday establishments Americans visit nearly a century later. Second, we broaden research on redlining beyond its well-documented consequences for residential sorting, wealth, and neighborhood investment, showing that its legacy extends from where people live and accumulate resources to how they navigate and experience the city in their daily lives. Third, our findings contribute to broader economic-sociological theories of classification by showing, consistent with research on classification situations (Fourcade and Healy, 2013; Scott, 1998), how HOLC grades not only described urban differences nearly a century ago but also helped institutionalize a ranked spatial order that became embedded in neighborhood resources, demographic patterns, and everyday mobility.

From Residential Neighborhoods to Activity Spaces

Residential segregation is a foundational dimension of U.S. urban inequality and stratification (for its recent dynamics, see Elbers, 2024). Neighborhood studies in sociological research since the early Chicago School have often placed residential space at center stage in explaining inequality-related outcomes (Massey and Denton, 1993). By unevenly concentrating resources and risks across neighborhoods, residential segregation by race (and income) has been linked to a range of negative outcomes, including high incarceration rates, exposure to violence, low academic achievement and upward social mobility, and poor housing stability among residents of segregated neighborhoods (Sharkey and Faber, 2014). Yet this long tradition of neighborhood research typically represents residential neighborhoods and their surrounding contexts using predefined census geographies, such as census tracts and block groups. While these units make residential neighborhoods measurable, recent studies increasingly question its capacity in capturing, accurately and completely, the full contextual influence of place-based social environments (Kwan, 2012): opportunities for intergroup contact, for example, depend both on who lives nearby and on who shares daily destinations (Cagney et al., 2020; Wong and Shaw, 2011). In this regard, neighborhood-based studies of segregation inevitably neglect the non-home spaces that people encounter outside their residential areas while working, conducting daily economic transactions, attending events, and visiting other neighborhoods (Cagney et al., 2020; Pinchak et al., 2022).

As residential neighborhoods may not capture the totality of social exposure and exchange that individuals experience, recent scholarship has proposed a shift toward studying *activity spaces* outside residential neighborhoods, where a large share of waking hours is effectively spent (Jones and Pebley, 2014). Early work using the Los Angeles Family and Neighborhood Survey shows that, while individuals still experience substantial racial segregation in their daily activities that largely mirrors the characteristics of their residential neighborhoods, activity spaces overall are more racially and economically diverse than residential neighborhoods (Jones and Pebley, 2014). More recent studies using large-scale GPS-based mobility data (Athey et al., 2021) and geotagged social media data (Candipan et al., 2021) across the country similarly show that the racial segregation individuals experience in their daily mobility is substantially lower than residential segregation, although the two measures are highly correlated. In particular, places such as restaurants and public parks exhibit greater racial mixing than their surrounding neighborhoods, while religious and educational institutions remain among the most segregated venues (Jones and Pebley, 2014). Similar findings of lower but persistent ethnic and/or racial segregation in daytime mobility than in residential space have been reported outside the U.S., including in Europe (Järv et al., 2015; Silm and Ahas, 2014), New Zealand (Hołubowska and Poorthuis, 2025), and Israel (Schnell and

Haj-Yahya, 2014)—with a parallel Chinese literature tracing the socioeconomic differentiation of daily routines (Wang et al., 2012)—suggesting a broader regularity in which everyday mobility creates opportunities for intergroup mixing without fully dissolving the ethnic and racial boundaries embedded in residential space.

Before turning to variation in the segregation that persists within activity spaces in the next section, it is important to clarify how previous research has operationalized activity space segregation. Previous studies have adopted three broad approaches, distinguished by their unit of observation; we follow the last. In *people-based* measures, researchers link individual- or device-level visits across destinations and summarize the racial composition of the environments encountered, often weighting destinations by visit frequency or duration (Athey et al., 2021; Moro et al., 2021; Wong and Shaw, 2011). In *flow-based* measures, scholars aggregate individual trips into origin–destination flows, usually between pairs of neighborhoods; segregation is assessed by whether the resulting mobility network disproportionately connects neighborhoods or populations with similar rather than different racial compositions (Candipan et al., 2021; Phillips et al., 2021). Finally, in *place-based* measures, which this study adopts, researchers pool the visits received by each establishment or location and characterize that place according to the racial composition of its visitors, often inferred from their home neighborhoods (Baldenius et al., 2025; Müürisepp et al., 2022). Each approach is suited to different questions about mobility segregation, without any being definitively superior to the others; importantly, segregation detected by one measure is generally informative about related segregation under another, and we accordingly review these strands of literature together below.[2] We adopt the place-based approach to examine whether destinations located in neighborhoods targeted and shaped by historical exclusionary practices continue to reproduce segregation in the daily encounters they host, including through visitors arriving from elsewhere in the metropolis, without claiming to reconstruct individuals' complete activity spaces or encounters across their trajectories.

How Activity Spaces Become Segregated

Despite the overall elevated social mixing in activity spaces, previous studies have by no means downplayed the persistent segregation that remains in daily encounters and its important implications for urban stratification. Krivo et al. (2013), for example, highlighted comparable social isolation experienced by residents of both highly disadvantaged and highly advantaged neighborhoods, as the two groups, compared with residents in the middle range, spend time in largely non-overlapping parts of the city. Being Black in highly disadvantaged neighborhoods incurs additional penalties, as their daily routines are disproportionately concentrated in areas that are both more racially and economically segregated. Similarly, Pinchak et al. (2022) found that Black adolescents experience higher levels of racial segregation and violent crime, along with lower levels of collective efficacy, in their daily routines outside the home than White youth, even after accounting for a host of individual- and home-neighborhood-level characteristics. Activity space, in this regard, becomes another

[2] For example, when research identifies high racial segregation using people- or flow-based measures, destination *places* are also likely to have visitor pools whose racial composition differs from that of the wider metropolitan population, reflecting disproportionate visitation from particular neighborhoods of origin.

dimension of urban inequality: while the scarcity of establishments such as pharmacies, grocery stores, and conventional banks in high-poverty, minority neighborhoods may push residents to seek resources in more advantaged communities (Browning et al., 2021), persistent activity space segregation continues to limit their exposure to demographically diverse populations and resources.

Accordingly, recent studies have increasingly shifted from documenting segregation in daily mobility to explaining why it persists, emphasizing several *present-day* sources of variation. The first operates at the macro level through the urban landscape: population density, transit infrastructure, and the spatial distribution of amenities structure which destinations are accessible and where different groups can converge. Consistent with this account, Athey et al. (2021) find lower experienced racial segregation in denser metropolitan areas and in those with greater public-transit use. Relatedly, travel time plays a first-order role in segregated consumption, and racial sorting in neighborhood-to-neighborhood mobility diminishes substantially when destinations are matched on driving, walking, and public-transit travel times (Davis et al., 2019; Vachuska, 2023). The amenity landscape also shapes opportunities for mixing: restaurants and shared civic spaces can attract relatively diverse visitors, whereas religious and educational institutions remain markedly segregated (Klinenberg, 2018); metropolitan areas with more abundant and accessible social infrastructure provide greater opportunities for cross-racial encounters (Athey et al., 2021).

Meso-level factors, such as residential segregation, and micro-level characteristics, including individual income, have also been widely discussed as important sources of variation in racial segregation across activity spaces. At the meso level, neighborhoods with disproportionately large Black populations are frequently avoided by white visitors (Anderson, 2015). Part of this sorting reflects visit homophily and racialized preference: even after accounting for travel time and neighborhood socioeconomic conditions, people are more likely to visit neighborhoods resembling their own (Vachuska, 2023). This tendency is especially pronounced among residents of white neighborhoods and may be stronger for destinations organized around durable memberships and recurring social ties, such as churches, than for those organized around relatively anonymous transactions (Dougherty, 2003; Gordon, 1964). At the individual level, unequal income and purchasing power—both closely intertwined with racial inequality—shape which destinations people can afford and reach, producing class-differentiated patterns of consumption and mobility traces that overlap substantially with racial segregation (Davis et al., 2019; Moro et al., 2021). These factors, operating at different levels, are neither mutually exclusive nor necessarily competing; rather, they may jointly reinforce the persistent racial segregation observed across activity spaces.

What remains largely missing from this literature is an *institutional* perspective on activity space segregation: whether patterns commonly attributed to present-day dynamics are also durable products of policy and formal governance, and whether those institutional effects are plausibly causal. This perspective carries important implications as contemporary preferences and constraints do not operate on a neutral urban landscape, but are embedded in a geography produced by historical institutions and racial regimes (Baker, 2022). Mobility-based segregation, after all, aggregates moving trajectories between residential origins and activity destinations, and exclusionary policies that reorganized where groups live and where resources are located may leave lasting imprints on the resulting encounters. Salient examples include racialized zoning, which used density restrictions to impede minority entry into majority-white neighborhoods while concentrating manufacturing activity in minority communities (Shertzer et al., 2016), and HOLC's grading of American neighborhoods in the 1930s, whose causal effects on contemporary residential segregation have been recently documented (Aaronson et al., 2021b; Faber, 2020). HOLC grading that this paper focuses on provides a particularly revealing and empirically tractable case: by formalizing racialized assessments of neighborhood risk, it could redirect investment and development, alter the distribution of organizational and amenity resources, and reinforce racialized residential patterns, including Black ghettoization (Faber, 2020)—all of which may continue to structure who visits activity locations across communities.

Redlining and Its Long Reach

Created in 1933 as a temporary New Deal agency to refinance distressed home mortgages and prevent foreclosure, the Home Owners' Loan Corporation (HOLC) issued more than one million loans between 1933 and 1936 (Fishback et al., 2024; Hillier, 2005). As its lending program wound down, HOLC undertook the Federal Home Loan Bank Board's City Survey Program, producing confidential "Residential Security" maps for 239 cities between 1935 and 1940. Prepared with input from local appraisers, lenders, and real-estate professionals, the maps classified residential areas into four grades of perceived long-term investment risk: A ("Best," green), B ("Still Desirable," blue), C ("Definitely Declining," yellow), and D ("Hazardous," red). HOLC staff based these boundaries on housing conditions, property values, land uses, and economic trends, but the appraisals also explicitly considered residents' racial, ethnic, and immigrant composition and the prospect of "racial infiltration" (Fishback et al., 2023; Hillier, 2005). Because the lowest-rated neighborhoods were shaded red and typically had substantial African American populations, these maps came to be associated with the practice of "redlining," through which residents of such neighborhoods were often denied access to credit for home purchases and improvements until racial discrimination in housing was prohibited by the Fair Housing Act of 1968. Credit may also have been restricted in neighborhoods receiving the next-lowest grade, marked yellow—a practice that has received far less public and scholarly attention but has been shown to have similarly persistent long-term effects on homeownership and housing values ("yellowlining") (Aaronson et al., 2021b). Hereafter, we use "redlining" as shorthand for HOLC's grading practice and, in our empirical analyses, for the combined effect of assignment to a C ("yellowlined") or D ("redlined") grade rather than a B grade.

While scholars caution that about 90% of HOLC loans in the 1930s preceded the mapping program and disagree about how widely the maps circulated (Hillier, 2003), historians such as Woods (2012) argues that federal regulations required member institutions (local banks) to adopt similar appraisal standards and construct maps of their own lending geographies, which essentially replicated HOLC maps and diffused its classification rationale. Thousands of real-estate professionals involved in producing the maps also remained active after the Second World War, providing an additional channel through which HOLC grading standards outlasted the program (Aaronson et al., 2021b). In this regard, although lenders' direct use of the maps or their underlying information remains actively debated, HOLC grades nonetheless codified, through an authoritative federal classification, a broader public–private appraisal regime that treated racial composition as an indicator of investment risk (Fishback et al., 2024) and attached lasting stigma to lower-graded neighborhoods (Faber, 2020).

Even as this historical debate continues, a growing body of quasi-experimental research finds that HOLC grading did more than codifying preexisting disadvantage: it altered subsequent neighborhood trajectories in various dimensions that widened racial and income inequality. Comparing actual grade boundaries with propensity-weighted counterfactual boundaries, Aaronson et al. (2021b) find that lower grades reduced homeownership, house values, and rents while increasing racial segregation through the concentration of Black population over subsequent decades. These effects were often at least as large at B–C ("yellowlined") boundaries as at C–D boundaries. Similarly, using a city-level difference-in-differences design, Faber (2020) shows that HOLC-appraised cities became more Black–white segregated than unappraised cities after 1930, with the divergence persisting through 2010. Redlining left its imprint through at least two mutually reinforcing pathways: the diversion of investment from Black to white communities and the reinforcement of racial biases among home seekers, real-estate professionals, and local residents (Faber, 2020). Such channels may likewise mediate the potential effect of redlining on racial segregation in activity spaces, as we elaborate later.

The consequences of redlining also extended across generations and into neighborhood social and physical environments. Applying boundary designs similar to those in Aaronson et al. (2021b) to Opportunity Atlas data, Aaronson et al. (2021a) identify adverse effects on adult income and upward mobility, family structure, and credit profiles among cohorts raised decades after the maps were drawn. Using the same strategies after linking children enumerated in the 1940 Census to later IRS and 2000 Census records, Aaronson et al. (2023) further document lower educational attainment and adult income, as well as residence in adulthood in neighborhoods characterized by lower educational attainment, higher poverty, and more single-parent households. Extending the evidence to public safety, Anders (2023) combine the population cutoff governing which cities were mapped with a spatial regression-discontinuity design in Los Angeles, finding that HOLC mapping increased present-day crime partly through greater segregation and lower educational attainment. Finally, comparing properties within narrow buffers on opposite sides of grade boundaries across 202 cities, Salazar-Miranda et al. (2024) connect lower grades to greater current and projected heat and flood risks, partly through reduced tree canopy and ground permeability.

Few studies, however, have examined whether historical exclusionary institutions such as redlining shaped racial segregation in activity spaces, although the two pathways emphasized by Faber (2020) imply such a legacy: disinvestment may have altered the type, quality, and metropolitan reach of local establishments and their capacity to attract diverse patrons, while institutional stigma may have reinforced Black residential concentration and discouraged outside visitation (Besbris et al., 2015). Direct evidence remains sparse. Associational studies link HOLC grades to city-specific changes in pandemic-era park use (Li et al., 2024a) and former sundown towns to wider racial disparities in park visitation (Rigby et al., 2026), but neither identifies the HOLC effect on racial sorting across ordinary metropolitan establishments. The closest design-based precedent as ours, Aaronson et al. (2025), documents persistent grade- and class-based sorting in neighborhood flows but does not assess racial segregation within establishment visitor pools. As activity spaces constitute an increasingly consequential dimension of urban stratification, we ask whether redlining plausibly caused lasting changes in their racial composition relative to the surrounding metropolitan population and whether any effect operates through channels such as changes in neighborhood amenities and residential racial composition. We use national mobility data introduced below linking visitors' home neighborhoods to establishments to assess these questions.

Data, Measures, and Methods

Data

SafeGraph mobility data. To measure racial segregation among visitors at each place, we use mobility data from SafeGraph's monthly Patterns files for 2019, obtained through the SafeGraph COVID-19 Data Consortium. For each of roughly four million points of interest (POIs) in 384 U.S. metropolitan statistical areas (MSAs)—restaurants, grocery stores, churches, schools, health facilities, and other establishments—the data record monthly visit counts and the census block groups (CBGs) where visitors reside, inferred from GPS pings of a panel of anonymous mobile devices covering roughly 10% of the U.S. population (Kang et al., 2020). We restrict every analysis to visitors whose home CBG lies in the same MSA as the POI. Tourists, cross-metro travelers, and foreign visitors therefore do not enter our measures: what we capture is how residents of a metropolitan area sort across that area's activity locations. Each POI carries a North American Industry Classification System (NAICS) category, which we use as fixed effects throughout.

Sampling bias and data quality. Because SafeGraph's device panel over-represents some neighborhoods and under-represents others—and because its GPS-based assignment of each device's home CBG carries some error—we apply the post-stratification correction recommended by SafeGraph: each home CBG j receives an adjustment factor equal to its population share divided by its device share, $w_j = (\mathrm{pop}_j/\sum_j \mathrm{pop}_j)\,/\,(\mathrm{dev}_j/\sum_j \mathrm{dev}_j)$, computed from the Patterns home panel summary and census population counts, and visitor counts from CBG j are multiplied by w_j before any aggregation. Over-represented CBGs are thereby down-weighted, and all measures reported below use these reweighted flows; results are essentially unchanged relative to the unweighted measure (Appendix I). A distinct quality concern attaches to the establishment records themselves, which are known to contain misclassified categories, duplicate listings, and closed establishments (Grigoropoulou and Small, 2025). Because our outcome is computed from visitor origins rather than from the category label, and every model conditions on category fixed effects, such errors are unlikely to drive the boundary contrast; we return to this limitation in the Discussion.

Census and neighborhood data. CBG-level demographic and socioeconomic covariates come from the 2019 American Community Survey five-year estimates. For the historical analysis we use the Home Owners' Loan Corporation (HOLC) security maps digitized by the Mapping Inequality project (Nelson et al., 2023) and, for the boundary design described below, the boundary files and 1910–1930 census tract covariates constructed by Aaronson et al. (2021b) from full-count historical censuses. County-level land-use regulation that we control in the descriptive analysis (national gradient) is measured with the Wharton Residential

Land Use Regulatory Index (Gyourko et al., 2008), which combines a series of restrictions on the use of space—local zoning, density limits, and affordable-housing requirements, among others.

Measuring Activity Space Segregation

Definition and unit of analysis. We define activity space segregation as the degree to which the racial composition of visitors to an activity location deviates from the racial composition of its metropolitan area. Our unit of observation is the location, not the individual: devices are not identified across locations in the Patterns data, so individual activity spaces cannot be reconstructed, and we accordingly measure segregation *in* activity locations rather than the segregation of any person's full activity space (for the same design choice, see Baldenius et al., 2025). Within the typology of activity-space research, ours is a standard place-based approach (Müürisepp et al., 2022). We discuss potential limitations of the approach at the end.

The integration benchmark. Because SafeGraph provides no demographic information on individual devices, we infer visitor composition from the racial makeup of visitors' home CBGs, following prior work (Athey et al., 2021; Candipan et al., 2021; Chang et al., 2021; Moro et al., 2021). Let v_{jk} denote the (bias-adjusted by w_j) visit flow from CBG j to POI k, and let γ_{ji} denote the share of racial group i (non-Hispanic white, Black, Hispanic, Asian, and mixed/other) among residents of CBG j.[3] We compare the visitor share of group i at POI k against the metropolitan group composition:

$$y_k^m \;=\; \frac{5}{8}\sum_{i=1}^{5}\left|\frac{\sum_j \gamma_{ji} v_{jk}}{\sum_j v_{jk}} - \delta_i^m\right|, \tag{1}$$

where y_k^m is the segregation of POI k relative to its metropolitan area m, δ_i^m is group i's share of that metropolitan population, and the factor 5/8 normalizes the index to the range of 0 to 1. A value of zero means the location's visitors mirror the metropolitan composition: full *integration* in the sense of classic metropolitan segregation indices, which likewise take the areal composition as the reference point; larger values mean the location draws a visitor mix increasingly unrepresentative of its metro. Figure 1 illustrates the two steps: how a location's visitor composition is assembled from its visitors' home neighborhoods, and how its distance from the benchmark defines segregation. This composition counts every same-metropolitan-area visitor to a location, including those who live in its own neighborhood; because such local residents are a small share of visits, excluding them leaves both the measure and its estimated response to redlining essentially unchanged (Appendix A).

In Appendix E, we also compute a second measure that replaces δ_i^m in Eq. (1) with equal shares ($\delta_i = 1/5$), which measures deviation from ideal *diversity* (as opposed to our main ideal

[3] γ_{ji} is the CBG's fractional racial composition, so each visit contributes its origin neighborhood's full racial mix, exactly as Figure 1 illustrates. Assigning each CBG its largest group instead ($\gamma_{ji} \in \{0, 1\}$), a common simplification, roughly doubles the level of the index and yields a proportionally larger boundary effect; ecological-inference methods that relax the assumption that groups within a CBG travel alike likewise leave the effect intact (Appendix B).

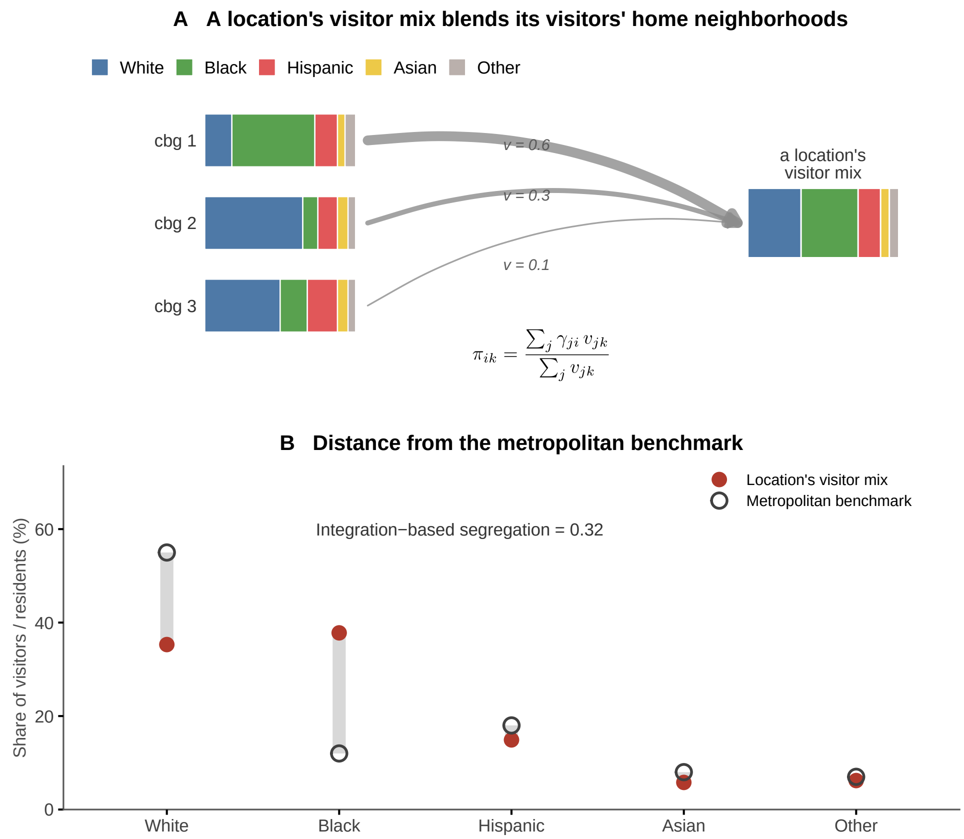

Figure 1: Constructing the integration measure. *Panel A*: a location's visitor composition π_{ik} is a flow-weighted blend of the racial compositions of its visitors' home neighborhoods (CBGs), each weighted by its visit flow v_{jk} to the location. *Panel B*: integration segregation is the summed absolute distance between that visitor composition and the metropolitan benchmark δ_i^m (grey bars mark the per-group deviations; shares are illustrative).

integration measure) and captures distance from maximal group balance rather than from the metropolitan composition (cf. the random-pairing index underlying the national experienced-diversity dataset of Xu et al., 2024). The two measures can and do diverge—a 20%-white location in a majority-white metro may sit closer to even shares (high diversity) while sitting far from the metro composition (low integration). We treat the integration benchmark as focal on theoretical grounds corresponding to the classic metropolitan-segregation literature (metropolitan indices such as the dissimilarity index D and the multigroup entropy index H) that take the metropolitan composition as their reference (Massey and Denton, 1988; Reardon and Firebaugh, 2002), alongside recent experienced-segregation work that normalizes by metropolitan racial composition (Baldenius et al., 2025). Importantly, applying the boundary design to this diversity measure yields no redlining effect (Appendix E), underscoring that the policy's imprint is specific to each metropolitan area's racial context.

Catchment size. Alongside visitor composition, the Patterns files report for each establishment the median distance from home traveled by the visitors whose home location

SafeGraph identifies. We use this quantity as a measure of *catchment size*: the geographic reach of a location's clientele, and thus how locally a place draws. Because trip distances are heavily right-skewed, we use it in ranked rather than absolute form, sorting establishments into six equal-sized bins when we examine where the boundary effect concentrates (Section). It offers a continuously measured alternative to the recorded establishment category, reading localness off visitor travel rather than off a commercial label.

Redlining and Covariates

Our focal historical exclusionary policy is redlining. A POI is coded as redlined if it falls in an area graded C ("declining," colored yellow) or D ("hazardous," colored red) on the 1930s HOLC security maps, as against the higher grades A ("best," green) and B ("still desirable," blue). Grade-A areas enter none of our estimation samples: the descriptive comparison sets redlined POIs against grade-B and unmapped neighborhoods, and the boundary design below compares grade-B locations to their C or D neighbors; we separate the two lower grades where the distinction matters.

The pre-treatment covariates that discipline the boundary design come from the complete-count decennial censuses of 1910, 1920, and 1930, harmonized to 2010 census-tract boundaries by Aaronson et al. (2021b); the 2010 tract is thus the unit at which these covariates are observed throughout. They comprise the characteristics HOLC's assessors weighed: each tract's Black population share, homeownership rate, foreign-born share, and literacy rate in all three census years, together with house value, rent, and radio ownership in 1930. The design below uses them not as regression controls but as balancing targets: real and counterfactual boundaries are matched on their pre-map cross-border gaps in these characteristics, so that only differences arising after the maps were drawn remain in the estimate.

A fully controlled descriptive specification, which we report below to show that the national gradient is robust to present-day differences between neighborhoods, additionally includes the Wharton Residential Land Use Regulatory Index, a county-level measure of the stringency of local zoning and land-use regulation (Gyourko et al., 2008), and a vector of contemporary community characteristics measured at the CBG level: majority racial group; median household income quartile; national walkability index; proximity to a primary highway (within 1,000 m); central-city location (whether the POI lies in its metropolitan area's principal city rather than a suburb); percent with high school education or above; percent aged 18–34; percent in the labor force; percent foreign born; and percent commuting by public transit. These contemporary controls do not enter the main boundary difference-in-discontinuities identification strategy: the county-level regulatory index is identical on a boundary's two sides and absorbed by the boundary fixed effects, while most CBG characteristics are themselves consequences of grading (i.e., post-treatment controls; we instead explore some of their mediation effects below), which the design nets out by balancing *historical* covariates instead. All models include POI-category fixed effects, and the national models include MSA fixed effects.

Analytic Strategy

Our analysis proceeds in three steps: a descriptive national gradient, a boundary identification design, and a mechanism analysis.

Step 1: Descriptive models. We first estimate POI-level regressions comparing redlined (C/D) POIs with grade-B "still desirable" areas (in consistency with our boundary design below) together with the neighborhoods HOLC never mapped (thus largely unaffected by exclusionary policies),

$$y_k^m \;=\; \beta_1 D_k + \mu_{m(k)} + \theta_{\ell(k)} + \sum_p \beta_p X_{kp} + \epsilon_k, \tag{2}$$

where $D_k = \mathbb{1}(\text{redlined})$ marks C/D locations, μ_m are metropolitan-area fixed effects, θ_ℓ are POI-category fixed effects, and X_{kp} are contemporary characteristics described above; standard errors are clustered by county.

While the national gradient model remains descriptive and we will proceed with a boundary identification design, we supplement the descriptives with a Oster (2019) bound analysis that reports the strength of selection into redlining treatment based on unobservables (relative to contemporary observables, even if they are post-treatment characteristics) that would be needed to drive the coefficient to zero, with $R_{\max} = 1.3\tilde{R}$; we report the detailed bound analysis in Appendix D.

Step 2: Boundary difference-in-discontinuities. The descriptive gradient cannot identify redlining's causal effect: HOLC graded neighborhoods that already differed—attending explicitly to residents' race and class (Hillier, 2003)—so greater present-day segregation on the lower-graded side need not stem from the map. To separate the grade's effect from these pre-existing differences, we adapt the counterfactual-boundary design of Aaronson et al. (2021b), comparing neighbors just across an HOLC border. Its identifying assumption is parallel trends: absent the maps, segregation would have evolved similarly on the two sides of a border. While activity space segregation itself cannot be observed in the 1930s, the assumption remains largely plausible as the important covariates that determined redlining are observed in 1910-1930 and, through the reweighting below, matched between real and comparison borders in every census in and before 1930—in levels and therefore in trends. The design requires a discontinuity fixed at a known historical moment, which is why redlining is our focal policy: HOLC's grades were drawn once, in the 1930s, at boundaries never redrawn. Other exclusionary levers offer no such fixed line: zoning regimes, for instance, have been revised continuously over the intervening decades, leaving no single historical moment at which treated and untreated space divide. We therefore treat exclusionary zoning as descriptive context (Appendix J) and leave its causal analysis—for which historical zoning maps offer a promising foothold (Shertzer et al., 2016)—to future work.

Specifically, we inherit Aaronson et al. (2021b)'s border *segments*—short pieces of the grade line, each at least a quarter mile long (median 0.37 miles), so that every comparison is confined to the two sides of the same stretch of border—and treat one segment, with a quarter-mile buffer on its higher-graded (B) and lower-graded (C or D) side, as a "boundary" b.[4] Because our 1910–1930 balancing covariates exist only for 2010 census tracts, this design assigns each POI to a side through its tract: a tract is linked to a side when its area overlaps that side's buffer; because tracts are typically larger than the buffer, 42% of tract–boundary pairs straddle the border and enter both sides.[5] A POI standing on the B side can therefore be counted among the C- or D-side observations whenever its tract reaches the opposite buffer. Such blurring can only mix the two sides and attenuate the discontinuity toward zero, and weighting each tract by its buffer-overlap share, as Aaronson et al. (2021b) do, changes the coefficient by only about 1.6%. We report this tract-based design first out of fidelity to Aaronson et al. (2021b), whose harmonized data define its boundaries, covariates, and counterfactual grid.

The second assignment drops the tract detour where it is not needed. The outcome, unlike the covariates, is observed at the POI itself, so we read each POI's grade from the HOLC polygon it occupies and place it on a side by its own position within the quarter-mile buffer; every establishment then enters unambiguously as B or as C/D, and only its reweighting covariates remain tract-level. Freed of side-blurring, this more direct measurement of the boundary contrast yields a larger estimate (Column 5 of Table 1), implying that the tract-based headline is, if anything, conservative.

To net out pre-map differences, we benchmark each real boundary against counterfactual (or placebo) boundaries where the grade never changed: segments of a half-mile grid, laid over each city, that fall within a single grade (B–B, C–C, or D–D)—but that differ discontinuously in covariates after reweighting (details below). Lacking a natural worse side, each is assigned a pseudo–lower side at random, averaged over 100 draws. After dropping boundaries split by a river or railroad and those without visitor data on both sides, the estimator pools 998 real B–C/B–D segments with 8,740 grid placebos; the POI-level analysis we described above instead uses whole B–(C/D) polygon borders (3,187 real, 17,161 in all, Column 5 of Table 1).[6] Figure 2 illustrates the design, drawing the border inventory of the second, POI-level strategy.[7]

[4] Aaronson et al. (2021b) study adjacent-grade borders only (C–B and D–C). Our B–C borders are their C–B; we also include the smaller number of B–D borders, which pair grades two steps apart, because our treatment contrasts redlined areas (C or D) with still-desirable areas (B). Given the small number of B–D borders, our results are similar even if only B–C borders are studied.

[5] Because these covariates are observed for whole 2010 tracts, a tract picked up by the quarter-mile buffer may also reach into other HOLC grades, so the balance condition below matches tract-measured rather than strictly buffer-local pre-map gaps.

[6] The inventories differ for concrete reasons. Aaronson et al. (2021b)'s harmonized files cover 144 cities and contain 13,163 border segments of all grade pairs, but only 1,750 of these are B–C or B–D: most grade lines separate other pairs (C–D, A–B), which do not form our treatment contrast. Excluding segments split by a railroad or river leaves 1,171, and requiring tracts with same-MSA visitor data on both sides leaves 998. Each design carries its own placebo inventory built by the same recipe as its real borders: the segment design draws on the 17,438 same-grade (B–B, C–C, D–D) segments of Aaronson et al. (2021b)'s counterfactual grid, of which 8,740 survive the same screens, whereas the POI-level design lays its own half-mile grid and keeps the 13,974 cell edges with same-grade establishments within a quarter mile on both sides. Its real-border count is larger than the segment count because it is built from the full digitized HOLC map rather than the 144 harmonized cities, and a whole polygon border needs only establishments (not covariate-bearing tracts) on its two sides.

[7] Results are similar if boundaries split by a river or railroad are not excluded; we report the estimates

The estimating equation pools the real and counterfactual boundaries:

$$y^m_{kb} \;=\; \gamma\, D_{kb} \;+\; \beta_{DD}\, D_{kb}\, R_b \;+\; \alpha_b \;+\; \theta_{\ell(k)} \;+\; \epsilon_{kb}, \tag{3}$$

where $D_{kb} = \mathbb{1}$(lower-graded side) and $R_b = \mathbb{1}$(real boundary); α_b are boundary fixed effects (so every comparison is across the two sides of one border) and $\theta_{\ell(k)}$ are POI-category fixed effects. We estimate Eq. (3) by weighted least squares,

$$\big(\hat\gamma, \hat\beta_{DD}, \hat\alpha, \hat\theta\big) \;=\; \underset{\gamma,\,\beta,\,\alpha,\,\theta}{\arg\min} \sum_{k,b} W_{kb}\,\big(y^m_{kb} - \gamma D_{kb} - \beta\, D_{kb} R_b - \alpha_b - \theta_{\ell(k)}\big)^2, \tag{4}$$

with weight $W_{kb} = w_b\, v_{kb}$, where v_{kb} is the POI's log visit count and w_b is the entropy-balancing weight on boundary b (Eq. (5) below; $w_b \equiv 1$ for real boundaries), which reweights the counterfactual pool to match the real boundaries' pre-map covariate gaps. Standard errors are clustered by boundary. The coefficient of interest, β_{DD}, is the excess segregation jump at a real graded line over a reweighted same-grade line matched on observed pre-map characteristics.

The reweighting makes "observationally equivalent" precise. For each boundary we form the within-boundary gap $z_b^{k,t} = x^{k,t}_{lgs,b} - x^{k,t}_{hgs,b}$ between its lower- and higher-graded side in characteristic k and year t: the Black share, homeownership, foreign-born share, and literacy rate for 1910–1930 (principal covariates), and house value, rent, and radio ownership for 1930: fifteen gaps in all. Real boundaries carry systematic gaps (the C/D side was poorer, Blacker, less owner-occupied); same-grade grid boundaries carry only idiosyncratic ones. Where Aaronson et al. (2021b) weight comparison boundaries by a propensity score, we use entropy-balancing weights (Hainmueller, 2012): the minimum-divergence weights w_b on counterfactual boundaries satisfying

$$\sum_{b\,\in\,\text{counterfactual}} w_b\, z_b^{k,t} \;=\; \bar z^{\,k,t}_{\text{real}} \quad \text{for all } (k,t), \tag{5}$$

so that the reweighted placebo pool reproduces the real boundaries' covariate gaps exactly (gaps unobserved in early censuses enter as zeros with a matched missingness indicator). Balance is thus exact in every draw, whereas propensity weighting leaves the 1930 house-value and 1920 Black share gap imbalanced (Figure 3; our main estimates are, however, even larger under propensity-score weighting); balancing only the four principal 1930 covariates yields nearly identical results. Inference combines clustered standard errors with the randomization distribution of β_{DD} across the 100 draws.

To corroborate our identification and estimates, in Appendix F, we complement the design with a second identification that needs no comparison group: restricting to the treated boundaries whose pre-map gaps are similar to placebo gaps and that were often demarcated to complete a redlining polygon (Aaronson et al., 2021b). At these borders grading was effectively as-good-as-random, allowing us to compare the two sides directly. We show the estimates as well as the parallel-trends evaluation for this design there.

with such boundaries excluded because neighborhoods beside such boundaries might have developed more differently even with pre-treatment covariates balanced, due to natural causes.

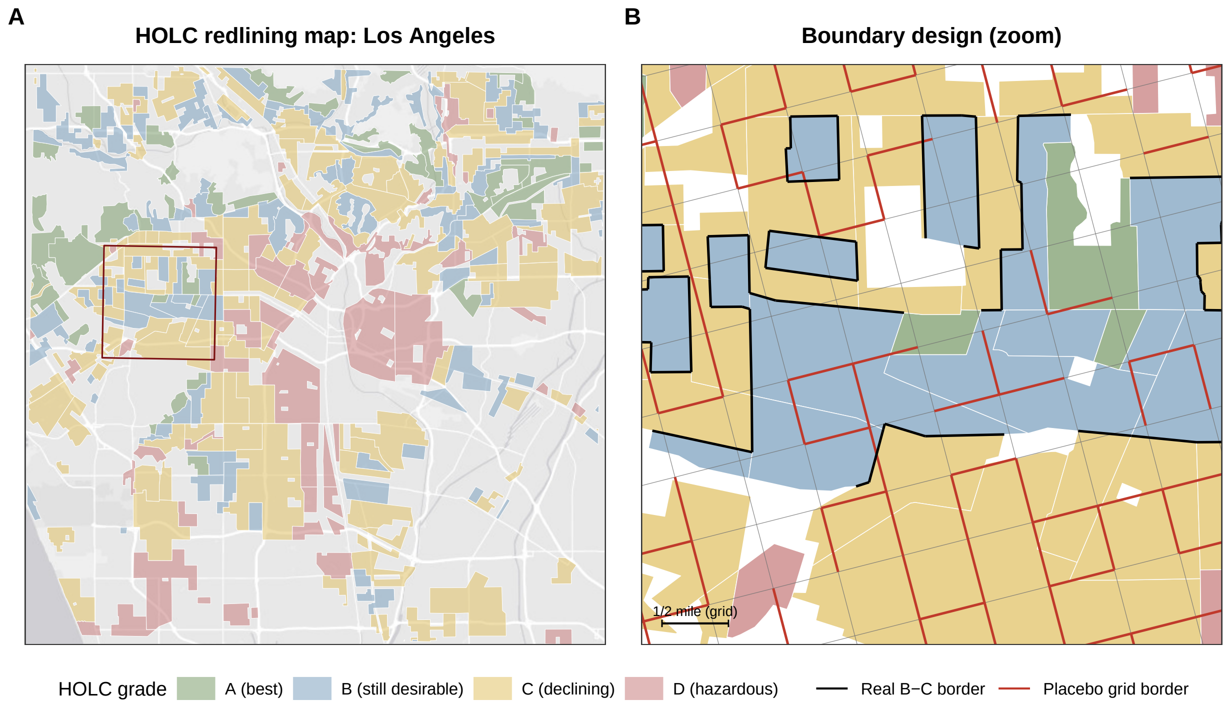

Figure 2: The boundary difference-in-discontinuities design. *Panel A*: HOLC security map of Los Angeles (grades A–D) overlaid on a present-day base map; the dark-red box marks the zoom window shown in Panel B. *Panel B*: real HOLC borders between grade-B ("still desirable") and grade-C ("declining") areas (solid black), the half-mile counterfactual grid laid over the mapped area (grey; drawn in the equal-area projection in which the analysis anchors it, hence its slight rotation), and every same-grade grid border serving as a placebo within this window (red). The panel depicts the POI-level design's border inventory (Column 5 of Table 1); the tract-based design uses Aaronson et al. (2021b)'s analogous border segments and counterfactual grid. POIs within a quarter mile of a border are assigned to its higher- or lower-graded side, and the design compares the segregation discontinuity at real graded lines with that at observationally equivalent lines that were never drawn.

Step 3: Mechanisms. Finally, we explore potential mechanisms through which historical redlining changed activity space racial segregation by focusing on its downstream effect on *residential racial composition*, drawing on both previous studies of the redlining effect on neighborhood segregation (Faber, 2020) and daily mobility preferences across neighborhoods of different racial composition (Anderson, 2015) (e.g., a preference against majority-Black neighborhoods from non-Black population). Specifically, we use the overall distance of the POI's neighborhood (its census block group) from the metropolitan composition, the residential counterpart of our outcome, $M_k = 5/8 \sum_i |\text{share}_{i,\text{nbhd}(k)} - \text{share}_{i,\text{MSA}(k)}|$ summed over the five groups, as the potential mediator. Within the boundary design we estimate the mediation following the classic framework of Baron and Kenny (1986): redlining's effect on M (path a), the association of M with racial segregation in activity space given treatment (path b), and the indirect effect $a \times b$, whose significance we assess with a Sobel test (Sobel, 1982), using the main estimator's fixed effects, weights, and clustering, and then decompose

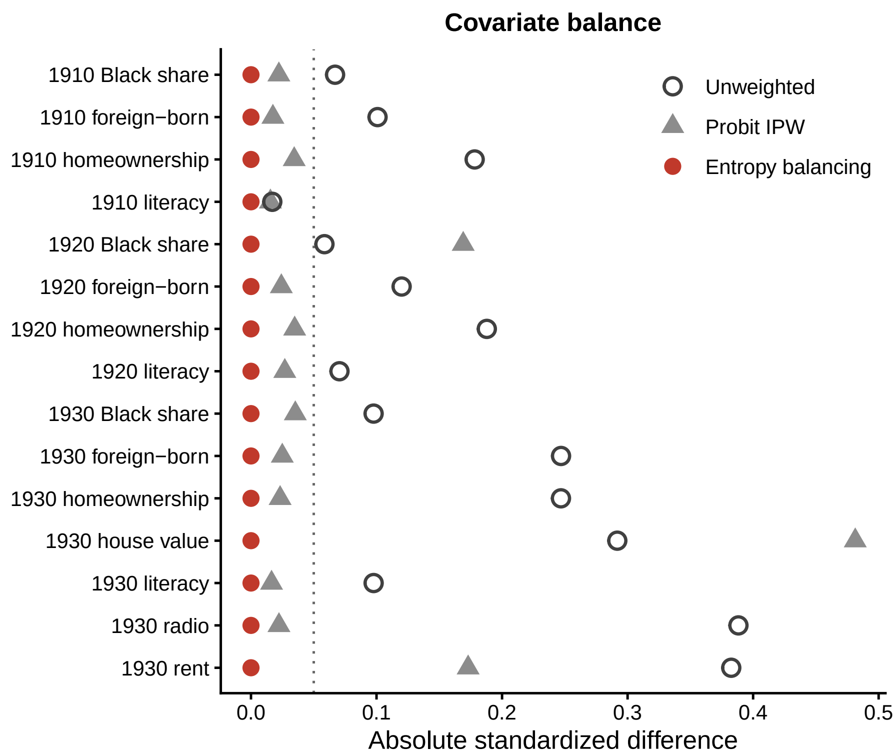

Figure 3: Covariate balance across weighting schemes. Each point is the absolute standardized difference, between real and reweighted counterfactual boundaries, in the lower-minus-higher-graded within-boundary gap for one 1910–1930 historical covariate, averaged over the 100 placebo-orientation draws. Under entropy balancing every gap is matched exactly (points at zero); probit inverse-propensity weighting reduces but does not eliminate the gaps, leaving several (most visibly the 1930 house value) above the conventional 0.05 threshold (dotted line). The unweighted counterfactual pool is shown for reference.

M by racial group to see which groups' displacement drives it (Figure 4).

We also weigh three alternative channels. The first is the *amenity mix*: some kinds of establishments (churches, restaurants) draw more segregated visitor pools than others everywhere, so redlining could raise a neighborhood's measured segregation simply by changing which kinds it hosts (e.g., through diverted resource investments), without changing who visits any given kind. To separate the two, we split each location's segregation into the part its establishment type carries everywhere—the type's national average—and the part specific to that location, and re-estimate the boundary design on each part. The second is the perceived *quality* of redlined places rather than their racial composition: matching each POI to its nearest Place Pulse 2.0 street-image ratings within 500 meters (27,877 rated images across 13 U.S. cities, each scored on six perceptual dimensions—how safe, wealthy, beautiful, lively, depressing, or boring a scene appears; Appendix H), we test within the same boundary design whether being perceived as less safe or lower in quality—and thus avoided by the average visitor—mediates the effect. The third is *income* rather than racial sorting, which we probe in two ways: by applying the same boundary design to an income analog of our segregation measure (the distance of a location's visitor income distribution from its metro's), and by checking that the racial effect persists once the neighborhood's median income is added as a covariate.

Results

A National Gradient in Activity Space Segregation

How much more segregated are the activity spaces of redlined places? We begin descriptively. Comparing points of interest in redlined (grade C or D) areas with grade-B and ungraded neighborhoods—net of metropolitan-area and establishment-category fixed effects and the extensive contemporary control vector described above (neighborhood income, education, majority race, walkability, and land-use regulation)—redlined POIs draw visitor mixes that are 0.020 further from their metropolitan composition (Table 1, Column 1; $p < 0.001$). This is not an artifact of too thin a control set: the POI-category fixed-effects-only gap is larger still, 0.025, meaning conditioning on the many ways redlined places differ today, including those post-treatment controls, removes only about a fifth of the association. While such estimates remain descriptive, the gap is substantively sizable, amounting to about 23% of the pooled standard deviation of POI-level segregation (0.08), a meaningful cross-sectional difference for a location-level measure.

To formally test the sensitivity of the estimate to possible unobserved characteristics that may sort neighborhoods into redlining, benchmarked by the observed characteristics we controlled for above. Specifically, we use the Oster (2019) selection bound (see Appendix D) and find $\delta^* = 2.67$, meaning that selection on unobservables before redlining would have to be more than two and a half times as strong as selection on our observed contemporary controls—themselves an unusually complete contemporary set, including racial composition, median income, and the share of foreign-born residents, characteristics HOLC's own assessors weighed in the 1930s—to reduce the gradient to zero. Note that this is a *national* gradient—a broad association across all redlined places—while the boundary design below that yields a smaller estimate instead identifies the *local* effect at the grade line, leaving only the discontinuity at the border. We therefore read this larger national gradient as suggestive of the scale redlining's imprint may reach nationally if it works through this broader channel and not only at the margin that we present below, even if it remains descriptive.

Identifying the Effect at Redlining Boundaries

We proceed with our main difference-in-discontinuity design and compare racial segregation in activity spaces at the POI level in the buffer area of the two sides of an HOLC boundary against the reweighted placebo treatment effects (Figure 2). Column 2 of Table 1 reports the raw boundary discontinuity: activity spaces on the lower-graded side are 0.0030 more segregated ($p < 0.001$). Only a small fraction of this effect is explained by pre-treatment discontinuity at the border: Column 4 of Table 1 compares this raw estimate with placebo effects estimated at same-grade boundaries reweighted to match the real boundaries' pre-map covariate gaps using entropy weighting, and finds an effect of 0.0025—after netting out pre-existing differences through placebo adjustments.

Importantly, the effect of historical exclusionary redlining on contemporary racial segregation at place-based activity spaces is insensitive to the weighting algorithm (Column 3): the difference-in-discontinuities is 0.0030 (an even larger estimate) under an alternative inverse-propensity weighting, and both weighting schemes return positive estimates in every one of 100 placebo-orientation draws (significant at the five-percent level in 85 and 66 of them, respectively). Column 5 reports what we regard as the sharpest measurement of the boundary contrast: each POI enters unambiguously as B or as C/D by its own distance to the border, with only its balancing covariates remaining tract-level. This assignment recovers many boundaries the coarser tract buffer discards and raises the estimate to 0.0036—consistent with the tract-based blurring of sides attenuating Columns 3–4, and confirming that the headline estimate is, if anything, conservative.

Substantively, this point-level (POI-buffer) estimate of 0.0036 is about 6% of the within-boundary standard deviation of POI-level segregation (0.058, net of boundary and establishment-category fixed effects)—not trivial for so local a contrast. To gauge it against a familiar yardstick, we apply the identical boundary design to the present-day neighborhood economic conditions through which redlining's long-run legacy has been traced: household income and rent (Appel, 2016; Krimmel, 2018). Redlining's imprint on the racial composition of a place's visitors is of the same order as its imprint on these local economic conditions at the same borders (Appendix C; 5% on income and 12% on rent). A grade line abandoned for generations still meaningfully shapes who shows up by day, carrying the policy's documented economic legacy into the domain of daily mobility.

It is important to note that the estimated treatment effect of redlining (on treated neighborhoods) is, *by design*, local at the boundary. It effectively compares neighborhoods that abut one another, and differences away all characteristics the two sides share, including the many ways they are similar in their immediate environment that has largely converged through historical events, movements, and urban developments (Bayer et al., 2007; Black, 1999). Accordingly, the boundary design recovers largely the local discontinuity as a conservative floor; and given that the effect of exclusionary policies may spill over and be more profound in neighborhoods further away from the immediate border that is more dilution-resistant, we note the scale of the national gradient presented above as suggestive of the broader scale the policy may carry.

Mechanism

Why does grading raise integration-based racial segregation in activity spaces? Of the downstream outcomes linked with historical redlining in previous studies (Aaronson et al., 2021b; Faber, 2020), we focus on the changes in neighborhood racial composition induced by racialized exclusionary practices that may consequently shape racial segregation in activity spaces around the border. Specifically, we evaluate whether the neighborhood's (POI's CBG) overall distance from the metropolitan composition—the residential counterpart of our outcome, summed across the five groups (Figure 4)—mediates our main estimates. Using the canonical mediation framework of Baron and Kenny (1986), we find that redlining widens this compositional distance ($a = 0.020$), which strongly predicts racial segregation in activity spaces ($b = 0.061$), for an indirect effect of 0.0012 (Sobel $p \approx 0.002$), about 44% of the total. In other words, a redlined neighborhood's activity space inherits the residential displacement the maps produced.

We further decompose the compositional change by racial group and find that it is specifically a Black–white shift (Figure 4). Grading pushes the neighborhood's Black share above its metro's (by 0.012, $p < 0.001$) and its white share below (by 0.017, $p < 0.001$), while the Hispanic and Asian shares barely move (Panel B). In separate mediation models, the Black distance from the metropolitan composition accounts for about 40% of the boundary effect and the white distance about 31% (each $p \leq 0.01$); because the two shares move as near mirror images these percentages are highly overlapping rather than additive, and with the Hispanic and Asian distances contributing essentially nothing, the compositional channel is in effect a Black–white one. This pattern is consistent with a long line of work on the racialized sorting of which neighborhoods non-Black Americans consider desirable places to live (Charles, 2003; Krysan et al., 2009)—preferences that can carry over into the racially patterned everyday mobility that shapes who shares a place by day (Anderson, 2015; Candipan et al., 2021).

Three rival channels might account for the same pattern. The first is the amenity mix (Athey et al., 2021). Applying the decomposition described above, the channel is a precise null: of a total effect of roughly +0.0027, the amenity-mix component contributes essentially nothing (1%; $p = 0.80$), and the cross-border shifts in category shares are themselves small (under one percentage point of visits) and offsetting. Virtually the entire effect arises *within* establishment types—which is also why the main estimate is insensitive to the category fixed effects. Redlining, in short, changed who visits a given kind of place, not which kinds of places exist across the line.

The second rival is the perceived *quality* of redlined places rather than who lives in them. Using the Place Pulse 2.0 street-image collection for the metros it covers, we enter crowd-rated perceptions as mediators: a composite combining all six rated dimensions (safe, wealthy, beautiful, and lively, with depressing and boring reverse-coded), and each dimension on its own. Neither the composite nor any single dimension shows a detectable indirect effect (mediated shares of about 0–2%), and the difference in perceived safety across the line is small (Appendix H). The third rival is economic rather than racial sorting: because redlined areas are poorer, the racial gap we estimate could be a byproduct of economic sorting that merely correlates with race. Three checks weigh against this. Applying the boundary design to an analogous *income*-segregation outcome—built exactly as the racial measure but on income—returns essentially no effect (+0.0005, $p = 0.70$); the racial effect is little changed by controls for the neighborhood's median income (+0.0023, versus +0.0026) or for income segregation itself; and it does not vary significantly with neighborhood income (interaction $p > 0.3$).

Where the Effect Concentrates

Previous studies show that racial segregation in activity spaces varies systematically by establishment type and by the geographic reach of a location's clientele (Candipan et al., 2021; Small et al., 2021). Redlining's imprint may therefore vary similarly across locations. Indeed, if redlining operates in part by reshaping neighborhood racial composition—and if awareness of that composition influences where people of different racial groups choose to go—its effect should be most pronounced at locations whose clientele is drawn from and

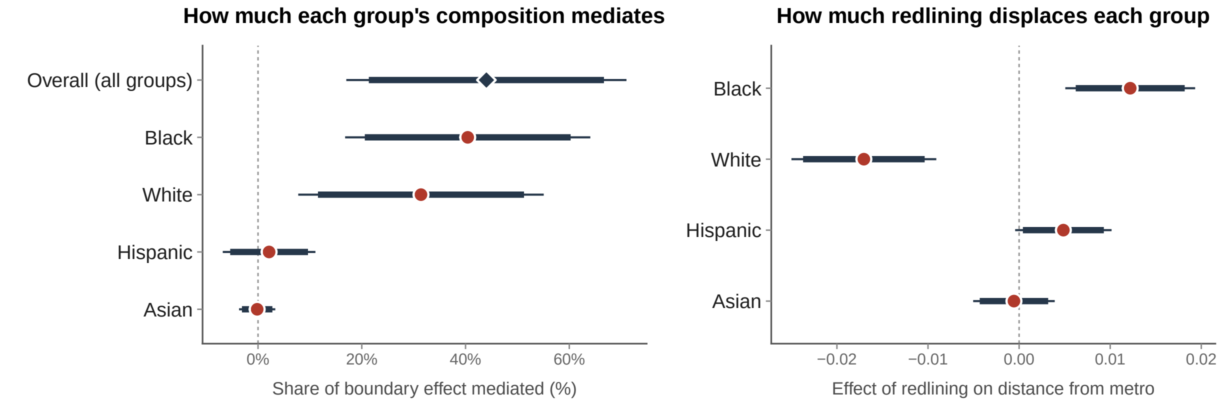

Figure 4: Mechanism: the compositional channel and its racial decomposition. *Panel A*: share of the entropy-balanced redlining effect on integration segregation mediated by the neighborhood's overall distance from the metropolitan composition ("Overall," about two-fifths) and by each racial group's distance from that composition (indirect effect as a percentage of the total). The channel runs on a Black–white axis; the Hispanic and Asian distances mediate essentially none. *Panel B*: redlining's effect on each group's *signed* distance from the metropolitan share ($\text{share}_{\text{nbhd}} - \text{share}_{\text{MSA}}$)—positive values indicate grading pushed the neighborhood above its metro's share for that group, negative below. Thick and thin bars are 90% and 95% confidence intervals; estimates averaged over placebo-orientation draws.

familiar with the surrounding neighborhood: everyday, locally patronized establishments with small catchment areas. Conversely, the effect may attenuate at locations that draw a metropolitan clientele from many neighborhoods. We examine this expectation using two cuts of the data (Figure 5).

Establishment type. We group establishments by their SafeGraph top-level category into broader establishment types (Table G1) and estimate the boundary difference-in-discontinuity within each. Figure 5A shows seven broad establishment types. Religious organizations lead (+0.0049, $p < 0.01$): American congregations are among the most racially homogeneous institutions in American society (Dougherty, 2003; Emerson and Smith, 2000), and a place of worship likely draws a congregation that mirrors the residing community—whose racial composition is itself largely shaped by redlining—as well as nearby residents whose choice of where to worship reflects their perception of the racial makeup of the church and the neighborhood. Everyday grocery, pharmacy, and general retail follow (+0.0036, $p < 0.01$)—the routine shopping people do close to home, over which they likely hold habitual and potentially racialized preferences about place and neighborhood racial composition—with education, dining, and recreation following behind. By contrast, the effect is indistinguishable from zero for types whose patronage is less organized by residential proximity or by frequent and routine visits (the remaining five are in Appendix G): finance, real estate, and lodging, whose offices and destinations draw from across the metropolis, and health care, where patients follow insurance networks, referrals, and specialty.

Importantly, this pattern is not merely mechanical: it holds even when we drop each POI's own block-group residents from the visitor mix (Appendix A), so a place's activity

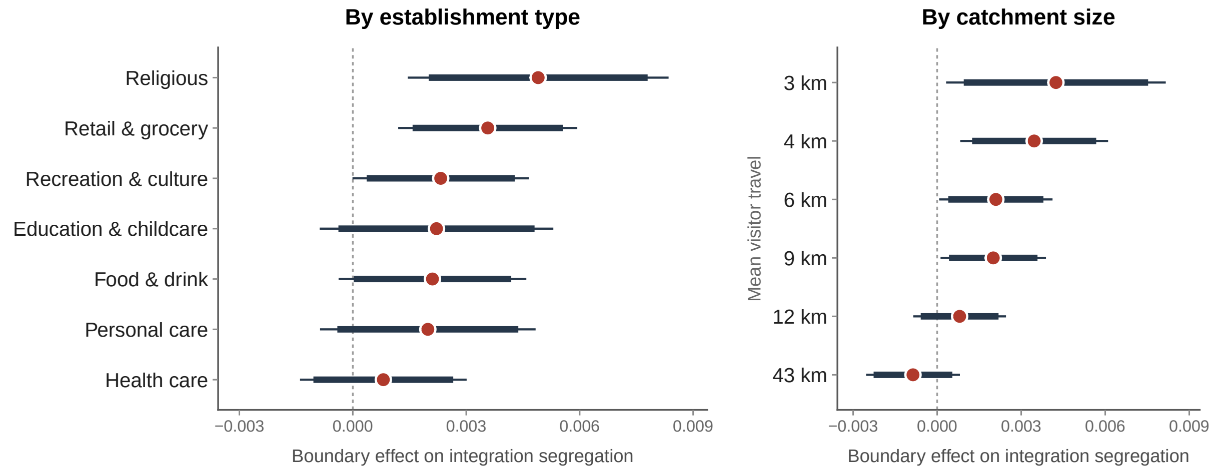

Figure 5: Where the boundary effect concentrates. *Panel A*: entropy-balanced boundary difference-in-discontinuity within establishment types, showing the seven types with the largest effects, ordered by size; the full set of types with estimates, and the categories set aside, are in Appendix G (Table G1). *Panel B*: the same estimate within six equal bins of POI catchment size (the median distance from home traveled by a location's visitors; local-draw locations at top). Points are estimates; thick and thin bars mark the 90% and 95% confidence intervals; estimates averaged over placebo-orientation draws.

space segregation is produced not only by the residents of its immediate neighborhood but by the wider set of nearby visitors who may know its racial character. In this sense, what separates the two ends is not the nominal category, but likely how locally a place draws its patrons—how far its routine patronage is woven into daily life where racialized preferences take hold—which the continuous catchment measure gauges directly below.

Catchment. Measuring the geographic reach of a location's clientele directly yields the same gradient (Figure 5B). We split POIs into six equal-sized bins by *catchment size*, defined in the Methods as the median distance from home traveled by a location's visitors and ranked to accommodate the heavy right-skew in trip length, and re-estimate the boundary difference-in-discontinuity within each. The effect falls monotonically as the draw widens, from +0.0042 ($t = 2.1$) for the most locally drawn sixth (mean catchment $\approx$ 2.5 km) to a slight, insignificant negative for the widest ($\approx$ 43 km), and the gap between the most- and least-local bins is significant ($p \approx 0.02$). As with our findings on POI-category heterogeneity, a location's visitor mix reflects the area its patrons are drawn from—a locally drawn clientele samples the redlining-shaped surroundings, while a regional establishment averages over the metropolis. We also show again that the gradient survives dropping each POI's own local residents (Appendix A), suggesting that the visitors who travel in are themselves racially sorted by the racial makeup of the place and its surrounding neighborhood—a sorting that fades as the draw widens and familiarity with the local racial context thins.

Robustness

Balanced boundaries: An Alternative Identification. A complementary design drops the comparison group entirely and leans on a feature of how the maps were drawn. HOLC rendered its grades as contiguous polygons over a city, and to close off a polygon's shape some boundary segments were run through terrain that was internally uniform—placing near-identical blocks on opposite sides of a grade line largely for cartographic convenience rather than because the two sides genuinely differed. We therefore follow Aaronson et al. (2021b) and isolate these segments empirically: for each real boundary we measure the pre-map (1910–1930) cross-border gap in the characteristics HOLC weighed—Black share, homeownership, house value, and foreign-born share—and keep the half whose gaps are near zero and, crucially, *flat* across all three pre-map censuses, so the two sides show no differential pre-trend before the maps. Because activity space segregation cannot be observed before 1930, this flat pre-map trajectory is the closest available analog to a parallel-trends test. On these balanced boundaries a direct two-sided comparison (no placebo group, no reweighting) returns +0.0049 ($p = 0.045$) on the most-balanced 40% of boundaries, similar in magnitude to the main estimate; on the less-balanced half, where the downgraded side was already the more disadvantaged and its gap was still widening toward 1930, the effect is more than twice as large (+0.011, $p < 0.001$). That a similar imprint appears even where neighborhoods began statistically alike is hard to reconcile with pre-existing sorting, and the gap between the two halves suggests redlining both created segregation at the margin and compounded it where disadvantage already existed (Appendix F).

Specification. The estimate is otherwise insensitive to the analytic choices behind it (Appendix A, Table A1). Balancing only the four 1930 covariates Aaronson et al. (2021b) emphasize, rather than the full set of fifteen historical controls, leaves it essentially unchanged (+0.0028 versus +0.0025); and because the main measure already corrects SafeGraph's device-panel bias by post-stratification reweighting of the visitor flows, we also recompute it from the raw, unweighted flows and find nearly identical estimates (Appendix I). Nor does it hinge on how visitor race is inferred from home neighborhoods: assigning each block group its largest racial group rather than its full distribution yields a proportionally larger effect on a coarser index, and relaxing through ecological-inference methods the assumption that groups within a block group travel alike leaves the effect intact (Appendix B).

Discussion and Conclusion

This article traces the racial segregation of everyday activity spaces—the visitor compositions assembled at roughly four million American establishments—to a bureaucratic act of classification—redlining—completed nearly a century before the mobility we observe. Descriptively, establishments in formerly C- and D-graded neighborhoods draw visitor pools that depart considerably more sharply from their metropolitan racial composition than those in higher-graded or unmapped areas, a national gradient that survives a rich vector of contemporary controls. We then move beyond the descriptive evidence, adapting the boundary difference-in-discontinuities design of Aaronson et al. (2021b) to show that, relative to reweighted same-grade lines carrying comparable observed pre-map covariate gaps, receiving the lower grade increased racial segregation (relative to metropolitan composition) in activity spaces at B–C and B–D borders by roughly six percent of the within-boundary standard deviation—an estimate that is robust across weighting schemes, spatial assignments, and measurement variants, and that reappears at boundaries whose two sides were statistically indistinguishable across every pre-map census. The stakes of this pattern follow from what segregated co-presence plausibly forecloses: access to organizations and information, cross-group exposure, and the shared institutions that might take root beyond the residential sphere (Chetty et al., 2022; Small, 2009). The reported effect, we stress, is a local estimate at the boundaries used in our design, not the total effect of redlining, which may be larger farther from those boundaries.

Our mediation evidence locates the principal pathway in neighborhood racial composition, which accounts for approximately 44% of the boundary effect and runs almost entirely along the Black–white divide—a pattern consonant with evidence that lower grading intensified residential sorting and that neighborhood racial composition shapes mobility over and above what travel time and income would predict (Davis et al., 2019; Faber, 2020; Vachuska, 2023). Neighborhoods left with a larger Black presence, in other words, draw proportionally fewer non-Black visitors, suggesting that the racialized preferences documented over residential choice (Charles, 2003; Krysan et al., 2009) extend to everyday destinations, and that the imprint we estimate may rest on the racialized reputation and stigma these places carry (Besbris et al., 2015), themselves downstream consequences of redlining. Importantly, the rival channels we can measure, which might otherwise displace this compositional account, receive no support: the boundary estimate is little changed when neighborhood income is accounted for; crowd-rated perceptions of place quality mediate nothing detectable; and the amenity mix itself contributes essentially none of the effect. This mediation exercise is nevertheless assumption-dependent rather than exhaustive: contemporary neighborhood racial composition may itself summarize decades of accumulated investment, institutional presence, reputation, and accessibility; and null estimates for the alternatives we can measure cannot exclude finer-grained economic or amenity channels operating below the resolution of our categories. What the compositional pathway does suggest, however, is the durability of racial prejudice previous research has long documented: the stigma borne by Black space and its avoidance by non-Black Americans (Anderson, 2015; Vachuska, 2023), now reaching beyond where people choose to live to where they are willing to visit by day.

For research on activity space segregation—which has established that movement beyond the home increases cross-group exposure without erasing sorting (Athey et al., 2021; Cai et al., 2024; Candipan et al., 2021; Moro et al., 2021)—the central implication concerns the origins of that residual sorting. Our contribution is to show that this sorting has an identifiable historical and institutional source: the accessibility constraints, income differences, and preferences through which contemporary scholarship accounts for experienced segregation operate within a residential and symbolic landscape that earlier classification helped produce. Activity space segregation, on this account, is neither independent of residential segregation nor reducible to it; it is one arena in which a historically structured residential order is re-enacted, daily, at the establishments where metropolitan populations assemble.

For the literature on redlining's legacy, the findings extend a body of design-based work linking lower grades to homeownership, property values, residential segregation, economic mobility, and a broader syndrome of neighborhood disadvantage (Aaronson et al., 2021a,b; Anders, 2023) to an outcome that this literature has largely overlooked: the racial clientele assembled at specific establishments. The study closest to ours, Aaronson et al. (2025), principally links residents' origin grades to the grades and socioeconomic characteristics of the neighborhoods they visit, alongside incoming-visit and city-level analyses that do not consider racial segregation in visitation patterns. We provide a complementary, place-level account of racial integration and carry the reach of institutional persistence from housing markets and household attainment into the everyday settings in which groups encounter—or fail to encounter—one another. By further showing that narrow-catchment, locally embedded establishments bear the largest imprints of the legacy, our results align with the relational accounts of urban segregation and racial homophily (Fiel, 2021) and reveal how historical exclusionary institutions created and reinforced such relational inequality.

None of this implies that HOLC single-handedly created the geography we observe. Discriminatory lending preceded the agency, the Federal Housing Administration conducted appraisals of its own, and lenders, real-estate professionals, and governments at every level jointly produced a racialized system of valuation and exclusion (Fishback et al., 2023, 2024; Markley, 2024); our boundary estimate accordingly cannot decompose the respective contributions of mortgage supply, private discrimination, stigma, and subsequent public and private investment. What the estimate does establish is of theoretical consequence in its own right: consistent with scholarship on the social power of classification (Fourcade and Healy, 2013; Scott, 1998), a visible and durable administrative label—applied once and long since repudiated—can coordinate expectations, validate unequal treatment, and become so embedded in demographic and investment trajectories that its imprint remains legible, far beyond its original scope of housing credit, in daily behavior generations after the originating program was dissolved.

We acknowledge and discuss several limitations that warrant caution in interpreting our results. First, our measures from SafeGraph capture co-presence of visitors rather than actual interaction in each place: we show which groups patronize the same establishments, not whether contact occurs there or with what quality—only that grading durably altered the demographic *preconditions* under which such contact *can* take place. A second set of limitations concerns the data themselves. SafeGraph is a nonprobability device panel, and although the post-stratification correction we apply aligns its geographic coverage with the census population, it cannot fully address selection among individuals within the same block group or recover visits the panel never observes (Li et al., 2024b). The establishment files are also known to contain category errors, duplicate listings, and establishments that have closed (Grigoropoulou and Small, 2025). We note, however, that because the outcome is computed from visitor origins rather than from the category label, and our main models condition on category fixed effects and weight each place by its log visit count, such errors are unlikely to drive the boundary contrast. Finally, visitor race is inferred from the composition of home neighborhoods rather than observed individually. We address the concern by examining alternative attributions in Appendix B—assigning each block group its largest group, and estimating group-specific visit propensities by ecological regression—both of which support the same conclusions.

Beyond the data, readers should be aware of the assumption behind our causal inference—that absent the maps, segregation would have evolved in parallel across the two sides of a border—and of the U.S.-centered scope of our results. The parallel-trends assumption cannot be tested in the outcome itself: activity space segregation is unobservable before the 2010s, and the flat pre-map covariate trajectories on which our designs rest are at best a plausible approximation to, not a substitute for, such a test; causal interpretation accordingly assumes that no unobserved discontinuity differentially generates the 2019 pattern. We partly address this concern with the balanced-boundary design of Appendix F, which confines the comparison to borders whose two sides show near-zero and flat pre-map gaps across all three censuses and returns an estimate at least as large as our main one. The estimates are also bounded in scope. Identified at American boundaries under an American racial order, they need not transfer to societies where ethnicity, migration, class, and governance organize space differently, and our null result for income segregation does not substitute for a genuinely intersectional measure that classifies visitors jointly by race and class—a design especially pertinent in settings where income rather than race is the primary axis of spatial sorting (Tammaru et al., 2016). Longitudinal mobility trajectories, richer destination records, and comparable institutional discontinuities outside the United States would address these limits directly, and we regard each as a well-defined direction for future research.

We close by insisting on where these conclusions assign responsibility: to institutions, not to the residents of historically graded neighborhoods. Segregation estimates can stigmatize the places they describe when they are read as attributes of local people or cultures rather than as sedimented outcomes of political and economic history. The persistence we document is neither an indictment of the communities that inhabit formerly redlined space nor a warrant for treating historical policy as destiny; it demonstrates, rather, that everyday mobility—for all the exposure it adds beyond the residential neighborhood—does not of itself dissolve inherited spatial inequality. Explaining when the daily round of destinations reproduces that inheritance and when it interrupts it is, we believe, essential to understanding how urban integration is made possible or denied.

Table 1: The effect of redlining on integration segregation, from naive association to causal boundary estimates. Dependent variable: integration segregation (Eq. 1), computed from the post-stratified 2019 origin flows (each home block group reweighted by its census-population-to-device ratio; Appendix I) with each visit carrying its home block group's fractional racial composition. Column 1 compares lower-graded (C/D) POIs with all others—grade-B areas and neighborhoods HOLC never mapped—net of MSA and category fixed effects *and* the extensive contemporary control vector (neighborhood income, education, majority race, walkability, and land-use regulation); before those controls the fixed-effects-only gap is 0.025, and the gradient withstands an Oster selection bound (see text). Columns 2–5 are boundary estimates. Standard errors (clustered by county in Column 1, by boundary in Columns 2–5) in parentheses. The reported coefficient is the lower-graded-side effect in Columns 1–2 and the difference-in-discontinuities in Columns 3–5.

	(1)	(2)	(3)	(4)	(5)
	OLS	Border	Difference-in-discontinuities		
	national	(no placebo)	IPW	Entropy bal.	EB, POI buffer
Lower-graded (C/D)	0.0196***	0.0030***	0.0030*	0.0025*	0.0036*
	(0.0025)	(0.0009)	(0.0014)	(0.0010)	(0.0015)
Placebo discontinuity	—	—	0.0000	0.0005	0.0011
POIs	3,246,415	105,118	1,072,020	1,072,020	327,771
Boundaries	—	998	9,738	9,738	17,161
MSA fixed effects	Yes				
Boundary fixed effects		Yes	Yes	Yes	Yes
Category fixed effects	Yes	Yes	Yes	Yes	Yes
Contemporary controls	Yes	No	No	No	No
Placebo reweighting	—	None	IPW	Entropy	Entropy
Buffer assignment	—	Tract	Tract	Tract	POI-level

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. The difference-in-discontinuities in Columns 3–5 is positive in 100% of 100 placebo-orientation draws (significant at the five-percent level in 66%, 85%, and 86% of draws, respectively). Observations are POI–boundary pairs: the tract-buffer columns (2–4) admit a POI to every boundary its tract abuts (about six on average), whereas the POI-level column (5) assigns each POI to its single nearest boundary within a quarter mile, so its count approximates the number of distinct POIs.

Appendix

A Robustness of the Boundary Estimate

Table A1 re-estimates the entropy-balanced difference-in-discontinuities under three alternatives to the baseline. The first balances only the four 1930 covariates Aaronson et al. (2021b) emphasize (Black share, homeownership, house value, foreign-born) rather than the full fifteen. The second narrows the buffer on each side of the border from a quarter mile to an eighth. The third divides the estimation sample by establishment type: a *daily-routine* group, read off the SafeGraph sub-category label and comprising groceries, supermarkets and convenience stores, restaurants and other food service, pharmacies and drug stores, general and department-store retail, and fuel; and its complement, the *less frequently visited* group, which holds every remaining establishment—religious congregations, health care, recreation and culture, personal services, finance, and lodging among them. The two groups are mutually exclusive and together exhaust the sample, with roughly 265,000 and 807,000 of its 1,072,020 POI–boundary observations. The split classifies establishments by what kind of place they are, a coarse proxy for how often any individual patronizes them; the log-visit weighting carried throughout is the more direct adjustment for visit frequency.

The estimate is stable across the covariate set and across both halves of the category split (+0.0025 and +0.0023), so the result does not rest on the less frequently visited establishments. At the eighth-mile buffer it attenuates to +0.0008 and is no longer distinguishable from zero. This is what the design's known blurring implies rather than evidence against the effect: because side assignment runs through whole 2010 tracts, which are far larger than the buffer itself, narrowing the buffer widens the mismatch between the tract that fixes a POI's side and the sliver of space the buffer defines, mixing the two sides further and pulling the discontinuity toward zero—the same mechanism that makes the POI-level assignment of Column 5 larger than the tract-based headline. The difference-in-discontinuities at the standard quarter-mile buffer remains significant and passes its placebo test.

Table A1: Robustness of the entropy-balanced boundary difference-in-discontinuities. Each estimate averages over placebo orientations, as in the main text; the baseline (all covariates) reproduces the headline estimate (Table 1, Column 4, +0.0025). The two category rows partition the estimation sample: daily-routine establishments (groceries, food service, pharmacies, general retail, fuel) and all remaining establishments.

Specification	DiD estimate	p
Baseline (all 15 historical covariates)	+0.0025	0.013
Key-4 1930 covariates only	+0.0028	0.007
Eighth-mile buffer (all covariates)	+0.0008	0.418
Daily-routine POI categories	+0.0025	0.049
Less-frequent POI categories	+0.0023	0.018

Excluding local residents. A location's own residents visit it disproportionately, so a measure built from all visitors could in principle inherit neighborhood composition mechanically rather than through any broader sorting of travel. To test this we recompute visitor race composition, and the segregation index, directly from the post-stratified monthly origin flows to 2.6 million establishments (those whose visitor origins carry home-panel coverage)—first for all visitors, then after dropping every visitor whose home block group (or home tract) is the POI's own. Local visitors are a small minority of traffic: on average 5.9% of a location's visits originate in its own block group and 10.7% in its own tract (medians 2.7% and 6.3%). Excluding them barely moves the segregation level—mean integration segregation falls only from 0.170 (all visitors) to 0.167 (excluding the home block group) and 0.165 (excluding the home tract). Re-running the entropy-balanced boundary difference-in-discontinuities on each version leaves the effect essentially intact: +0.0046 ($p = 0.01$) for all visitors, +0.0043 ($p = 0.02$) excluding home–block-group residents, and +0.0042 ($p = 0.02$) excluding home-tract residents (averaged over five orientations). The compositional mediation is likewise stable, with neighborhood racial composition carrying 38% of the excluding-home-block-group effect. Redlining's imprint on activity space segregation is therefore not an artifact of local residents patronizing local places; it reflects the racially sorted travel of visitors drawn from the surrounding metropolitan area.

The gradient survives, too. Excluding local residents leaves not only the pooled effect but its structure intact. Re-estimating the establishment-type and catchment gradients on the excluding-home-block-group measure reproduces the pattern reported in the main text (Table A2): the effect stays largest for neighborhood-tied types—religious congregations (+0.0082) and everyday retail and grocery (+0.0054), both $p < 0.05$—and for the most locally drawn establishments, and falls to zero or below for regionally drawn types and wide-catchment locations, tracking the all-visitor gradient almost exactly. Where the effect concentrates is thus a property of who travels to a place, not merely of who lives beside it.

B Inferring Visitor Race: Ecological-Inference Robustness

Because the Patterns data carry no demographic information on devices, visitor race is inferred from the racial composition of visitors' home block groups. This is an ecological inference in the classic sense (Goodman, 1953; King, 1997): the measure attributes to each visit the composition of its origin, and it would be biased if racial groups living in the same block group sorted differently across establishments: if, say, the Black residents of a mixed block group patronized one set of places and its white residents another. We address the concern in two steps, each recomputing the segregation index directly from the raw monthly origin flows to 3.3 million establishments—isolating the attribution choice; the post-stratification weights of Appendix I change these comparisons negligibly—and re-estimating the entropy-balanced boundary difference-in-discontinuities as in the main analysis (Table B1).

Table A2: Heterogeneity of the boundary effect with local residents excluded. Within-establishment-type and within-catchment-bin boundary difference-in-discontinuities, computed on the from-scratch visitor measure for all visitors and after dropping every visitor whose home census block group is the POI's own. Averaged over three placebo-orientation draws; significance stars refer to the excluded-residents column. $^{\dagger}p < 0.10$, $^{*}p < 0.05$.

	All visitors	Excl. own block group
By establishment type		
Religious	+0.0085	+0.0082*
Retail & grocery	+0.0058	+0.0054*
Education & childcare	+0.0050	+0.0042
Recreation & culture	+0.0049	+0.0045*
Food & drink	+0.0042	+0.0042†
Other retail	+0.0031	+0.0030
Personal & household services	+0.0027	+0.0026
Government & civic	+0.0025	+0.0012
Health care	+0.0021	+0.0014
Lodging & travel	+0.0009	+0.0014
Finance & insurance	−0.0010	−0.0013
Real estate & rental	−0.0025	−0.0026
By catchment size (mean visitor travel)		
≈ 2.4 km (most local)	+0.0073	+0.0063†
≈ 4.2 km	+0.0064	+0.0060*
≈ 5.9 km	+0.0039	+0.0041*
≈ 8.1 km	+0.0035	+0.0031†
≈ 11.2 km	+0.0022	+0.0017
≈ 37.5 km (widest)	−0.0010	−0.0009

Largest-group attribution. The main analysis attributes to each visit its origin block group's full racial distribution from the 2019 American Community Survey (Table B03002), so that a visit from a block group that is 60% Black and 40% white contributes six-tenths of a visitor to the Black share and four-tenths to the white share. A common simplification in this literature instead assigns each block group its largest racial group, a coarse coding under which within-neighborhood heterogeneity is set aside by construction. Because that coding removes the blending of groups within origins, the index's level roughly doubles, from a mean of 0.173 to 0.214 in the boundary sample, though the two versions are highly correlated across establishments ($r = 0.85$). Under the specification of this appendix the boundary effect is +0.0039 ($p = 0.01$) on the fractional measure against +0.0055 ($p = 0.02$) on the largest-group coding; relative to each measure's own mean the two are the same size (2.2% and 2.6%), so the coarser coding inflates the absolute effect without altering the proportional one—which is why the choice of attribution does not bear on any substantive conclusion.

Ecological regression. The preceding check retains the assumption that, within a block group, every racial group travels to a given establishment at the same per-capita rate, what the ecological-inference literature calls the neighborhood model. The classic alternative is Goodman's constancy model, the foundation of King (1997)'s estimator: each racial group has its own propensity to visit a given establishment, constant across the establishment's origin block groups, so that differential sorting by race within origins is permitted and is identified from variation across origins in how their racial makeup relates to their per-capita visit rates. We estimate this model establishment by establishment. For each establishment we regress the monthly per-capita visit rate of each origin block group (its visits to the establishment divided by its total outflow that month) on the block group's five racial shares, without intercept, weighting by outflow and constraining the propensities to be non-negative; the establishment's visitor composition is then the propensity of each group multiplied by that group's population at risk across its origins. The regression requires enough origins to be estimable, and we fit it for the 78% of establishments with at least twenty-four origin-months (2.6 million establishments), retaining the proportional composition for the rest. Because the regressions rest on modest numbers of origins, the resulting index is noisier and its level higher (mean 0.29), but this imprecision has no reason to differ systematically across the two sides of a grade line. The boundary effect on the ecological-regression measure is +0.0038 ($p = 0.04$), essentially identical to the proportional estimate, and +0.0035 ($p = 0.07$) when the analysis is confined to the establishments for which the model was fit. That the effect is the same size under the two polar assumptions about within-neighborhood travel indicates that redlining's imprint is not an artifact of inferring visitor race from home neighborhoods.[8]

Table B1: Boundary effect under alternative inferences of visitor race. To isolate the attribution choice, each row recomputes the integration-segregation index from the raw 2019 origin flows under the stated attribution of race to visitors (the main measure additionally applies the post-stratification weights of Appendix I, which leave these comparisons essentially unchanged) and re-estimates the entropy-balanced boundary difference-in-discontinuities with boundary and establishment-type fixed effects, weights $w_{\text{att}} \times \log(\text{visits} + 1)$, standard errors clustered by boundary, averaged over five placebo orientations.

Attribution of visitor race	Mean of index	DiD estimate	p
Full racial distribution of home block group (main attribution)	0.173	+0.0039	0.013
Largest group of home block group	0.214	+0.0055	0.016
Ecological regression (Goodman constancy model):			
all establishments	0.291	+0.0038	0.038
establishments with an estimable regression (78%)	0.312	+0.0035	0.072

C Benchmarking the Boundary Effect Against Economic Outcomes

To judge the size of the boundary effect on activity space segregation, we compare it against a common yardstick: redlining's effect, at the *same* boundaries and under the *identical*

[8] King (1997) refines the constancy model by combining it with the deterministic bounds of the method of bounds. At the establishment level those bounds are uninformative, because a single establishment absorbs a negligible share of any block group's race-specific outflow, so the regression estimator is the operative component of the approach here.

entropy-balanced difference-in-discontinuity, on the present-day (2010) neighborhood economic outcomes through which its long-run legacy has been documented (Appel, 2016; Krimmel, 2018). We draw median household income, gross rent, and the homeownership rate for each boundary tract from the same census panel used to construct the historical covariates, and estimate the discontinuity in each. Because these outcomes are measured on different scales, we express each effect relative to that outcome's own *within-boundary* standard deviation (its residual variation net of boundary and establishment-category fixed effects, the same yardstick used for segregation in the main text), so the magnitudes are directly comparable.

Table C1 reports the results. Redlining's discontinuity in activity space segregation (4–6% of a within-boundary standard deviation, depending on tract- versus POI-level assignment) is of the same order as its discontinuities in neighborhood income ($\approx 5\%$) and rent ($\approx 12\%$), and smaller than its discontinuity in homeownership ($\approx 19\%$), the most direct housing outcome. Every economic discontinuity is in the expected direction—the lower-graded side is poorer, with lower rents and less homeownership—consistent with redlining's documented economic imprint. In this entropy-balanced border sample, where the two sides begin observably similar, the income and rent gaps are estimated imprecisely, whereas the segregation and homeownership effects are statistically distinguishable from zero. The effect on activity space segregation is thus comparable in magnitude to redlining's imprint on the local economic conditions long central to segregation research, and—unlike the income and rent gaps—precisely identified.

Table C1: Redlining's boundary discontinuity in activity-space segregation versus present-day neighborhood economic outcomes, at the same HOLC boundaries. Each entry is the entropy-balanced difference-in-discontinuities (lower- minus higher-graded side, benchmarked against the counterfactual boundaries), averaged over placebo-orientation draws. "% of within-boundary SD" expresses the effect relative to that outcome's residual standard deviation net of boundary and category fixed effects.

Outcome	DiD estimate	% of within-bdy. SD	Precision
Activity space segregation (tract buffer)	+0.0025	3.9%	significant
Activity space segregation (POI buffer)	+0.0036	6.1%	significant
Median household income	−\$1,100	5.4%	n.s.
Gross rent	−1.0%	12.0%	marginal
Homeownership rate	−1.2 pp	18.9%	$p < 0.01$

D Selection on Unobservables: Oster Bounds

The national gradient (Table 1, Column 1) is descriptive, but its stability under an extensive contemporary control vector is informative about how much unobserved confounding would be required to overturn it. We formalize this with the bounding procedure of Oster (2019), which sharpens the familiar "coefficient stability" intuition by pairing the movement of the coefficient with the movement of the R^2 as controls are added. Let $\mathring{\beta}$ and $\mathring{R}$ be the coefficient and R^2 from the model with fixed effects only, and $\tilde{\beta}$ and $\tilde{R}$ those from the fully controlled model. Assuming selection on unobservables is proportional (by a factor δ) to selection on observables, and that a hypothetical regression on *all* determinants, observed and unobserved, would attain $R^2 = R_{\max}$, the statistic δ^* is the value of δ that would drive the coefficient to zero:

$$\delta^* \;=\; \frac{\tilde{\beta}\,(\tilde{R} - \mathring{R})}{(\mathring{\beta} - \tilde{\beta})\,(R_{\max} - \tilde{R})}. \tag{6}$$

A value $\delta^* > 1$ means unobservables would have to be *more* important than the observed controls to explain the effect away; the larger δ^*, the more robust the estimate. Following Oster (2019) we set $R_{\max} = 1.3\,\tilde{R}$ (the benchmark under which roughly 90% of results from randomized designs survive) and also report the far stricter $R_{\max} = 1$ for reference. The controls are the full contemporary vector described in the main text (neighborhood income, education, racial composition, land-use regulation, and the remaining community characteristics), with the sample held fixed across the restricted and full models.

Table D1 reports the inputs and δ^*. For *integration* segregation the coefficient barely moves as the rich control set is added ($+0.025 \rightarrow +0.020$) while the R^2 rises only modestly, yielding $\delta^* = 2.67$: selection on unobservables would have to be more than two and a half times as strong as selection on our observed controls, themselves an unusually complete set, to reduce the gradient to zero; under the punitive $R_{\max} = 1$ the statistic is 0.47. For the *diversity* benchmark the picture reverses: the coefficient more than halves when controls are added, giving $\delta^* = 0.30$, far below one: the diversity gradient is fragile, consistent with the null boundary estimate for that measure (Appendix E). We stress that these bounds concern the *descriptive* association: because several controls are themselves post-treatment, a large δ^* shows only that the gradient is not an artifact of the observed differences between neighborhoods, not that it is causal. Causal identification rests on the boundary design.

Table D1: Oster (2019) selection bounds for the national gradient.

Measure	$\mathring{\beta}$ ($\mathring{R}$)	$\tilde{\beta}$ ($\tilde{R}$)	$R_{\max}$=1.3$\tilde{R}$	δ^*	δ^* ($R_{\max}$=1)
Integration	+0.025 (0.29)	+0.020 (0.37)	0.48	2.67	0.47
Diversity	−0.084 (0.56)	−0.032 (0.66)	0.86	0.30	0.17

E The Diversity Benchmark

Our focal measure benchmarks a location's visitor composition against its *metropolitan* racial composition. A natural alternative benchmarks it against *equal* group shares, replacing δ_i^m in Eq. (1) with $\delta_i = 1/5$. Following segregation research, we call the result *deviation from ideal diversity*: it measures distance from maximal group balance rather than from the population actually present in the metro. The two benchmarks answer different questions and can diverge sharply—a majority-Black location in a majority-white metropolis sits close to even shares (high diversity) yet far from its metro composition (low integration).

Redlining's effect is specific to the integration benchmark. Descriptively, redlined POIs are if anything *closer* to even group shares than still-desirable ones (a raw gradient of −0.084

net of metropolitan and category fixed effects), and this descriptive gap is not robust to selection on unobservables (Oster $\delta^* = 0.30$, far below the $\delta^* > 1$ threshold; Appendix D). Applying the entropy-balanced boundary difference-in-discontinuities to the diversity measure returns a precise null: the difference-in-discontinuities is 0.0000 (median $p = 0.77$; significant in none of the 100 placebo-orientation draws), against the +0.0025 estimated for integration segregation. The raw discontinuity at real boundaries is itself only a trivial −0.0009—redlined sides draw visitor pools marginally *closer* to even group shares—and once benchmarked against counterfactual boundaries it is indistinguishable from zero. Substantively, redlining does not make activity spaces less diverse in the abstract; it pulls them away from the racial mix their metropolitan population would imply. This specificity is itself informative: the policy's imprint operates through displacement relative to the metropolitan population—the quantity classic segregation indices track—not through any general loss of group balance.

F A Second Identification: Balanced Boundaries

Our main design builds a counterfactual comparison group. A second, independent strategy needs none. If two adjacent neighborhoods were statistically identical before the maps, then which side HOLC graded lower cannot reflect a pre-existing difference between them, and the present-day gap in segregation across the line is a direct estimate of the grading's effect. We therefore isolate the boundaries at which grading was, in effect, random, and compare their two sides directly—the balanced-boundary logic Aaronson et al. (2021b) apply to individual economic outcomes.

For each real B↔C/D boundary we measure the pre-map (1930) cross-border gap (lower- minus higher-graded side) in the characteristics that shaped HOLC's assessments: Black share, homeownership, median house value, and foreign-born share. Boundaries whose gaps are near zero are those where the two sides were indistinguishable before grading. Among the boundaries with complete 1930 records (960 of the 2,138), we take the half with the smallest such gaps and, on their visitors, estimate

$$y_{kb}^{m} = \beta\, D_{kb} + \alpha_b + \theta_{\ell(k)} + \epsilon_{kb},$$

where D_{kb} again marks the lower-graded side and α_b, $\theta_{\ell(k)}$ are boundary and establishment-category fixed effects. POIs are assigned to a boundary by their own distance to it (within a quarter mile); no placebo or reweighting enters, so identification rests entirely on the balance of the selected boundaries.

Table F2 shows the selection succeeds: across the balanced boundaries, every pre-map gap is an order of magnitude smaller than in the full sample and negligible, with a 1930 Black-share gap of +0.001 against +0.012 on the less-balanced half. On the most-balanced 40% of boundaries the effect is $\hat{\beta} = +0.0049$ ($p = 0.045$), close to the main estimate; nearby cutoffs give +0.004 to +0.005. On the less-balanced half, where the lower-graded side was already the more disadvantaged, it is more than twice as large and highly significant (+0.011, $p < 0.001$).

Crucially, this balance is not confined to 1930. On the balanced boundaries the cross-border gap is near zero and *flat* across all three pre-map censuses, so the two sides show no differential pre-trend before the maps (Table F1); on the less-balanced half the gaps are sizable and, for race and foreign-born share, already *widening* through 1930. Because activity space segregation is unobservable before 1930, this flat pre-map trajectory in the balanced sample is the closest available analog to a parallel-trends test, and mirrors the "no pre-trend" that Aaronson et al. (2021b) report for their low-propensity borders.

Table F1: Pre-map trajectory of the cross-border gap (lower- minus higher-graded side), 1910–1930, by balance stratum. On the balanced boundaries every gap is near zero and flat; on the less-balanced boundaries the racial and foreign-born gaps widen toward 1930.

Within-boundary gap	1910	1920	1930
Balanced half			
Black share	+0.000	−0.000	+0.000
Homeownership	−0.001	+0.000	−0.001
Foreign-born share	+0.001	+0.001	+0.000
Less-balanced half			
Black share	+0.008	+0.006	+0.012
Homeownership	−0.042	−0.041	−0.038
Foreign-born share	+0.013	+0.011	+0.017

The two results together sharpen the interpretation. Redlining raised activity space segregation even where the two sides began alike—which pre-existing sorting cannot explain—and it did so more strongly where disadvantage was already present, consistent with a policy that both created inequality at the margin and compounded it where it existed.

Table F2: Balanced-boundary identification. Top panel: mean pre-map (1930) cross-border gap (lower- minus higher-graded side) by balance stratum. Bottom panel: the direct across-boundary estimate on each stratum (boundary and category fixed effects). The balanced column keeps the 40% of real borders with the smallest standardized pre-map gaps; the less-balanced column is the half with the largest. A cutoff sweep appears in the appendix text.

	All boundaries	Most-balanced 40%	Less-balanced half
Pre-map cross-border gap, 1930			
Black share	+0.006	+0.001	+0.012
Foreign-born share	+0.008	+0.001	+0.014
Homeownership rate	−0.027	−0.006	−0.045
Median house value ($)	−1,079	−130	−1,978
Redlining effect $\hat{\beta}$	+0.0045	+0.0049	+0.0106
p-value	< 0.001	0.045	< 0.001
Boundaries	2,138	384	480

G Establishment-Type Grouping

The establishment-type heterogeneity (Figure 5A) sorts POIs into broad types by matching keywords against their SafeGraph top-level category label. Table G1 gives the full mapping together with each type's within-type boundary difference-in-discontinuity; the seven types plotted in the main figure are those with the largest effects. Across the ranking the effect is largest for establishments used close to home (religious congregations, groceries and pharmacies) and falls to zero or below for those whose clientele is regional or institutional (specialty health care, finance, lodging, real estate). We treat these type contrasts as a coarse illustration and rely on the continuous catchment measure, which reads localness directly off visitor travel, for the sharper test. The grouping is coarse: SafeGraph's top-level category is a broad four-digit label, so the types are wide: some establishments in a "local" type (department stores, universities, hospitals) draw regionally. We report a type only where it retains enough boundary coverage to estimate (at least 200 POI–boundary observations with both graded sides present), and we exclude three kinds of establishment rather than force them into a type: automotive and fuel services (gasoline, repair, parts), whose routine local use is mixed with highway through-traffic; automobile dealers, which draw regionally; and a residual "not classified" group dominated by non-public, business-to-business establishments (wholesalers, manufacturing, contractors, freight, telecommunications carriers, professional and consulting offices) that are not visit destinations in the sense the analysis requires.

H Perceived Place Quality (Place Pulse)

To probe whether the boundary effect runs through the perceived quality of places rather than through who lives near them, we draw on the Place Pulse 2.0 collection, in which online participants made pairwise comparisons of Google Street View images on six perceptual dimensions: how safe, wealthy, beautiful, lively, depressing, or boring a scene appears. Because each rated image is a geolocated street scene rather than a point of interest, the ratings must be matched to POIs spatially. We first collapse the six dimensions into a single quality score for each rated location (the mean of the standardized safe, wealthy, beautiful, and lively ratings, net of the standardized depressing and boring ratings) and then assign each boundary POI the score of its nearest rated image, keeping only matches within a distance threshold (500 m in the baseline). Boundaries are retained when they have matched POIs on both sides, and the entropy-balancing weights are recomputed within this subsample. Place Pulse covers only a limited set of U.S. cities, so most boundary POIs lie far from any rated image (the median nearest-image distance is about 3 km), and the analysis is confined to the roughly 290 boundaries (under a third of the national sample) with an image nearby. Among the POIs that do match, however, the image is close: the median POI-to-image distance is about 160 m.

Entered as a mediator of the boundary effect on integration segregation, the composite quality score carries essentially none of it (0–2%), and the redlining coefficient is left essentially unchanged when it is included. The same holds dimension by dimension: entering each of the six perceptual ratings on its own, none mediates more than about one to two percent of the boundary effect, and redlining's effect on each is statistically indistinguishable from

Table G1: SafeGraph top-level categories mapped to establishment types, with the number of distinct POIs in the boundary estimation sample and each type's within-type boundary difference-in-discontinuity. Seven broad types are plotted in Figure 5A. Stratified within-type estimates, averaged over placebo-orientation draws; $^{\dagger}p < 0.10$, $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Type	SafeGraph top-level categories	POIs	DiD
Religious	Religious organizations	17,826	$+0.0048^{**}$
Retail & grocery	Grocery, supermarkets, convenience, specialty food; pharmacies and drug stores; health/personal-care stores; clothing; department and general merchandise; shoe, book, hardware, furniture, florist, jewelry	27,893	$+0.0033^{*}$
Recreation & culture	Amusement and recreation; museums and historical sites; fitness; golf; bowling; arts and performing arts; gambling; zoos; nature parks	13,069	$+0.0023^{\dagger}$
Education & childcare	Elementary and secondary schools; child day-care; colleges and universities; educational support	13,194	+0.0021
Food & drink	Restaurants and other eating places; drinking places; snack and beverage bars; caterers	49,761	+0.0021
Other retail	Sporting-goods, hobby, and music stores; electronics and appliances; building materials and garden supply; miscellaneous retailers	12,977	$+0.0018^{\dagger}$
Personal & household services	Personal-care services; beauty salons; barber shops; nail salons; hair; spas; laundry and dry cleaning; goods repair; veterinary; death care	10,788	+0.0018
Government & civic	Justice and public order; public administration; postal service; social assistance	1,759	+0.0014
Health care	Physicians, dentists, other practitioners; hospitals; nursing and residential care; outpatient, mental-health, chiropractic, optometric, diagnostic	21,675	+0.0009
Lodging & travel	Traveler accommodation; travel arrangement	1,677	+0.0002
Finance & insurance	Depository credit (bank branches); insurance agencies; other financial services	5,936	−0.0002
Real estate & rental	Real-estate lessors and agents; rental and leasing	3,088	−0.0010

zero (the largest, a marginally lower perceived-safety rating on redlined sides, still transmits about one percent). No facet of how these places *look* carries the effect.

The null is not an artifact of the matching radius. Repeating the analysis at thresholds of 250, 500, 750, and 1,000 m (which vary the matched sample from about 80,000 to 102,000 POIs across 286 to 298 boundaries, with median match distances of 140–160 m) leaves perceived quality mediating essentially none of the boundary effect at every threshold (between −1% and 0%).

Three limitations temper this null. First, the perceptual ratings are a coarse proxy for amenity quality, filtered through raters' own associations (including, potentially, race), so they may absorb part of the very neighborhood composition we treat as the mechanism. Second, coverage is partial and urban, and the images postdate the historical grading by decades. Third, a null on a limited subsample cannot exclude a quality channel that a larger or more precisely measured sample might detect. We therefore read this analysis as corroborating, not decisive: it is consistent with an effect that operates through neighborhood composition rather than perceived place quality, but it does not on its own settle the question.

I Sampling Bias and Post-Stratification

The SafeGraph device panel covers roughly ten percent of the population but over-represents some neighborhoods. The main measure therefore applies SafeGraph's recommended correction, described in the Data section: each home census block group is weighted by the ratio of its census population share to its device share before visitor compositions are formed. The resulting adjustment factors have an interquartile range of about 0.86 to 1.43 (median 1.12), so a typical over-represented block group is down-weighted by roughly ten percent. The correction is immaterial to every conclusion: recomputing every visitor composition from the raw, unweighted flows yields a nearly identical index (the two versions correlate at 0.98), and re-estimating the full set of models on the raw-flow version reproduces each result with marginally smaller boundary estimates (for example, $+0.0035$ versus the reported $+0.0036$ in the POI-buffer design). The correction addresses coverage imbalance across block groups; it cannot, however, recover any selection among individuals within the same block group, a limit we note in the Discussion.

J Zoning and Gentrification

Redlining is the paper's focal policy, but two other exclusionary levers appear in our data, and Table J1 reports their descriptive associations with integration segregation—estimated at the POI level with MSA and establishment-category fixed effects and county-clustered standard errors, with and without the contemporary control vector of the main text. Neither lever admits the fixed historical boundary that identifies the redlining effect, so we present associations, not effects.

Zoning. Locations in counties with more restrictive present-day land-use regulation (the 2018 Wharton Residential Land-Use Regulatory Index) draw visitor pools slightly *closer* to their metropolitan composition: a one-standard-deviation increase in restrictiveness is associated with -0.007 integration segregation net of controls, about nine percent of the pooled standard deviation. The index varies at the county level and measures contemporary regulation rather than any historical treatment, so we draw no causal conclusion from this association.

Gentrification. Following the gentrifiability design of Hwang and Ding (2020), POI neighborhoods are classed as gentrifying, gentrifiable but nongentrifying, or nongentrifiable. The raw contrasts track baseline disadvantage—relative to the disadvantaged gentrifiable baseline, both gentrifying and nongentrifiable neighborhoods draw visitor pools closer to the metropolitan composition—and once the contemporary controls enter, the gentrifying–nongentrifying gap disappears entirely (-0.0009, $p = 0.53$) while the nongentrifiable contrast shrinks to a modest -0.006. What the data reveal, in short, is the compositional baseline rather than any distinct imprint of gentrification; the estimates are thin precisely because gentrification is a continuous, contemporary process with no discontinuity to exploit. A composition-conditional, activity-space analysis of gentrification in the spirit of Hwang and Ding (2020), and causal designs built on historical zoning maps (Shertzer et al., 2016), remain well-defined directions for future work.

Table J1: Descriptive associations of contemporary exclusionary levers with integration segregation. OLS at the POI level with MSA and establishment-category fixed effects; the second column adds the contemporary control vector of Table 1 (neighborhood income, education, age structure, foreign-born share, transit, walkability, highway proximity, and majority-race community). Gentrification classes follow the gentrifiability design of Hwang and Ding (2020); the omitted category is gentrifiable but nongentrifying neighborhoods. Standard errors, clustered by county, in parentheses. $^{***}p < 0.001$.

	FE only	+ contemporary controls
Land-use restrictiveness (WRLURI 2018, per SD)	-0.0090^{***}	-0.0074^{***}
	(0.0023)	(0.0019)
Gentrifying	-0.0152^{***}	-0.0009
	(0.0020)	(0.0014)
Nongentrifiable	-0.0213^{***}	-0.0056^{***}
	(0.0028)	(0.0010)
POIs (zoning / gentrification)	2,569,871 / 3,263,829	

References

Aaronson, D., Faber, J., Hartley, D., Mazumder, B., and Sharkey, P. (2021a). The long-run effects of the 1930s HOLC "redlining" maps on place-based measures of economic opportunity and socioeconomic success. *Regional Science and Urban Economics*, 86:103622.

Aaronson, D., Han, J. K., Hartley, D., and Mazumder, B. (2025). The lasting impact of historical residential security maps on experienced segregation. *Real Estate Economics*, 53(1):134–163.

Aaronson, D., Hartley, D., and Mazumder, B. (2021b). The effects of the 1930s HOLC "redlining" maps. *American Economic Journal: Economic Policy*, 13(4):355–392.

Aaronson, D., Hartley, D., Mazumder, B., and Stinson, M. (2023). The long-run effects of the 1930s redlining maps on children. *Journal of Economic Literature*, 61(3):846–862.

Allport, G. W. (1954). *The Nature of Prejudice*. Addison-Wesley, Cambridge, MA.

Anders, J. (2023). The long run effects of de jure discrimination in the credit market: How redlining increased crime. *Journal of Public Economics*, 222:104857.

Anderson, E. (2015). The white space. *Sociology of Race and Ethnicity*, 1(1):10–21.

Appel, I. (2016). Pockets of poverty: The long-term effects of redlining. *SSRN Electronic Journal*.

Athey, S., Ferguson, B., Gentzkow, M., and Schmidt, T. (2021). Estimating experienced racial segregation in US cities using large-scale GPS data. *Proceedings of the National Academy of Sciences*, 118(46):e2026160118.

Baker, R. S. (2022). The historical racial regime and racial inequality in poverty in the american south. *American Journal of Sociology*, 127(6):1721–1781.

Baldenius, T., Klauber, H., and Koch, N. (2025). Heat and experienced racial segregation. *Journal of Environmental Economics and Management*, 133:103210.

Baron, R. M. and Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6):1173–1182.

Bayer, P., Ferreira, F., and McMillan, R. (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4):588–638.

Besbris, M., Faber, J. W., Rich, P., and Sharkey, P. (2015). Effect of neighborhood stigma on economic transactions. *Proceedings of the National Academy of Sciences*, 112(16):4994–4998.

Black, S. E. (1999). Do better schools matter? parental valuation of elementary education. *The Quarterly Journal of Economics*, 114(2):577–599.

Browning, C. R., Calder, C. A., Boettner, B., Tarrence, J., Khan, K., Soller, B., and Ford, J. L. (2021). Neighborhoods, activity spaces, and the span of adolescent exposures. *American sociological review*, 86(2):201–233.

Cagney, K. A., Cornwell, E. Y., Goldman, A. W., and Cai, L. (2020). Urban mobility and activity space. *Annual Review of Sociology*, 46(1):623–648.

Cai, L., Browning, C. R., and Cagney, K. A. (2024). Exposure of neighborhood racial and socio-economic composition in activity space: A new approach adjusting for residential conditions. *Social Forces*, 103(1):350–373.

Candipan, J., Phillips, N. E., Sampson, R. J., and Small, M. (2021). From residence to movement: The nature of racial segregation in everyday urban mobility. *Urban Studies*, 58(15):3095–3117.

Chang, S., Pierson, E., Koh, P. W., Gerardin, J., Redbird, B., Grusky, D., and Leskovec, J. (2021). Mobility network models of COVID-19 explain inequities and inform reopening. *Nature*, 589(7840):82–87.

Charles, C. Z. (2003). The dynamics of racial residential segregation. *Annual Review of Sociology*, 29(1):167–207.

Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., Johnston, D., Koenen, M., Laguna-Muggenburg, E., Mudekereza, F., Rutter, T., Thor, N., Townsend, W., Zhang, R., Bailey, M., Barberá, P., Bhole, M., and Wernerfelt, N. (2022). Social capital I: Measurement and associations with economic mobility. *Nature*, 608(7921):108–121.

Davis, D. R., Dingel, J. I., Monras, J., and Morales, E. (2019). How segregated is urban consumption? *Journal of Political Economy*, 127(4):1684–1738.

Dougherty, K. D. (2003). How monochromatic is church membership? racial–ethnic diversity in religious community. *Sociology of Religion*, 64(1):65–85.

Elbers, B. (2024). Explaining changes in US residential segregation through patterns of population change. *Nature Cities*, 1(3):194–204.

Emerson, M. O. and Smith, C. (2000). *Divided by Faith: Evangelical Religion and the Problem of Race in America*. Oxford University Press, New York.

Faber, J. W. (2020). We built this: Consequences of new deal era intervention in america's racial geography. *American Sociological Review*, 85(5):739–775.

Fiel, J. E. (2021). Relational segregation: A structural view of categorical relations. *Sociological Theory*, 39(3):153–179.

Fishback, P. V., LaVoice, J., Shertzer, A., and Walsh, R. (2023). The HOLC maps: How race and poverty influenced real estate professionals' evaluation of lending risk in the 1930s. *Journal of Economic History*, 83(4):1019–1056.

Fishback, P. V., Rose, J., Snowden, K. A., and Storrs, T. (2024). New evidence on redlining by federal housing programs in the 1930s. *Journal of Urban Economics*, 141:103462.

Fourcade, M. and Healy, K. (2013). Classification situations: Life-chances in the neoliberal era. *Accounting, Organizations and Society*, 38(8):559–572.

Goodman, L. A. (1953). Ecological regressions and behavior of individuals. *American Sociological Review*, 18(6):663–664.

Gordon, M. M. (1964). *Assimilation in American Life: The Role of Race, Religion, and National Origins*. Oxford University Press, New York.

Grigoropoulou, N. and Small, M. L. (2025). Are large-scale data from private companies reliable? an analysis of machine-generated business location data in a popular dataset. *Social Science Computer Review*, 43(5):1087–1113.

Gyourko, J., Saiz, A., and Summers, A. (2008). A new measure of the local regulatory environment for housing markets: The Wharton Residential Land Use Regulatory Index. *Urban Studies*, 45(3):693–729.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, 20(1):25–46.

Heine, C., Abbiasov, T., Santi, P., and Ratti, C. (2025). The role of urban amenities in facilitating social mixing: Evidence from stockholm. *Landscape and Urban Planning*, 254:105250.

Hillier, A. E. (2003). Redlining and the Home Owners' Loan Corporation. *Journal of Urban History*, 29(4):394–420.

Hillier, A. E. (2005). Residential security maps and neighborhood appraisals: the home owners' loan corporation and the case of philadelphia. *Social Science History*, 29(2):207–233.

Hołubowska, O. and Poorthuis, A. (2025). Examining the impact of daily urban activity on spatial segregation: Disparities in the proportion of foreign-born residents across residential areas and visited locations. *Journal of Ethnic and Migration Studies*, 51(1):324–345.

Hwang, J. and Ding, L. (2020). Unequal displacement: Gentrification, racial stratification, and residential destinations in philadelphia. *American Journal of Sociology*, 126(2):354–406.

Järv, O., Müürisepp, K., Ahas, R., Derudder, B., and Witlox, F. (2015). Ethnic differences in activity spaces as a characteristic of segregation: A study based on mobile phone usage in Tallinn, Estonia. *Urban Studies*, 52(14):2680–2698.

Jones, M. and Pebley, A. R. (2014). Redefining neighborhoods using common destinations: Social characteristics of activity spaces and home census tracts compared. *Demography*, 51(3):727–752.

Kang, Y., Gao, S., Liang, Y., Li, M., Rao, J., and Kruse, J. (2020). Multiscale dynamic human mobility flow dataset in the U.S. during the COVID-19 epidemic. *Scientific Data*, 7(1):390.

King, G. (1997). *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*. Princeton University Press, Princeton, NJ.

Klinenberg, E. (2018). *Palaces for the People: How Social Infrastructure Can Help Fight Inequality, Polarization, and the Decline of Civic Life*. Crown, New York.

Krimmel, J. (2018). Persistence of prejudice: Estimating the long-term effects of redlining. *SocArXiv preprint*.

Krivo, L. J., Washington, H. M., Peterson, R. D., Browning, C. R., Calder, C. A., and Kwan, M.-P. (2013). Social isolation of disadvantage and advantage: The reproduction of inequality in urban space. *Social Forces*, 92(1):141–164.

Krysan, M., Couper, M. P., Farley, R., and Forman, T. A. (2009). Does race matter in neighborhood preferences? results from a video experiment. *American Journal of Sociology*, 115(2):527–559.

Kwan, M.-P. (2012). The uncertain geographic context problem. *Annals of the Association of American Geographers*, 102(5):958–968.

Li, Y., Huang, Y., Li, R., and Zhang, K. (2024a). Historical redlining and park use during the COVID-19 pandemic: Evidence from big mobility data. *Journal of Exposure Science & Environmental Epidemiology*, 34(3):399–406.

Li, Z., Ning, H., Jing, F., and Lessani, M. N. (2024b). Understanding the bias of mobile location data across spatial scales and over time: A comprehensive analysis of SafeGraph data in the United States. *PLOS ONE*, 19(1):e0294430.

Loo, B. P., Fan, Z., and Moro, E. (2024). Residential and experienced social segregation: the roles of different transport modes, metro extensions, and longitudinal changes in hong kong. *Humanities and Social Sciences Communications*, 11(1):1439.

Markley, S. (2024). Federal "redlining" maps: A critical reappraisal. *Urban Studies*, 61(2):195–213.

Massey, D. S. and Denton, N. A. (1988). The dimensions of residential segregation. *Social Forces*, 67(2):281–315.

Massey, D. S. and Denton, N. A. (1993). *American apartheid: Segregation and the making of the underclass*. Harvard University Press, Cambridge, MA.

Moro, E., Calacci, D., Dong, X., and Pentland, A. (2021). Mobility patterns are associated with experienced income segregation in large US cities. *Nature Communications*, 12(1):4633.

Müürisepp, K., Järv, O., Tammaru, T., and Toivonen, T. (2022). Activity spaces and big data sources in segregation research: A methodological review. *Frontiers in Sustainable Cities*, 4:861640.

Nelson, R. K., Winling, L., et al. (2023). Mapping inequality: Redlining in new deal america. American Panorama: An Atlas of United States History, Digital Scholarship Lab, University of Richmond. https://dsl.richmond.edu/panorama/redlining/.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Phillips, N. E., Levy, B. L., Sampson, R. J., Small, M. L., and Wang, R. Q. (2021). The social integration of american cities: Network measures of connectedness based on everyday mobility across neighborhoods. *Sociological Methods & Research*, 50(3):1110–1149.

Pinchak, N. P., Browning, C. R., Calder, C. A., and Boettner, B. (2022). Racial inequalities in adolescents' exposure to racial and socioeconomic segregation, collective efficacy, and violence. *Demography*, 59(5):1763–1789.

Reardon, S. F. and Firebaugh, G. (2002). Measures of multigroup segregation. *Sociological Methodology*, 32(1):33–67.

Rieger, S., Baldassarri, D., and Abascal, M. (2026). Diversity and prosociality in NYC neighborhoods: Evidence from a lost wallet experiment. *American Journal of Sociology*. Advance online publication. https://doi.org/10.1086/741818.

Rigby, D., Esposito, M. H., Brown, T. H., Lee, H., and Clark, S. (2026). Sociohistorical contexts of racial violence: Sundown towns and the durability of racialized public space. *Ethnic and Racial Studies*, 49(2):480–495.

Salazar-Miranda, A., Conzelmann, C., Phan, T., and Hoffman, J. (2024). Long-term effects of redlining on climate risk exposure. *Nature Cities*, 1(6):436–444.

Schnell, I. and Haj-Yahya, N. (2014). Arab integration in Jewish-Israeli social space: Does commuting make a difference? *Urban Geography*, 35(7):1084–1104.

Scott, J. C. (1998). *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed*. Yale University Press, New Haven, CT.

Sharkey, P. and Faber, J. W. (2014). Where, when, why, and for whom do residential contexts matter? moving away from the dichotomous understanding of neighborhood effects. *Annual Review of Sociology*, 40(1):559–579.

Shertzer, A., Twinam, T., and Walsh, R. P. (2016). Race, ethnicity, and discriminatory zoning. *American Economic Journal: Applied Economics*, 8(3):217–246.

Silm, S. and Ahas, R. (2014). The temporal variation of ethnic segregation in a city: Evidence from a mobile phone use dataset. *Social Science Research*, 47:30–43.

Small, M. L. (2009). *Unanticipated gains: Origins of network inequality in everyday life*. Oxford University Press, New York.

Small, M. L., Akhavan, A., Torres, M., and Wang, Q. (2021). Banks, alternative institutions and the spatial–temporal ecology of racial inequality in US cities. *Nature Human Behaviour*, 5(12):1622–1628.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology*, 13:290–312.

Tammaru, T., Marcińczak, S., van Ham, M., and Musterd, S., editors (2016). *Socio-Economic Segregation in European Capital Cities: East Meets West*. Routledge, London.

Vachuska, K. (2023). Racial segregation in everyday mobility patterns: Disentangling the effect of travel time. *Socius: Sociological Research for a Dynamic World*, 9:23780231231169261.

Wang, D., Li, F., and Chai, Y. (2012). Activity spaces and sociospatial segregation in Beijing. *Urban Geography*, 33(2):256–277.

Winling, L. C. and Michney, T. M. (2021). The roots of redlining: Academic, governmental, and professional networks in the making of the new deal lending regime. *Journal of American History*, 108(1):42–69.

Wong, D. W. S. and Shaw, S.-L. (2011). Measuring segregation: An activity space approach. *Journal of Geographical Systems*, 13(2):127–145.

Woods, L. L. (2012). The federal home loan bank board, redlining, and the national proliferation of racial lending discrimination, 1921–1950. *Journal of Urban History*, 38(6):1036–1059.

Xu, W., Wang, Z., Attia, N., Attia, Y., Zhang, Y., and Zong, H. (2024). An experienced racial-ethnic diversity dataset in the United States using human mobility data. *Scientific Data*, 11(1):638.