The (Express)Way to Segregation: Evidence from Chicago*

Sara Bagagli†

February 2025

Click here for the latest version

Abstract

How do man-made barriers influence racial segregation within cities? I study the long-term effects of expressway construction on racial segregation. These multilane roads (i) produce a local shock to residential amenities and (ii) divide the areas they cross through, creating local barriers to the interaction of nearby communities. I find that expressways affect within-city racial segregation through two main channels. First, a price or disamenity channel: Racial segregation increases because of income differences between Black and white residents, which on average lead the two groups to react differently to changes in the quality of life near expressways. Second, a physical barrier channel: Following expressway construction, racial sorting is affected by changes in accessibility to different parts of the city and, in turn, to neighborhoods with distinct demographic compositions. Motivated by these findings, I develop a structural urban model with racial preference spillovers across locations. The model estimates racial preference parameters and undertakes counterfactual experiments to inform current public policies targeting the social issues of transport infrastructures in US cities.

*I thank Ralph Ossa, Lorenzo Casaburi, and Ed Glaeser for their guidance and support throughout the project. I also thank Milena Almagro, Adrien Bilal, Greg Crawford, Don Davis, Ben Faber, Alessandro Ferrari, Gabriel Kreindler, Mathilde Le Moigne, Guilherme Lichand, Jeffrey Lin, Bob Margo, Dennis Novy, Daniel Ramos-Menchelli, Benoit Schmutz, Daniel Sturm, Nick Tsivanidis, and seminar and conference participants at NBER SI Urban, EAYE Paris, University of Rome Tor Vergata, Bank of Italy, LSE, UC Riverside, Cemfi, University of Geneva, PSE, UEA Toronto, CHS Stanford GSB, EUI, NBER Economics of Transportation, MIT CRE, RIEF Brussels, RUSH, SEA Basel, Università della Svizzera Italiana, NYUAD, University of Lucerne, University of St. Gallen, Harvard, and UC Berkeley for helpful discussions and feedback at various stages of this project. All errors are mine.

†London School of Economics, Department of Geography and Environment. s.bagagli@lse.ac.uk

1 Introduction

> *"[The Interstate Highway System] was a program which the twenty-first century will almost certainly judge to have had more influence on the shape and development of American cities, the distribution of population within metropolitan areas and across the nation as a whole, the location of industry and various kinds of employment opportunities (and, through all these, immense influence on race relations and the welfare of Black Americans) than any initiative of the middle third of the twentieth century."*
>
> Senator Daniel Moynihan, 1970

Did expressways increase racial segregation in cities? Within two decades, the US built an integrated network of thousands of miles of roads, connecting the country from side to side and dramatically reducing interstate travel times. However, these roads did more than reshape national connectivity – they transformed metropolitan areas, permanently changing the structure of cities and the urban landscape. This paper investigates the long-run effects of these man-made barriers on racial segregation within cities.[1]

I exploit the construction of expressways in Chicago as a source of variation in neighborhoods' quality and connectivity.[2] Expressways are heavily trafficked and large multilane roads, difficult to cross. Crucially, they were built when the city was already racially mixed. This context hence provides a rare opportunity to study the effects of the expressway construction on urban spaces through the lenses of both an immediate barrier effect and an eventual sorting mechanism. As noted by Emily Badger and Carla Cameron in their Washington Post article (dated July 16, 2015): "Look at racial maps of many American cities, and stark boundaries between neighboring black and white communities frequently denote an impassable railroad or highway, or a historically uncrossable avenue."[3]

I develop an empirical strategy to causally estimate the effects of expressways on neighborhoods' racial composition across two dimensions.[4] The first considers expressways as a source of permanent reductions in residential amenities in the neighborhoods they cross through (Brinkman and Lin, 2022; Robinson, 1971). By degrading the quality of life through increased pollution, noise, and other disamenities in nearby locations, expressways affect the socio-demographic compoisition of the neighborhoods, as reflected in changes in prices and resident populations in impacted areas. The second dimension examines how expressways influence racial sorting through their roles as urban barriers. The variation in accessibility induced by the creation of urban barriers within a city allows me to characterize the barrier effect of expressways. By physically separating neighborhoods on either side of the road, expressways alter exposure to parts of the city characterized by different racial

[1] I follow the current Associated Press's style, which recommends to capitalize "Black" and not "white" when referring to race, ethnicity, or culture more generally (see https://apnews.com/article/archive-race-and-ethnicity-9105661462).

[2] Throughout the paper, I refer to expressways as controlled-access roads only (equivalent to interstates).

[3] https://www.washingtonpost.com/news/wonk/wp/2015/07/16/how-railroads-highways-and-other-man-made-lines-racially-divide-americas-cities/

[4] These urban transportation infrastructures also reduce commuting costs, facilitating a separation between workplace and residential locations. Given the decadal frequency and long-time horizon of my analysis, I do not observe commuting patterns in this setting. While it might be that what I generally define as "disamenity effect" embeds in part the "access benefits effect" of expressways (leading to decentralization), I perform a series of empirical exercises that are consistent with the role of expressways as local disamenities – despite access benefits effects may also be at play.

configurations, thus reshaping patterns of racial sorting. I use these two sources of within-city variation – proximity to the road and changes in accessibility – as identification strategies in both the reduced-form and the structural analyses.

I begin by empirically estimating the disamenity and barrier effects of expressways. The first set of results shows that expressways permanently affect the demographic composition, population size, and property values of nearby neighborhoods. On average, the share of Black residents in neighborhoods near expressways increases by 15 percentage points immediately after their opening and by 20 percentage points in the following decades. At the same time, the total residential population in these areas drops by approximately 30% relative to the full sample mean, compared to unaffected locations.[5] Consistent with the idea that expressways create a negative shock to residential amenities, I also find that neighborhoods closer to expressways exhibit lower house value, land value, and college share (as a proxy for income). The main results are estimated using the two-way fixed effects estimator of de Chaisemartin and D'Haultfoeuille (2020) and satisfy the assumptions for causal identification.[6] These findings are robust to a range of checks, including different choices of treatment and control group bandwidths, the exclusion of certain portions of the city, and alternative empirical specifications.

Next, I measure the impact of the barrier effect of expressways. I define a barrier effect as the increase in the cost of crossing an expressway, measured by travel time. By creating within-city barriers, expressways affect accessibility to different parts of the city, altering the racial mix of areas a neighborhood is exposed to. If expressways create a barrier along racial lines, we may expect neighborhoods to become racially more similar to the areas they are more exposed to – once these urban barriers are in place. I leverage this intuition to test the barrier effect of expressways using a long-difference empirical specification. To estimate how changes in exposure to Black residents influence changes in the racial composition of neighborhoods in Chicago, I construct a measure of accessibility to races that captures spatial spillovers generated by the racial composition of nearby locations. This metric is a location-specific weighted average of the share of Black residents in each neighborhood. The weights decrease with bilateral travel time between locations and depend on the road network: higher weights are placed in locations that are more easily accessible.[7]

I find that higher exposure to Black areas in the city (i) increases the likelihood that a neighborhood becomes more Black over time and (ii) reduces its valuation in the long run. Both effects are sizable. On the one hand, a one standard deviation increase in exposure to Black areas is associated with a 0.16-0.20 standard deviation increase in the share of Black residents in the neighborhood, de-

[5]The change in the racial composition of nearby neighborhoods over time is not simply driven by a drop in total residential population – holding Black population fixed. The results show that these locations experience a change in the racial mix because of both an outflow of the (white) population and an inflow of the Black population.

[6]In the baseline analyses, census tracts in the treatment groups are those within 1 km from the closest expressway; census tracts in the control group are those further than 3 km away.

[7]This summary measure captures, for each origin neighborhood, changes in exposure to Black residents over time. These changes are driven both by sorting and by changes in transportation infrastructure, which alter how accessible other neighborhoods in the city effectively are. An instrumental variable strategy addresses the specific concern that changes in exposure to Black residents may be correlated with unobserved shocks affecting neighborhood composition dynamics. The instrument holds the racial composition of neighborhoods fixed to the pre-period, isolating variation in exposure that is solely due to changes in travel times between locations.

pending on the specification. On the other hand, it leads to a 0.24-0.32 standard deviation decrease in land value, which I use as a proxy for neighborhood valuation. The results remain robust to the inclusion of a rich set of controls, including proximity to the nearest expressway and a measure capturing changes in exposure to rich areas in the city.[8]

A primary concern with the causal interpretation of these results is the potential endogeneity of expressway placement. The decision of where to locate expressways may have been influenced by the socio-demographic characteristics of the areas, as historically argued (Mohl, 2004; 2008; Archer, 2020). I address this concern with a series of checks in the empirical analyses. First, in estimating the disamenity effects of expressways, I show that the assumptions required for identification are valid. The estimated coefficients in the pre-periods are consistent with the assumption of parallel trends in outcomes between treated and control groups. Second, I complement these results with instrumental variable regressions to address the potentially non-random selection of locations in the treatment group. Using instruments widely used in the literature for the assignment of transportation infrastructures that plausibly satisfy the exclusion restriction (Redding and Turner, 2015), I show that the results remain comparable across specifications. Finally, in estimating the barrier effect of expressways, I document that the results are not driven by locations that may have been intentionally isolated. The finding that higher exposure to Black areas increases the likelihood of a neighborhood becoming more Black remains invariant to (i) excluding historically Black neighborhoods from the analysis and (ii) removing the central areas of the city, where the presence of expressways alongside Black communities was arguably more salient.

To quantify the overall impact of expressways on racial segregation in Chicago, I develop a quantitative spatial urban model that incorporates racial preferences in location choices. The setup adheres to a monocentric city model (Alonso, 1964; Muth, 1969; Mills, 1967) but features an internal city structure following Ahlfeldt et al. (2015). The city is populated by an endogenous number of residents from one of four socio-demographic groups (race and educational attainment, in a two-by-two classification). Individuals work in the city center, and commuting is costly. They choose their residential location alongside the consumption of floorspace and final good to maximize utility. The utility of living in a given area depends on residential amenities, land prices, and idiosyncratic shocks. The model produces a heterogeneous type-specific demand system for residential neighborhoods. Expressways enter the model by providing a source of variation in neighborhood quality and accessibility, affecting both amenities and commuting times. Total residential amenities depend on location fundamentals, neighborhood demographics (residential externalities), and the disamenity of expressways. Using analogous sources of variation of the reduced-form analyses, I structurally estimate the parameters governing residential externalities and the disamenity effects of expressways.

GMM estimation results reveal large and statistically significant racial preference and disamenity parameters. The racial preference parameters show substantial heterogeneity by socio-demographic type, particularly between Black and white individuals. The estimates are consistent with Black and white residents exhibiting higher utility for living near same-race neighbors. Residential externali-

[8] I also conduct a series of robustness checks to confirm that the results are not driven by potentially concerning locations for identification, such as historically Black neighborhoods or the downtown area.

ties appear to be an important agglomeration force, particularly in relation to the concentration of same-race residents in the vicinity. For white residents, the elasticity of amenities with respect to the concentration of nearby same-race neighbors is notably higher than for different-race neighbors, ranging from three times as large for low-educated white individuals to six times as large for high-educated white individuals. In contrast, estimates for Black residents indicate strong agglomeration forces linked to the density of same-race residents in the surrounding neighborhoods while suggesting congestion forces concerning the density of nearby white residents.

In addition, residential externalities are highly localized – and appear more localized for Black households than white households.[9] Other things equal, residential externalities fall to zero after around 10 minutes of travel time for Black residents and after around 20 minutes for white residents. For both Black and white groups, the rate of spatial decay of racial preferences goes towards zero faster among the highly educated types. Finally, the size of the disamenity on average appears larger for white residents (attaching 23.9% lower amenities in proximity to the expressway) than for Black residents (22% inferior amenities), attenuating by 95% at 3.8 km from the expressway.[10]

I use the theoretical framework to run counterfactual experiments to evaluate the relation between urban forms and racial sorting. These exercises consist of assuming alternative values of location characteristics or model parameters and solving for the model's counterfactual equilibrium. In all these counterfactuals, I choose the reservation level of utility in the wider economy in the post-period to ensure that the total (type-specific) population in Chicago is equal to its value in 1990. I begin by using counterfactuals to provide further evidence of the model's fit. I analyze the extent to which the effects of the construction of expressways on neighborhood demographics can be explained by the endogenous forces of the model rather than by changes in location fundamentals over time. I find that the model with endogenous forces is able to explain the observed changes in neighborhoods' demographics well (the correlation between the distribution of the share of Black in the observed equilibrium and its counterfactual value is 0.815).

I then run a counterfactual where I study the implications of removing the racial bias. I set the elasticity of amenities with respect to the concentration of different-race residents in the surrounding areas equal to the elasticity of amenities with respect to the concentration of one's race. I find that the counterfactual population of Chicago lives in more integrated neighborhoods. The share of individuals living in neighborhoods with 90% or more Black share drops from 30% in the observed equilibrium to around 5%, while the share of residents living in a neighborhood with at most 10% Black share drops from 50% to less than 30%. The population appears more evenly distributed across all neighborhood configurations.

[9] The rate of spatial decay of racial preferences is equal to 0.674 (s.e. 0.181) for low-educated Black individuals and to 0.747 (s.e. 0.167) for high-educated Black individuals. For low-educated white residents, it equals 0.229 (s.e. 0.047) and 0.291 (s.e. 0.048) for high-educated white residents. The average value of the rate of spatial decay of racial preferences, weighted by the population shares, is 0.342. To the best of my knowledge, this is the first estimate of this sort, making it difficult to benchmark its value. Nevertheless, it is reassuring to observe that it falls within the range of the most pertinent estimates found in the literature. The residential externalities parameter is equal to 0.76 in Ahlfeldt et al. (2015). The elasticity of consumption travel cost with travel times is estimated at 0.019 in Miyauchi et al. (2022).

[10] The magnitudes are comparable to the values recently found in the literature. Brinkman and Lin (2022) find freeway neighborhoods having 18.4% lower amenities, attenuating by 95% at 3.8 km from the expressway (I calibrate the parameter that governs the rate of decay from their work).

Finally, I conduct an additional counterfactual analysis, examining the implications of mitigating the neighborhood effects of expressways on racial sorting – while maintaining their access benefits. The exercise would hence correspond to a policy intervention that places the expressways underground. I investigate the counterfactual treatment effects of simultaneously removing (i) the disamenity and (ii) the barrier effect of expressways. I find that the distribution of races in the counterfactual equilibrium is characterized by a drop of 10 percentage points in the share of residents of Chicago living in a neighborhood with 90% or more Black individuals. At the same time, almost 20% of Chicago's total population would end up residing in integrated neighborhoods (characterized by around 30% Black share) – nearly a seven-fold increase relative to the observed equilibrium. Mitigating the neighborhood effects of expressways is associated with a reduction of 16.8% in racial segregation as measured with the dissimilarity index (moving from 0.844 in the observed equilibrium to 0.702 in the counterfactual equilibrium).

This research builds on three main strands of the economics literature. The first looks at the effects of investments in transportation infrastructure within cities, including Baum-Snow (2007), Baum-Snow et al. (2017), and Gonzales-Navarro and Turner (2018) on population and property prices; Tsivanidis (2022), Baum-Snow (2020), and Heblich et al. (2020) on city structure and welfare.[11] Close to this paper, Brinkman and Lin (2022) evaluate the local adverse effects of freeways on population and welfare in US cities. While their paper is the first to provide evidence of an increase in the cost of travel across highways using travel diary microdata, I fill a gap in this literature by emphasizing the role of expressways as urban barriers. By affecting the level of accessibility to different areas in the city, I show that this feature affects the spatial distribution of people across neighborhoods, on top of more traditional channels studied in the literature – namely, reductions in transportation and commuting costs (for a review, see Redding and Turner, 2015) and changes in relative amenity values.[12]

The second related literature looks at the causes and consequences of residential segregation. Many influential works explore the link between segregation in the US to schooling and labor market outcomes (Kain, 1968; Cutler and Glaeser, 1997; Cutler et al., 1999; Collins and Margo, 2000; Ananat, 2011). Others investigate the emergence of segregation in northern US cities, focusing on early (e.g., Shertzer and Walsh, 2019) or late periods of the Great Migration (Boustan, 2010; Derenoncourt, 2022; Shi et al., 2022).[13] To the best of my knowledge, this is the first paper to provide a long-run analysis of the causal effects of physical barriers on socioeconomic disparities within cities – a link that was,

[11] A large body of works looks at the impact of transportation infrastructure on various outcomes (Redding and Turner, 2015). For instance, long-distance transportation infrastructures (mainly railroads and highways) have been shown to affect land value (Donaldson and Hornbeck, 2016), regional output (Chandra and Thompson, 2000; Jedwab and Moradi, 2016; Ahlfeldt and Feddersen, 2018; Banerjee et al., 2020), trade (Donaldson, 2018; Faber, 2014; Duranton et al., 2014), urban development (Duranton and Turner, 2012; Baum-Snow et al., 2020), migration (Morten and Oliveira, 2023), the spatial sorting of heterogeneous residents and workers (Fretz et al., 2022; Weiwu, 2023), firms and labor markets (Gibbons et al, 2019; Michaels, 2008).

[12] This latter issue saw a surge in recent urban papers (Brinkman and Lin, 2022, Carter, 2023; Ahlfeldt et al., 2019; Anderson, 2020; Mahajan, 2023).

[13] A growing strand of this literature examines the importance of neighborhood effects and social networks on socioeconomic outcomes (Echenique and Fryer, 2007; Cutler et al., 2008), also thanks to advancements in GIS and GPS technologies (Chetty and Hendren, 2018a; Chetty and Hendren, 2018b; Chetty et al., 2016; Athey et al., 2020).

until this research, an anecdotal observation.[14]

Finally, this work speaks to the literature studying the dynamics of neighborhood sorting (Lee and Lin, 2018; Heblich et al., 2020; Heblich et al., 2021; Bayer et al., 2016), and to the growing literature on the spatial sorting of heterogeneous agents (Fajgelbaum and Gaubert, 2020; Davis and Dingel, 2020; Redding and Sturm, 2023; Davis et al., 2023; Gechter and Tsivanidis, 2023).[15] This paper makes two main contributions to this strand of the literature. First, I create a novel measure of exposure to races in the city that builds on the market access literature. This metric allows us to estimate the intensity of spatial proximity and the strength of racial preferences when space matters.[16] Second, I model the geography of Chicago in a general equilibrium framework to estimate racial preference parameters and undertake counterfactual experiments to inform of the continuous neighborhood effects of expressways.

The rest of the paper proceeds as follows. Section 2 discusses the relevant setting and the data. Section 3 introduces the empirical analysis. In Section 4, I estimate the disamenity effect of expressways. In Section 5, I estimate the barrier effect of expressways. Section 6 presents the theoretical framework. In Section 7, I describe the estimation procedure and the results. Finally, Section 8 concludes.

2 Background and data

In this section, I describe the historical context, highlighting the most important events that characterize the development of expressways and the dynamics of racial segregation in Chicago. Then, I provide a quick overview of the primary data used in the analyses and their sources.

2.1 Background

The surge in expressway construction in the Chicago Metropolitan Area was driven by the 1956 Interstate Highway Act. The national plan expanded the mileage of a previous plan, commissioned in 1947, to a 41,000-mile interstate system. The 1956 plan required the federal government to pay 90% of construction costs. The plan's purpose was primarily to improve the connection between major metropolitan areas in the US, to serve US national defense, and to connect with major routes in Canada and Mexico. Within metropolitan areas, the 1956 plan also incorporated some highways meant for local commuting. The construction was started after funding approval in 1956, and by 1975, the national system was mostly complete, spanning over 40,000 miles. By 1990, over 43,000 miles were in operation, and virtually the entire plan had been built throughout the country.

[14] In the sociology literature, a paper by Roberto and Hwang (2017) explores the correlation between physical boundaries and residential segregation, using 2010 block-level data in a cross-section of US cities.

[15] Closely related is also the literature evaluating the dynamics of segregation and tipping points (Schelling, 1971; Card et al., 2008; Dorn, 2008; Logan and Parman, 2017; Caetano and Maheshri, 2019; Gregory et al., 2022; Sethi and Somanathan, 2004; Christensen and Timmins, 2023).

[16] The concept of spatial proximity and the parametrization of racial preferences in this type of framework dates back to Schelling (1971) and has recently been studied by Logan and Parman (2017). In both works, the notion of segregation is based on next-door neighbors. A key contribution is estimating the parameter that governs the rate at which racial preferences decay spatially.

Figure 1 shows that the rollout of expressways in the Chicago Metropolitan Area followed the national trend. Each panel of the figure displays a snapshot in time of the completed portions of the expressway network in the respective census year. The first segment was completed in 1951, and by 1970 virtually all roads drawn in the plan were in operation. Only a few segments of suburban ring routes are of more recent construction. In central areas, the expressway network was completely laid out by 1970.

In this paper, I focus exclusively on expressways (interstate highways) due to their distinct physical attributes. Expressways are controlled-access roadways, meaning that entrances and exits are limited. All traffic merges on or off via ramps that connect expressways to highways, secondary roads, or tertiary roads. Therefore, any route can cross an expressway only through an overpass or underpass.[17]

In addition to connecting major cities, these multilane roads served the local network in metropolitan areas and contributed to suburban sprawl. Faster commuting times played a crucial role in reshaping the spatial distribution of the population in U.S. metropolitan areas between 1950 and 1990. Baum-Snow (2007) finds that highways account for approximately one-third of the total decline in central city residents relative to the metropolitan population. This pattern of suburbanization, which affected most American cities, also transformed Chicago. In 1950, the Chicago metropolitan area had a population of more than 5.1 million, with 3.6 million residing within city boundaries. By 1990, the city's population had declined to less than 2.8 million (a 22% drop). In contrast, the total metropolitan population grew by more than 50% over this period, reaching 8 million in 1990.

The creation of the Interstate Highway System also came with an intertwined history of infrastructure development and racial inequality. As former US Secretary of Transportation Anthony Foxx stated in 2016, these roads were often routed through disadvantaged neighborhoods, where the poorest residents — almost always racial minorities — lived. In some cases, they were even deliberately designed to separate neighborhoods.[18] "It became clear to me only later on that those freeways were there to carry people *through* my neighborhood, but never *to* my neighborhood."[19] Historical accounts document the link between racial configurations and the passage of highways across neighborhoods (Mohl, 2004; 2008; Archer, 2020).

Chicago was no exception. In the late 1940s, city officials estimated that the planned expressways would destroy over 8,100 housing units (Mohl, 2001). Figure A3 shows two sides of Troop Street in 1949, when the Eisenhower Expressway – running westwards from the Chicago Loop – was constructed. Near the city center, the road primarily cut through Italian and Greek communities, forcing many residents to relocate to the Northwest Side and the suburbs. The small Black community, in contrast, moved further west. Another notable example is the Dan Ryan Expressway, which runs north-south. In 1956, its alignment was shifted west, arguably to isolate the growing Black

[17] In contrast, other types of highways (such as US highways, state, and county highways) do have intersections, even with minor roads. While some sections may be limited-access, most of their routes function as free-access roads, allowing private drives to enter and exit directly. For more details, see Appendix A.1.1.

[18] See for instance the March 28, 2016 Washington Post article on the issue: https://www.washingtonpost.com/local/trafficandcommuting/defeating-the-legacy-of-highways-rammed-through-poor-neighborhoods/2016/03/28.html

[19] Anthony Foxx, before the Center for American Progress (Washington DC, March 30, 2016).

community on Chicago's South Side. Long considered as "lifeline of the South Side", the Dan Ryan Expressway created a zone of demarcation, as many neighborhoods along its route deteriorated.[20] As these cases illustrate, the placement of expressways within the urban landscape was unlikely to have been random. Local factors likely influenced their precise routing, as is generally the case in the assignment of transportation infrastructure to various locations (Redding and Turner, 2015).[21]

When expressways were laid out, Chicago was home to many minority groups, whose presence had grown significantly since the late 19th century. During the first wave of the Great Migration (1910–1920), more than 50,000 African Americans moved to Chicago from the rural South in search of job opportunities, increasing their share of the city's population to 4.1%. Foreign immigration declined during the Great Depression and World War II, while the Black population continued to grow. By 1944, nearly one in ten Chicago residents was Black.

As foreign-born communities declined in size, Black neighborhoods became increasingly prominent, giving rise to the so-called "Black Metropolis" on the city's South Side. By the late 1940s, over 90% of Chicago's Black population lived in this area, corresponding to the largest purple region in Figure A5 in Appendix A.1.3).[22] In contrast, Hispanic immigration grew steadily only in the second half of the 20th century. In 1960, just about 1.5% of Chicago's total population was Hispanic, but by 1990, this figure reached nearly 20%.[23] In population censuses — my primary source of demographic data — Hispanic origin was not separately identified until after 1970. In the following analyses, I distinguish only between Black and non-Black individuals; however, this categorization remains sufficient for studying the segregation of one minority group from others. Due to data limitations, individuals of Hispanic origin, along with other racial and ethnic minorities, are consistently grouped with the white population.

2.2 Data

The primary data sources are the decennial US census of population and housing covering the period between 1920 and 2010, combined with GIS data. The primary geographic unit used in the analyses is the census tract.[24] To ensure comparability over time, historical census tract boundaries are normalized to 2010 boundaries, following the procedure in Lee and Lin (2018). The geographic extent of the metropolitan area is determined by data availability in 1950. This approach results in 1,511 consistent boundary census tracts that partition the Chicago Metropolitan Area between 1950 and 2010. Of these, 780 tracts were already surveyed at the beginning of the 20th century, allowing for data coverage spanning a full century at 10-year intervals (1920–2010). The dataset provides fine-grained

[20] https://www.chicagotribune.com/news/ct-xpm-1998-03-01-9803010173-story.html

[21] I address this endogeneity concern in the empirical strategy and establish the assumptions necessary for causal identification.

[22] With the vast majority of the Black population concentrated in a narrow stretch of land on the South Side, Chicago was already a segregated city before expressways were built, as shown in the figure. However, this area was too small to accommodate the growing number of African Americans migrating to the city. Since then, the Black population has tripled, and the city's spatial structure has changed dramatically.

[23] Today, Chicago is among the most diverse US cities, with its largest racial and ethnic groups making up nearly equal shares of the population: https://fivethirtyeight.com/features/the-most-diverse-cities-are-often-the-most-segregated/

[24] A census tract typically covers an area of approximately 2 square kilometers and includes around 6,000 residents.

spatial information on a wide range of household socio-demographic characteristics and housing attributes. Below, I outline the main variables of interest and the procedure used to construct the time series, grouping them by data source. Summary statistics for 1950 are reported in Table A1 in Appendix A.

Censuses of population and housing The primary demographic of interest is the racial composition of each census tract. Given the long time horizon of the analysis and the changing demography of the period, the type of information available from the decennial censuses allows to consistently distinguish between two race categories: "Black" and "non-Black".[25] Starting in 1980, race classification in the census became more detailed. For these later periods, I define Black residents as those recorded as "Black, non-Hispanic" (following Logan et al., 2014). Consequently, in the analyses that follow, the complementary category to the share of Black residents is the share of non-Black residents, which includes white individuals, people of Hispanic origin, and other racial and ethnic groups.

Since census tract-level data on (self-reported) household income is available only from 1950 onward, I supplement granular demographic data with an estimate of the share of college graduates, consistently covering the period between 1940 and 2010. This statistic is computed as the share of individuals aged 25 and older who completed at least four years of college.

Land and road network data I supplement the US population and housing census data with additional sources.

First, to address the limitation that census data do not include information on multi-unit buildings, I use Olcott's Land Value Blue Book, a dataset unique to Chicago. This collection provides estimates of land values for every city block in Chicago for most of the 20th century. The data, digitized and made available by Ahlfeldt and McMillen (2014; 2018), gives spatial granularity at 300× 300-foot grid cell level.

Second, contemporaneous transport networks are sourced from the US Census Bureau. From the complete road network of Chicago, I extract expressways that traverse the study area.[26] Starting from the present-day network, I reconstruct snapshots of expressway expansion at 10-year intervals corresponding to census years. The opening dates of expressways are initially from Baum-Snow (2007), while the precise assignment of opening years for individual road segments within the city is based on Illinois and Indiana State Maps issued in the census years.

[25] An exception is the 1940 census, which recorded race only as "white" versus "non-white". To ensure consistency, I re-coded the 1940 categories to align with the rest of the time series. Since other non-white racial groups were a small share of the population at that time (Hispanic immigration remained limited before the 1960s), I classify all individuals recorded as "non-white" in 1940 as Black. To validate this assumption, I compared the resulting tract-level shares of Black residents with data from a 1934 population and housing census conducted exclusively in Chicago. The 1934 census classified race as "White", "Negro", and "Other", in line with other census periods. For comparability, I computed the tract-level share of Black residents using the same weighting scheme derived from the 1940 census data. Reassuringly, the correlation between the two measures (from 1934 and 1940) is 98.83%.

[26] One reason for restricting the analysis to expressways (excluding other types of highways and major roads) is their technical attributes. Expressways must adhere to specific standards, the most important being their controlled-access design, which limits entry and exit points. This feature makes them particularly relevant for identifying urban barriers, as it restricts movement across them and disrupts local connectivity.

Finally, the 1940 road network is obtained from the Urban Transition HGIS Project (Shertzer et al., 2016).[27]

2.2.1 Sample

The original sample consists of 1,511 census tracts with fixed boundaries (normalized to 2010) covering the area that was part of the Chicago Metropolitan Area in the 1950 census. Of these, 791 tracts lie within the boundaries of the city of Chicago, while the remaining 720 constitute the suburban area. For the earliest periods, the information is only available for the subset of 757 tracts that had already been surveyed at the beginning of the 20th century.

Notorious public housing projects, such as Cabrini-Green and the Robert Taylor Homes, were developed by the Chicago Housing Authority as part of mid-20th-century urban renewal efforts. These large high-rise housing developments were predominantly occupied by Black residents. To isolate the neighborhood effects of expressways, in my main analyses I exclude from the sample the census tracts that hosted these projects (see Map A4 in Appendix A) and conduct robustness checks using larger exclusion radii around them. For completeness, I also present results including public housing neighborhoods in the appendix. The results remain stable.

3 Overview of the empirical analysis

Expressways, primarily designed to connect distant locations, may also shape patterns of local segregation. In this section, I outline the empirical strategy and present reduced-form evidence of their impact on segregation along two dimensions. The first considers ex- pressways to create permanent reductions in residential amenities in the neighborhoods they cross through. By making some places worse than others due to increases in pollution, noise, and other disamenities in nearby locations, expressways may affect the area's socio-demographic composition through a decrease in its valuation, as reflected by differences in house prices. The second starts from the widely observed phenomenon that racial segregation in US cities tends to occur along man-made barriers such as railroads, highways, or large city roads. I focus on expressways because (i) they are heavily trafficked and large multilane roads that are difficult to cross, and (ii) they were built when the city was already racially mixed.

To assess the long-term effects of expressways on neighborhood composition, I leverage the time dimension of the empirical setting, accounting for dynamic treatment effects. First, I estimate the disamenity effect of expressways using a difference-in-differences specification with multiple periods. The analysis spans 1950-2010 (the first expressway in the city opened in 1955) for all units in my data, and extends back to 1920 for areas already surveyed at the beginning of the 20th century. Second, I estimate the barrier effect of expressways with a long-difference empirical specification, measuring the impact of changes in exposure to Black residents on changes in neighborhood racial composition. To the extent that expressways create within-city barriers that affect accessibility to different parts of

[27] https://s4.ad.brown.edu/Projects/UTP2/ncities.htm

the city, I test whether neighborhoods become racially more similar to the areas they are exposed to once these urban barriers are in place.

4 Disamenity effect of expressways

I employ multi-period difference-in-differences specifications to estimate the dynamic impact of proximity to expressways on the valuation of neighborhoods. I compare the average outcomes of census tracts near expressways to those of the control group. The estimating equation is of the following form:

$$Y_{it} = \alpha_i + \gamma_t + \sum_{j=-2}^{6} \beta_j D_i \times T_{i,t=t^*+j} + \epsilon_{it} \tag{1}$$

where i is a census tract; t is a census year; Y_{it}, the outcome variable of interest, measures the share of Black residents or total population (depending on the specification) in census tract i at time t; D_i is an indicator variable for being close to the expressway (in the baseline specification, it takes a value equal to 1 if the census tract's centroid is within 1 km from the nearest expressway – corresponding to the largest distance at which both pollution and noise are estimated to reach the benchmark levels of areas with no highways – and 0 if it is more than 3 km away from the closest expressway);[28] $T_{t \geq t^*}$ is an indicator for the post-construction period; α_i are census tract fixed effects; γ_t is a set of time fixed effects (further interacted with a set of baseline controls, as stated below each figure). All standard errors are clustered at the census tract level. In robustness checks, I show that the results remain quantitatively similar using alternative choices of the bandwidths defining treatment and control units. The baseline specification leaves a 2 km buffer between treatment and control units to address, in a reduced-form sense, potential spillover treatment effects that could contaminate nearby control locations. I also show that results remain mostly robust to more conservative clustering approaches (see Section 4.2 below for more details).

The coefficients of interest are β_j, which capture the effect of being exposed to treatment since j time periods. Each β_j estimates the difference between treatment and control group outcomes at event time j. Negative values of j allow us to check for the existence of pre-trends in the depen-

[28]In the baseline specification, I consider as treated all census tracts that lie within 1 km from the nearest expressway. This corresponds to the largest distance at which both pollution and noise are estimated to reach the benchmark levels of areas with no highways. Concerning pollution, Karner et al. (2010) – integrating the results of 41 monitoring studies of the dispersion of near-road air pollutant concentrations – report that the concentration of ultrafine particles (UFPs) achieves the background level (of no highways) at a distance of 910 meters from the source. As for the noise produced by expressways, it is estimated to vanish at a distance on average smaller than the pollution derived from the same source. In a free field (i.e., assuming equal sound propagation in all directions), noise obeys the inverse square law: sound intensity decreases by nearly 6 decibels for each doubling distance from the source. Under these conditions, the sound produced by expressways – estimated at around 75 decibels by the Federal Highway Administration – reaches the ambient noise level recommended by the WHO of 40 decibels at a distance of 320 meters. In the real world, in the presence of reflections of reverberations, sound propagation does not follow precisely this law, but the estimates seem relatively close. From the transportation noise map developed by the US Department of Transportation – representing the approximate average noise by mode – in 2018, the noise produced by expressways in Chicago faded at a distance of around 250 meters. For an illustration, a section of the interactive map produced by the Bureau of Transportation Statistics and its source is reported in Figure A8 in the Appendix A.3.

dent variable. All β coefficients are normalized relative to β_{-1}, the decade just before entering into treatment.

The results are computed following the two-way fixed effect estimator proposed by de Chaisemartin and D'Haultfoeuille (2020), which is valid even when the treatment effect is heterogeneous across groups or over time, as it may be in this setting. In two-way fixed effects regressions where the treatment effect is not constant, the estimated coefficient is equal to a weighted sum of several difference-in-differences that compare the evolution of the outcome between consecutive time periods across pairs of groups. However, since the control group may effectively be treated in both periods in some of these comparisons, some weights may be negative. This may be an issue when the average treatment effects are heterogeneous across groups or periods because the estimated coefficient may be of a different sign than the average treatment effects of the pairwise comparisons. To test the severity of this concern in my setting, I run the *twowayfeweights* command in de Chaisemartin and D'Haultfoeuille (2020) in my baseline specification with the full set of controls. I find that 86.1% of the average treatment effects on the treated receive positive weights (in the specification with no controls, 78.0%), with weights summing up to 1.041, while 13.9% receive negative weights (and their sum is equal to -0.041).[29]

The baseline specifications run regression (1) with a set of baseline variables (each interacted with year-fixed effects). First, I flexibly control for distance to the Central Business District (CBD). This variable is likely to affect both the outcome variables – Black households and minority groups more generally settled first in the central parts of the city – and the treatment status due to the radial structure of expressways. Second, I include a "city center" fixed effect – which effectively splits the sample into central and suburban areas.[30] This additional control is particularly useful to isolate the disamenity effect of expressways.[31] Under the assumption that distance to the expressways within the city does not affect the likelihood of moving to the suburbs, the city fixed effect isolates the change in relative residential amenities induced by the disamenity effect of expressways – net of suburban movements.[32] Third, I control for baseline population density since it likely correlates with both treatment status assignment and the primary outcome of interest (Black and minority groups more generally tended to live in denser areas). Finally, I control for baseline neighborhood characteristics

[29]The command also reports two summary measures of the robustness of the estimated coefficient. The first corresponds to the minimal value of the standard deviation of the treatment effect between the treated groups and time periods under which β and the average treatment effect on the treated could be of opposite signs. I find a value of 0.762 (after standardizing the outcome variable to make the magnitude clearer to interpret). Instead, the second summary measure corresponds to the minimal value of the standard deviation of the treatment effect across the treated groups and the time periods in which all the average treatment effects are of a different sign than β. The reported value is 3.889 (after standardizing the outcome variable). Reassuringly, both summary measures appear to be large, suggesting that β and the average treatment effects could be of opposite signs only if there is a lot of treatment effect heterogeneity across groups or time periods.

[30]Central areas are those within the administrative boundaries of the City of Chicago.

[31]Expressways change the relative amenity of the places they serve in two fundamental ways. The first is the so-called access benefit: expressways increase the relative residential amenity of farther away locations, which are now easier to reach and commute from. The second is the disamenity effect: expressways change the relative amenity of residential neighborhoods by making some places worse than others through increases in pollution, noise, and other disamenities. While these two dimensions are likely to go hand in hand, this paper focuses on the latter.

[32]One threat would be if displaced households near the expressway were disproportionately forced to move to the suburbs because of the lack of alternative housing opportunities downtown. However, this scenario does not seem to be supported by historical evidence.

proxied by the grades assigned by the Home Owners Loan Corporation (HOLC) in the 1930s (Fishback et al., 2020).[33] Given the history of discrimination and unequal treatment during the settlement of the Black population and European immigrants in industrialized US cities, minorities tended to live in disadvantaged and economically distressed neighborhoods. The HOLC maps reflected these long-lasting inequities.[34] Omitting the neighborhood effects of redlined areas in this context would likely bias the estimated coefficient upwards. In robustness checks in Appendix B, I report the results without controls and show that the effects remain quantitatively similar (Figures B1 and B19 respectively for share Black and population as outcome).

The empirical specification assumes that, in the absence of expressways, treatment and control group outcomes would have evolved similarly. However, the 1920s saw large population movements, raising concerns about the validity of the identification assumption for this early period. During this time, the city underwent rapid urban expansion, while the Black population remained primarily concentrated in the city center. As a result, identifying a reliable comparison group for this earliest period is challenging.[35] In what follows, I omit the year 1920 from the analysis. However, in Section B.1.9 of Appendix B, I also report results including 1920, showing that the pre-trend in outcomes for this early period disappears once I account for areas that experienced contemporaneous changes in population due to factors unrelated to expressways.

4.1 Main results

Following Brinkman and Lin (2022), I first report the effects of expressway proximity on changes in residential population, a measure that reflects shifts in the quality of life in affected neighborhoods. The estimated coefficients from regression (1) on the residential population are reported graphically in Figure 2. Event times to the right of the red vertical bar denote post-treatment periods, with event time 1 marking the decade in which the expressway opens to traffic. Event times to the left of the dotted red vertical bar instead correspond to pre-treatment periods. Finally, event times between the two vertical bars represent the period from the first expressway construction plan to the decade when the expressway became operational.

The results show a persistent decline in residential population in affected areas compared to control units. During the construction phase (before expressways opened to traffic), residential population decreases by an average of 331.52 (s.e. 118.77), corresponding to about 8% of the full sample mean. After expressways became operational, the residential population continues to decline by approximately 500 people each decade (equivalent to 12% of the full sample mean), reaching a total reduction of around 2,000 people by decade four into treatment. The average treatment effect is estimated at -1,359.84 (s.e. 216.25). Reassuringly, the estimated β coefficients for pre-treatment peri-

[33] See Appendix A.2 for a brief description of redlining and the boundaries drawn in Chicago.

[34] Recent studies have shown their long-term consequences in terms of home ownership, house values, and rents (Aaronson et al., 2021; Krimmel, 2020).

[35] Figure B14 in Appendix B plots the raw average residential population in (eventually) treated versus control areas over time using a binned regression (without controls). Between 1920 and 1930, the population of what would later constitute the control group doubled before stabilizing at a slower growth rate, raising concerns about its suitability as a counterfactual in this early period.

ods are consistent with the assumption of parallel trends in outcomes between treated and control groups.

Next, I examine the effects of expressway proximity on neighborhood racial composition. The estimated coefficients from regression (1) on the share of Black residents are reported graphically in Figure 3. Living near an expressway is associated with a sharp increase in the share of Black residents, all else equal. On average, areas near expressways experience an increase of 15.43 percentage points (s.e. 0.02) in the share of Black residents in the first decade after treatment and an average increase of around 20 percentage points across subsequent decades relative to the pre-expressway period. The average treatment effect across all post-treatment periods is 15.72 percentage points (s.e. 0.03). The estimated β coefficients for pre-treatment periods are consistent with the assumption of parallel trends in the outcome variable between treated and control neighborhoods before intervention. During the construction phase, eventually affected locations experience both a decline in residential population and an initial shift in neighborhood demographics, which intensifies in the post-treatment period.

4.2 Robustness checks

I show that the results remain robust and quantitatively similar across a series of robustness checks, including: (i) removing controls; (ii) using a balanced panel (restricting the analysis to observations with full data coverage); (iii) weighting observations by baseline population; (iv) varying treatment and control groups bandwidths; (v) excluding portions of the city; (vi) using alternative clustering levels for standard errors; (vii) employing alternative empirical specifications; and (viii) applying the semi-parametric difference-in-differences estimator of Callaway and Sant'Anna (2020). In addition, I present the results of a set of IV regressions designed to address potential non-random selection into treatment. The estimated effects are quantitatively similar to the event-study results.

Robustness exercises are described in detail in Appendix B and summarized here. First, I report results for both changes in the share of Black residents and residential population without controls, showing that the estimates remain quantitatively similar. Second, I restrict the analysis to a balanced panel of 760 observations per period. These tracts correspond to the portion of the City of Chicago enumerated at the start of the study period, aligning with the most historically significant sites. The results remain virtually unchanged. Third, to ensure that findings are not driven by low-population or non-representative census tracts, I weight observations by baseline population density. The results remain robust.

Fourth, I show that the results do not significantly vary with the choice of the bandwidth used to assign census tracts to the treatment and control groups. In my baseline specification, I conservatively account for potential spatial spillovers by leaving a 2 km buffer zone between treated observations (census tracts with centroids within 1 km of the nearest expressway) and control observations (centroids farther than 3 km from the nearest expressway). This approach yields nearly equal-sized groups (388 and 454 observations at baseline, respectively), but at the cost of excluding approximately 40% of sample units (those with centroids located between 1 and 3 km from the expressway). Nevertheless, Appendix B confirms that the results are robust to alternative definitions of treated and control units. When I compare spatially closer groups, particularly when I allow the control group to be located nearer to the expressway, the estimated coefficients attenuate but remain highly significant. These findings suggest that my empirical design effectively captures the impact of expressways on neighborhood racial composition and demographic changes.

Fifth, I provide additional evidence that the main results are not driven by specific areas that underwent significant changes during this period. I restrict the analysis to census tracts within the boundaries of the City of Chicago and find that the results remain virtually unchanged. The findings also hold after excluding the area that historically housed the vast majority of the Black population (i.e., the so-called Black Belt), confirming the treatment effects are not driven by these locations. For completeness, I also report results including areas where large public housing projects were developed. To account for potential spillover effects in nearby neighborhoods, I run robustness checks by removing census tracts within a certain radius (500 meters, 1 km) from a public housing project. In all cases, results do not change.

Sixth, I test the robustness of the results under different conservative clustering approaches to account for spatial correlation in the errors. For the sample within the boundaries of the City of Chicago, I show that the results remain largely strong after partitioning the city into 25 equally sized grid cells (6×6 km) and clustering standard errors at this fine spatial level. As an additional exercise, I cluster standard errors at a broader regional level—north, west, and south sides of Chicago—despite this leading to only three clusters. Even in this setting, the estimates remain strongly statistically significant. For the full sample, consisting of the metropolitan area, I cluster standard errors after dividing the city into 60 equally sized cells (8×8 km). While this increases standard errors, the main results remain robust.

Seventh, I further examine the negative relationship between the share of Black residents and proximity to expressways using a distance-based regression. I divide the continuous variable measuring distance to the nearest expressway into five roughly equal-sized distance bins (Appendix B.3). Across both the city and suburban areas, the share of Black residents decreases with distance from the expressway. In the city, the share of Black residents increases by between 9.6 and 33.5 percentage points (from the farthest to the closest bin) relative to suburban areas located more than 4 km from the nearest expressway (the omitted category). In the suburbs, the magnitudes are smaller, but the negative relationship between the share of Black residents and expressway distance remains evident (Figure B31).

Finally, I show that the estimated causal effects of expressway proximity remain robust when using the Callaway and Sant'Anna (2020) difference-in-differences estimator (Appendix B.4). Like the estimator developed by de Chaisemartin and D'Haultfoeuille (2020), this method produces unbiased estimates in settings with multiple time periods and variation in treatment timing. In two-by-two designs, it estimates group-specific average treatment effects across all periods while imposing a weaker parallel trends assumption.

4.2.1 IV results

The identification assumption for estimating regression (1) is that, in the absence of expressways, census tracts in the treatment and control groups would have evolved similarly, conditional on controls. Under this assumption, the estimated β coefficients capture any deviations from parallel trends in outcomes caused by the expressway rollout. The richness of temporal data in my empirical setting allows me to test for pre-trends in outcome evolution between treated and untreated units before treatment – thus providing empirical support for the parallel trends assumption, which is necessary for valid identification in my preferred specification.

As an additional robustness check, I also run a set of IV regressions, a common approach in the literature to address concerns about the potential non-random selection of locations into the treatment group. The results remain consistent across specifications. In Appendix B.5, I report IV results using instruments widely used in the literature to assign transportation improvements within cities that plausibly satisfy the exclusion restriction (Redding and Turner, 2015).

Various instrumental variables can be used to predict expressway routes in this setting. One approach relies on the development of a straight-line instrument. It isolates variation stemming from the fact that expressways were primarily designed as long-distance road infrastructures to connect cities nationwide, rather than to facilitate metropolitan development (Baum-Snow, 2007). The instrument is constructed as straight lines connecting Chicago to the cities targeted by the 1947 Interstate Highway System plan. As a result, neighborhoods receiving expressways were those that happened to lie in the direction of these targeted cities.

Another suitable set of instruments leverages historical routes. I use proximity to the 1898 railroad network as an instrument for the current location of expressways. Identification relies on the assumption that historical routes are unlikely to be correlated with current changes to neighborhood characteristics, conditional on baseline controls. Instrument relevance is justified by the fact that expressways tend to follow pre-existing railroads due to lower right-of-way costs.

Finally, unique to this context, I use proximity to routes shown in the 1909 Burnham Plan as an instrument for proximity to an actual expressway. Since the plan was developed before the Great Migration, it is plausibly uncorrelated with more recent changes in neighborhood demographics.

The IV results are reported in Appendix B.5. In the preferred specification that combines all available instruments (column 6), every additional kilometer from the nearest expressway is associated with a 0.04 percentage point decrease in the share of Black residents, all else equal. These results are quantitatively similar to the event study results.[36] The IV estimates are slightly larger in magnitude than the corresponding OLS estimates from the two-by-two design (reported for comparison in the same table), suggesting that the observed changes in neighborhood composition in areas close to the expressways may be somewhat understated. The larger IV estimates imply that expressways were more likely allocated to growing neighborhoods, amplifying their effects, rather than to declining neighborhoods (a result also found in Brinkman and Lin, 2022). While suggestive, these results alleviate concerns about negative selection in expressway placement, hinting that expressways were not primarily targeted toward locations expected to decline.

[36]The average distance from the closest expressway in the control group of the event study (census tracts with centroids more than 3 km away) is 5.3 km, while the average distance in the treatment group (tracts within 1 km) is 0.5 km. As a result, the implied difference in the share of Black residents between treated and control locations from the IV estimates is $(5.3 - 0.5) * 0.04 = 0.192$, or 19.2 percentage points, on average. The baseline event study specification reports an average treatment effect of 15.7 percentage points.

4.3 Effects by wave of expressway construction

A concern may, in principle, regard the non-random timing of expressways' construction – specifically, whether local factors influenced the decision of which segments to build first. However, two considerations suggest that this potential issue is unlikely to pose a serious threat here: (i) census data is available at 10-year intervals, meaning selection concerns apply only over larger time horizons; (ii) 97% of the eventually treated units in the sample were connected to the expressway network within just two consecutive census waves (i.e., with only one decade separating the two treatment groups). Additionally, all expressways near the city center were completed within the first two decades (see Figure 1). Only a few suburban ring routes are of the more recent period, between 1971 and 2010.

Since virtually all treated units entered treatment either in 1960 or 1970, I estimate separate leads and lags regressions for each group. This specification allows for the estimation of treatment effects in each census year by treatment group, capturing potential heterogeneity in the effects of expressway construction depending on the construction period.

The results are reported in Appendix B.6. To increase precision, observations are weighted by population density at baseline. Figure B36 panels (a) and (b) plot the estimated average treatment effects on the share of Black residents for units treated in 1960 and 1970, respectively, relative to never-treated (control) units. The results reveal that the impact of expressway proximity differs between units treated in 1960 and those treated in 1970, a pattern masked in the event study results.

For census tracts treated in 1960, the share of Black residents increased by approximately seven percentage points, on average, relative to 1940 levels, already in the decade before expressways opened to traffic, suggesting that demographic shifts began during the construction phase. This trend continued: By 1960, the share of Black residents in treated units had risen more than 20 percentage points relative to 1940, and stabilized at a 40 percentage point increase in subsequent years. Appendix Figure B37, panel (a), shows a systematic population decline in these tracts relative to pre-expressway periods but with no evidence of anticipation effects. Taken together, these findings suggest that non-Black residents tended to leave neighborhoods designated for expressway construction, and their places were fully occupied by incoming Black residents.

The results for 1970, however, depict a different story. The share of Black residents decreased consistently following expressway construction, declining by four percentage points as early as 1960, during the construction phase. This decline coincided with a substantial population drop – by more than 700 residents (19% of the full sample mean) already in 1960, the decade before these expressways were operational. These findings are consistent with those of Brinkman and Lin (2022), which documents that, from the mid-1960s onward, expressways were increasingly routed through Black neighborhoods, and Carter (2023), which studies displacement patterns in Detroit.

The evidence suggests that, by 1970, expressways were more likely to pass through densely populated areas, leading to greater displacement and disproportionately targeting Black neighborhoods. In contrast, expressways built in 1960 did not exhibit the same pattern: total population declined only after expressways became operational, not before. These results are consistent with expressways generating local disamenities that reduced the desirability of nearby areas, contributing to neighborhood deterioration. The findings suggest that expressways have persistent and reinforcing effects on neighborhood valuation – and, in turn, on residential population trends.

4.4 Additional evidence

Appendix B.7 reports additional results, summarized here.

First, I document changes over time in the number of housing units following expressway construction. In event time 0 (when construction was ongoing), affected locations saw an average reduction of 139.01 housing units (s.e. 29.57) relative to control group locations. The number of housing units continued to decline in subsequent decades, stabilizing at a reduction of approximately 600 units (44.6% of the full sample mean). While part of this drop is likely mechanical – since expressways physically displace buildings – the sustained decline suggests lower investment in these locations relative to comparable areas further away from expressways.

To further assess whether the decline in residential population reflects worsening neighborhood conditions and declining property values, I examine the effects of expressway proximity on house and land values.[37] In addition, to better understand compositional changes in affected neighborhoods, I analyze the impact of expressways on the share of college graduates. This additional set of results aligns with those of Brinkman and Lin (2022).

Figure B42 plots the estimated β coefficients from a regression using average house value (in real terms) as the outcome. The results show that in the first decade after expressways opened to traffic, self-reported house values remained stable relative to the pre-expressways period. However, one decade later, house values declined by an average of -\$24,072.40 (s.e. 4,871.94), equivalent to 16.25% of the full sample mean. The graph also suggests a mean reversion in later decades, with the estimated average treatment effect equal to -\$14,359.72 (s.e. 4,627.94). When comparing raw means between eventually treated and never treated units (Figure B43), separately for suburban and the central areas, the mean reversion observed in the event study appears to be driven by the 2000 housing bubble and the subsequent subprime mortgage crisis. House values declined sharply between the 2000 and 2010 census periods, with even steeper drops among control group units in both central and suburban areas.

Figure B44 plots the estimated β coefficients for land value. Since Olcott's land value data covers only the most central part of the Chicago metropolitan area, the analysis is restricted to census tracts within the City of Chicago. The results are relatively noisy. In general, locations that received the expressway were on an upward trend in land values before construction, a trend that was entirely offset once expressways became operational. On average (albeit suggestively), land value declined by around 0.2 log points in the first decades into treatment before becoming statistically indistinguishable from zero by decade four. As with house values, the estimated coefficients suggest a mean reversion effect, though here it starts as early as the second decade after treatment. Examining raw means between eventually treated and never treated units (Figure B45), this mean reversion in later periods may be driven by the housing boom – and potentially to gentrifying central neighborhoods.

Finally, I further explore the effects of expressway construction on neighborhood demographics,

[37]The sample correlation between average real house value and (log) land value is 0.58. Appendix Figure A9 plots binned scatter-plots of their relationship.

using the share of college graduates as the outcome of interest. I consider college share a proxy for income, as self-reported income was only included in census records starting in 1950.[38] The results are reported in Figure B46. The average treatment effect of expressways on the share of college graduates corresponds to a reduction of 0.03 percentage points (s.e. 0.01), equivalent to 17.6% of the full sample mean. Similar to the results on house values, the effect is strongly negative but does not appear until the second decade after treatment.

4.5 Discussion of the disamenity effect of expressways

The first empirical design of the paper isolates the dynamic impact of expressway proximity on racial segregation and neighborhood valuation. Using multi-period difference-in-differences specifications, with de Chaisemartin and D'Haultfoeuille (2020) two-way fixed effects estimator, I show that, on average, expressways are associated with an increase of approximately 15 percentage points in the share of Black residents living in nearby neighborhoods during the first decade of treatment and an average increase of 20 percentage points in subsequent decades. The effect of expressways on within-city racial sorting is persistent: longer exposure to expressways leads to a permanent increase in the share of Black residents in affected areas. Consistent with the idea that expressways create a negative shock to residential amenities, I also find that the residential population declines permanently in nearby affected places in the post-expressway period. On average, affected areas lose approximately 30% of the full sample population mean. In addition, neighborhoods closer to expressways tend to experience lower house values, land values, and college share (used as a proxy for income).

I conclude the section by briefly addressing a few remaining confounding factors. On the one hand, part of the decline in residential population and housing units following expressway construction is likely due to the physical displacement required to build these roads. In Appendix B.8, I compare changes in outcomes between locations treated early and those treated later, with the idea that these areas should be roughly comparable to one another in the process undergoing construction. The persistent and worsening decline in the residential population of affected neighborhoods cannot be explained by displacement alone. On the other hand, population and housing declines, particularly marked in central areas, may partly reflect increased demand for commercial space. This concern is addressed in Brinkman and Lin (2022). Using travel surveys conducted in Chicago (1956-2000) and in Detroit (1953-1994), they overall find "little evidence that central freeways caused local negative effects [on population growth] by attracting jobs."

5 Barrier effect of expressways

In this section, I examine the presence of a barrier effect of expressways. I define the barrier effect as an increase in the cost of crossing an expressway. These higher costs affect accessibility between different parts of the city, which are characterized by distinct racial compositions. To the extent that

[38] The two measures are highly correlated, with a sample correlation of 0.77. Besides being already available since 1940, college share is less prone to measurement error. Appendix Figure A10 plots binned scatterplots of their relationship.

expressways create within-city barriers, I test whether neighborhoods become racially more similar to the areas they are exposed to once these urban barriers are in place.

The hypothesis to test is based on the intuition that the barrier effect of expressways manifests itself by increasing racial divergence between areas on opposite sides of the road while reinforcing similarity between neighborhoods on the same side. In an ideal experiment, within each pair of census tracts cut by an expressway, the tract experiencing the largest increase in exposure to Black residents should become increasingly more Black over time, relative to the census tract on the opposite side of the road. The identification assumption required to causally estimate the barrier effect is that, in the absence of expressways, census tracts located on either side of a planned expressway would have evolved similarly.

I leverage this intuition and test for the barrier effect of expressways by running a long-difference empirical specification where I estimate the effect of changes in exposure to Black residents on the change in the racial composition of neighborhoods in Chicago. Areas experiencing the largest increase in exposure to Black residents are expected to see a greater increase in their Black population share over time.

To measure the change in exposure to Black residents, I compute a novel metric of accessibility, defined as a location-specific weighted average of the share of Black residents in each neighborhood. The weights are a decreasing function of bilateral travel time between each origin location and all other neighborhoods in the city. The weights depend on the development of the underlying road network – in particular, the construction of expressways – and assign higher values to locations that are more easily accessible. As a result, this summary measure captures, for each origin neighborhood, changes in exposure to Black residents over time, driven both by residential sorting and by changes in transportation infrastructure that affect how accessible other neighborhoods in the city are.

The estimating equation in first differences is the following:

$$\Delta y_i = \beta_s \Delta S_i + \beta_d Dist.\,expressway_i + City\,side\,FE + \gamma_c' Controls_i + \epsilon_i \tag{2}$$

where $\Delta y_i = y_{i,1990} - y_{i,1950}$ measures the change in the outcome variable (share of Black households, land value) between 1950 and 1990 (in the baseline specification); $\Delta S_i = \sum_{j \neq i} e^{-\rho \tau_{ij\,post}} share\,Black_{j\,post} - \sum_{j \neq i} e^{-\rho \tau_{ij\,pre}} share\,Black_{j\,pre}$ captures the change in exposure to Black residents, induced by both sorting and changes in transportation infrastructure; $Dist.\,expressway$ represents the distance (in km) between the centroid of each census tract and the nearest expressway.[39] City-side fixed effects (for being in the north, west, or south of the city) are always included. Baseline control variables are added sequentially, as stated below each table, to partially account for changes in observables that might be correlated with neighborhood dynamics. Standard errors are clustered at the census tract level unless indicated differently.[40]

[39] Based on the previous reduced-form results, this variable accounts for the disamenity effect of expressways and the identification of the coefficient of interest β_d relies only on the residual variation in the predicted change in exposure. The sample correlation between distance to the nearest expressway and change in exposure to Black areas is -0.28.

[40] To allow for arbitrary spatial correlation, I also report baseline results (i) clustering standard errors at a grid level that partitions the city into 25 equally sized squares; (ii) using Conley (1999) standard errors with a 3 km cutoff (results, not reported here, remain similar at different cutoff values, such as 5 km and 10 km).

A concern in estimating regression (2) is that the change in exposure to Black residents (ΔS) is likely correlated with unobserved shocks to neighborhood residential amenities, which are captured in the error term. This variable is constructed as a weighted average of exposure to Black areas, where weights are a function of the bilateral travel times between any pair of locations. As a result, it incorporates information on the racial composition of neighborhoods in both periods.[41] To address this issue, I instrument for the change in exposure to Black residents holding the racial composition fixed to the baseline period. The instrument hence isolates variation in exposure to Black areas that is only due to changes in travel time between locations, while accounting for sorting effects. Formally, the instrument is constructed as: $\Delta SMA_i = \sum_{j \neq i} share\ Black_{j\ pre}(e^{-\rho\tau_{ij\ post}} - e^{-\rho\tau_{ij\ pre}})$.

In the baseline specification reported in the main text, I calibrate the rate of spatial decay of the weights (ρ) from the literature. I set $\rho = 0.019$ as the estimated elasticity of consumption travel cost with respect to travel times from Miyauchi et al. (2022) – the most pertinent estimate in this context. To isolate the barrier effect, I also set the cost of crossing the expressway network to be infinitely high. As a result, I only capture changes in exposure within neighborhoods located on the same side of the expressway, helping identify the barrier effect. In the Appendix, I report additional results where I increase the parameter ρ to values that effectively set the weights used to compute changes in exposure to Black residents virtually zero beyond distances of 10 km and 20 km ($\rho = 0.167$ and $\rho = 0.083$, respectively). These exposure measures can thus be interpreted as iso-areas, i.e., network-based equivalents of buffer zones. The local exposure measure is a weighted average of the racial composition of neighborhoods within a given distance (10 or 20 km, depending on the specification) from the origin location, when traveling on the road network.

Figure 4, panels (a) and (b), display the change over time in exposure to Black residents (ΔS) and the instrumental variable (ΔSMA), respectively. Census tracts are grouped into deciles based on the value of the respective variable, with darker colors representing higher values. The maps reveal two key patterns. First, both variables exhibit substantial independent variation with distance from the expressway network: Being closer to the road does not necessarily result in larger changes in either variable. Second, they show a high degree of spatial correlation, reflecting the unequal distribution of racial groups across different parts of the city. The IV takes its highest values in the south and its lowest values in the north, consistent with the broader racial geography of Chicago.

To address the concern that the IV may be capturing spatially clustered features within the city, I conduct the following robustness checks. First, I always include region-fixed effects, ensuring that identifying variation comes exclusively from within the city's north, west, and south sides. Second, I introduce a control variable that captures changes in exposure to wealthy neighborhoods. This measure is computed as the weighted average of the share of college graduates in each location (serving as a proxy for neighborhood income levels), similar to the measure of exposure to Black residents.[42] The inclusion of this control accounts for the possibility that changes in neighborhood accessibility may affect not only exposure to Black areas in the city but also access to different sets

[41] To address the correlation between initial residential choices and local unobserved shocks, I exclude the location itself when calculating its predicted change in exposure.

[42] The weights used are the same as those applied in computing exposure to Black residents.

of residential amenities (e.g., parks or restaurants), correlated with resident income.[43] Finally, as mentioned above, I replicate the results using more localized exposure measures by adjusting the weights so that locations beyond 10 or 20 km from each origin effectively receive near-zero weights. These localized exposure measures exhibit lower spatial correlation, mitigating concerns that the findings may be influenced by broader citywide segregation patterns. The corresponding maps are presented at the beginning of Sections C.5 (20 km cutoff) and C.6 (10 km cutoff) in Appendix C.

5.1 Results

Table 1 presents the estimated coefficients from regression (2), where the outcome variable is the change in the share of Black residents. Region-fixed effects for being in the north, west, or south of the city are always included.

OLS results are reported in columns (1) to (7). Column (1) shows the sample correlation between the (standardized) change in exposure to Black residents and the (standardized) change in the share of Black residents in the origin neighborhood. The estimated coefficient is large and statistically significant at the 1% level: a one standard deviation increase in exposure to Black residents is associated with an average 0.496 standard deviation increase in the share of Black residents in the neighborhood. In column (2), I include distance to the nearest expressway (in km) as an additional regressor. The estimated coefficient for exposure to Black residents remains stable.

Columns (3) to (5) sequentially introduce additional controls to partially account for changes in observables that might correlate with changes in neighborhood composition. In column (3), I control for basic census tract characteristics: a quadratic polynomial of distance to the central business district (CBD), land area, and distance to the water. Column (4) extends the set of controls to capture historical conditions.[44] In column (5) I further control for changes in exposure to wealthy areas in the city (ΔY).[45] The point estimate remains largely stable after the inclusion of additional controls. The estimated coefficient for distance to the expressway reveals a strong negative relationship between the share of Black residents in a neighborhood and distance from the road: the share of Black residents decreases by approximately 0.2 standard deviations for every additional kilometer from the expressway, on average, holding all else fixed.

In columns (8) to (11), I instrument for the change in exposure to Black residents while holding the racial distribution within the city fixed to the pre-period. The estimated coefficient of ΔS captures what would be expected if accessibility to other parts of the city changed over time while the racial composition of all other neighborhoods remained constant. The variation isolated by the instrument can be interpreted as the immediate effect of introducing an uncrossable barrier in the city before residents adjust their location decisions. The estimated coefficient remains large and statistically

[43] This ensures that the estimated effects of exposure to Black residents are not confounded by changes in access to higher-income neighborhoods and their amenities.

[44] The complete set of historical controls includes distance to railroads in 1898, HOLC grade, historical outcomes in levels, and the change in population density between 1920 and 1940. Since the city was smaller at the beginning of the sample period, including historical controls reduces the sample size from 764 to 727 observations.

[45] Similar to ΔS, the variable is constructed as $\Delta Y_i = \sum_{j \neq i} e^{-\rho \tau_{ij\,post}} c_{j\,post} - \sum_{j \neq i} e^{-\rho \tau_{ij\,pre}} c_{jpre}$ where c_j is the share of college graduates in neighborhood j.

significant at the 1% level ($\beta = 0.158$). As expected, the point estimate is lower than in the OLS specification, confirming that the OLS estimates were largely inflated by sorting effects. In column (11), I additionally instrument for the change in exposure to wealthy neighborhoods in the city using an instrumental variable (ΔYMA) constructed analogously to ΔSMA.[46] Both the point estimates for the change in exposure to Black residents and the point estimate for distance to the expressway remain stable and highly statistically different from zero. Notably, the estimated coefficient for the change in exposure to wealthy areas (ΔY) is not statistically different from zero in the estimated regressions.

Next, Table 2 presents the estimated coefficients from regression (2), where the outcome variable is the change in land value – used as a proxy for neighborhood valuation – between 1950 and 1990. On average, an increase in exposure to Black residents reduces land value by approximately 0.3 standard deviations (s.e. 0.051) in the regression with the full sets of controls in column (5). After accounting for sorting effects, the results remain largely stable in the IV specifications, though smaller in magnitude (around -0.25 in column 11). In addition, the results show that the estimated coefficient for distance to the expressway is, on average, positive, but statistically different from zero only in the IV specifications. Point estimates above zero are consistent with the idea that expressways generate disamenity effects within cities.

To summarize, this estimation exercise yields four main takeaways. First, a higher exposure to Black residents increases the likelihood that a neighborhood becomes more Black over time and decreases its valuation in the long run. Second, changes in exposure to wealthy areas do not appear to affect a neighborhood's racial composition, after controlling for distance to the expressway and changes in exposure to Black residents. Third, both the disamenity and the barrier effect of expressways affect the racial distribution within the city. Fourth, land value appears to be more responsive to changes in neighborhood composition induced by the barrier effect than to the disamenity effect of expressways.

5.1.1 Long-difference results over time

This section reports the results of long-difference regressions estimated between 1950 and 1980 and between 1950 and 2000. One limitation in conducting these exercises is the lack of complete snapshots of the road network over time. I observe the road network only as of 2019 for the post-period and as of 1940 for the pre-period. As a result, the weights used to compute changes in exposure to Black areas in the city remain the same as in the baseline specification, while the changes in sorting and outcomes reflect neighborhood evolution over time. Despite this limitation, these additional results provide insight into neighborhood dynamics and the persistence of the estimated effects.

Results are presented in Sections C.3 and C.4 in Appendix C. On average, the estimated barrier effects of expressways remain quantitatively similar across periods. While the magnitudes of the estimated coefficients are not directly comparable – since they are obtained from different regressions

[46] $\Delta YMA_i = \sum_{j \neq i} c_{j\,pre}(e^{-\rho\tau_{ij\,post}} - e^{-\rho\tau_{ij\,pre}})$, where c_j is the share of college graduates living in neighborhood j. This instrument holds sorting fixed to the pre-period and isolates variation in exposure to rich areas induced solely by changes in travel times.

– the relative importance of key determinants of neighborhood change remains consistent over time. Both exposure to Black areas and distance from expressways continue to be significant predictors of changes in neighborhood racial composition and valuation. However, over time, the relative importance of proximity to expressways declines compared to exposure to Black areas in explaining changes in both Black share and land value.

5.2 Robustness checks

I conduct several robustness checks. First, I allow for arbitrary spatial correlation of errors across census tracts within the same grid cell or within a certain distance from each other. Second, I restrict the sample to census tracts that had fewer than 20% Black residents at the beginning of the century (95% of the data) to address concerns that a few historically Black neighborhoods might be driving the results. Third, I rerun the analysis using only census tracts located more than 5 km from the central business district (this exercise removes 12% of the data). The results remain robust to all these specifications. Fourth, in Appendix C, I report results using more localized exposure measures and show that the main findings still hold. Finally, I provide additional evidence using house value and college share as alternative outcomes. Below, I describe these robustness checks in more detail.

The remaining columns in Table 1 and Table 2 report the results of the robustness checks. Columns (6) and (7) present the results for the specification with the full set of controls, accounting for potential spatial correlation of errors. Columns (6) display the standard errors under the assumption that errors are spatially correlated within grid cells (for this exercise, I partitioned the city into 25 squares). In columns (7), I use Conley (1999) standard errors to allow for arbitrary spatial correlation of errors across tracts within 3 km of each other (results remain similar when using alternative cutoffs, e.g., 5 km and 10 km). The point estimates of interest remain highly statistically different from zero.

Columns (9) in Tables 1 and 2 present the results after removing the 5% of the sample that historically had a high concentration of Black residents. This exercise addresses concerns that the main results may be driven by a few census tracts that were already predominantly Black at the beginning of the twentieth century. At that time, the city was highly segregated, with Black neighborhoods concentrated in the south (and, to a lower extent, in the west). Reassuringly, the point estimates in both tables remain stable.

Columns (10) report the estimated coefficients after removing census tracts within 5 km of the central business district. This exercise helps account for the potential effects of gentrification in downtown areas and, more broadly, demonstrates that the results are not solely driven by central areas, which may have undergone recent redevelopments. The results become slightly stronger in the regression where the share of Black residents is the outcome, consistent with the idea that gentrification may influence the racial composition of neighborhoods in the city center. Similarly, the results for land value become stronger in absolute terms, suggesting that gentrification may partially offset negative neighborhood effects.

Appendix Sections C.5 and C.6 report results using the local exposure measures described above, with 20 km and 10 km cutoffs, respectively. The maps at the beginning of each section display changes in the regressor of interest (ΔS) and the IV. In all cases, the maps exhibit greater spatial variation than those relative to the primary analysis. This is because the weights for both types of local measures are set so that only areas within a certain (network-based) distance from each origin location receive high weights.

The increased spatial variation helps alleviate concerns that both the IV and the endogenous regressor may capture unobserved characteristics shared among geographically clustered units, which could influence neighborhood dynamics. At the same time, however, these variables are also more susceptible to bias from local unobservable characteristics. In the OLS specifications, the correlation between exposure to Black residents and the change in the neighborhood's Black share is even stronger. On average, across specifications, a one standard deviation increase in exposure to Black residents is associated with a high 0.8 standard deviation increase in outcome. Comparing these results to the even larger point estimates using a 10 km cutoff suggests that the racial composition of nearby neighborhoods strongly influences the racial mix of an area. However, it is reassuring to observe that when sorting is held fixed to the pre-expressway period, the point estimates drop substantially and become closer to those in the baseline specification. The estimated coefficients after instrumenting for the change in exposure to Black neighborhoods drop to approximately 0.2-0.3 standard deviations. In addition, the estimated effects of instrumented localized changes in exposure to Black residents on land value are qualitatively in line with the main results. However, the point estimates from the IV regression in the last column become strongly attenuated after instrumenting for changes in exposure to wealthy areas in both the 20 km and 10 km cutoff specifications.

Finally, in Appendix C.7, I estimate the barrier effect on two additional outcomes: house value and college share, which serve as proxies for changes in the relative valuation of neighborhoods. In both cases, the results indicate that house value and college share respond strongly to changes in exposure to wealthy areas in the city but remain unaffected by changes in exposure to Black areas once exposure to wealthy neighborhoods is controlled for.

5.3 Discussion of the barrier effect of expressways

The second empirical design evaluates the barrier effect of expressways. I find that higher exposure to Black areas increases the likelihood that a neighborhood becomes more Black over time. The effect is sizable: depending on the specification, a one standard deviation increase in exposure to Black residents is associated with a 0.16-0.20 standard deviation increase (in the IV specifications) in the share of Black residents living, on average. The results remain stable after including a rich set of controls, such as distance to the nearest expressway and a measure of changes in exposure to wealthy areas. Next, I estimate the barrier effect of expressways on changes in land value between 1950 and 1990, using land value as a proxy for neighborhood valuation. An increase in exposure to Black residents reduces land value by 0.24-0.32 standard deviations in the long run. Taken together, these findings indicate that both the disamenity and the barrier effects of expressways affect the racial distribution within the city. However, land value appears to respond more strongly to changes in neighborhood composition induced by the barrier than to the disamenity of expressway proximity.

6 A quantitative spatial urban model with racial preferences

The theoretical framework follows the canonical Alonso (1964), Muth (1969), and Mills (1967) monocentric city model, incorporating an internal city structure à la Ahlfeldt et al. (2015). I consider a city embedded within a wider economy. The city consists of discrete locations indexed by $j = 1, ..., J$. For simplicity, the time subscript is omitted, but all expressions hold in each period. Land K_j is allocated exclusively to residential use and is supplied by a competitive floorspace sector. The city is populated by an endogenous number of residents N^o of one of four types, indexed by $o \in \{WH, WL, BH, BL\}$ – a two-by-two classification based on race (B Black, W white) and educational attainment (H high-educated, L low-educated). Agents are perfectly mobile within the city and the wider economy. The outside option of living outside the city provides a reservation level of utility $\bar{U}^o$ for type o. Individuals decide whether to move to the city or not before observing idiosyncratic utility shocks for each location within the city. If an individual chooses to move to the city, they observe the realization of idiosyncratic utility and choose the residential location that maximizes their utility. Locations differ in residential amenities, residential land availability, and access to the transport infrastructure, which determines travel times between locations.

6.1 Preferences

The city is populated by an endogenous number of residents in each period, N^o, belonging to one of four types, $o \in \{WH, WL, BH, BL\}$. Individuals derive utility from residential amenities, a consumption index, and a location-specific idiosyncratic shock unique to each individual. This shock captures the idea that individuals may have private preferences for living in a specific neighborhood.

The utility of individual ω of type o living in j is given by:

$$U(\omega)^o_j = B^o_j C(\omega)^o_j z(\omega)_j \tag{3}$$

Residential amenities B^o_j capture common features that make a location a more or less desirable place to live.[47] The consumption index $C(\omega)^o_j$ depends on the consumption good $c(\omega)^o_j$ which is chosen as numeraire and residential land $L(\omega)^o_j$.

The consumption index $C(\omega)^o_j$ is assumed to take the Cobb-Douglas form:[48]

$$C(\omega)^o_j = \left(\frac{c(\omega)^o_j}{\alpha}\right)^{\alpha} \left(\frac{L(\omega)^o_j}{1-\alpha}\right)^{1-\alpha} \tag{4}$$

Individual heterogeneity is modeled as in structural urban models following McFadden (1974). For each worker ω of type o living in j, the idiosyncratic component of utility ($z(\omega)_j$) is drawn from a common independent Fréchet distribution:

[47] This assumption allows different types to place different values on amenities, reflecting heterogeneous preferences by race and education.

[48] In the baseline model, I assume that all types allocate the same share of income to residential floorspace, as captured by the common term α. This assumption could, however, be relaxed to allow for type-specific expenditure shares.

$$F(z(\omega)_j) = e^{-T_j(z(\omega)_j)^{-\epsilon}} \tag{5}$$

where the scale parameter $T_j > 0$ determines the average utility from living in j and the shape parameter $\epsilon > 1$ controls the dispersion of the idiosyncratic utility.[49]

Once the idiosyncratic utility for each residence location is revealed, households choose where to live to maximize their utility, taking residential amenities, prices, and the location decisions of other households as given. As a result, households sort into the city based on location characteristics and their idiosyncratic preference shocks.

I assume that all households supply one unit of labor in the Central Business District (C) in exchange for a type-specific wage w^o. All jobs are located in the city center. Labor is used to produce a final good that is exclusively traded in external markets, and the full revenues are shared between absentee entrepreneurs.[50] Commuting to work is costly and depends on the mode of transportation, which is assumed to vary by education level (denoted by the subscript H for highly educated and L for low-educated individuals). The effective wage of a household of type o residing in j is equivalent to $\frac{w^o}{d_{jC}^{H,L}}$. $d_{jC}^{H,L} = e^{\kappa \tau_{jC}^{H,L}}$ is the iceberg commuting cost, which increases with the travel time (τ_{jC}) between the location of residence (j) and employment (C).[51] Finally, the parameter κ controls the size of the commuting costs.

The indirect utility from living in j can hence be expressed in terms of the common component of amenities, effective wage, floorspace prices, and the idiosyncratic shock:

$$u(\omega)_j^o = \frac{B_j^o w^o R_j^{\alpha-1} z(\omega)_j}{d_{jC}^{H,L}} \tag{6}$$

Given that $z(\omega)_j$ is Fréchet distributed, also the indirect utility follows a Fréchet distribution. As a result, the probability that an individual ω of type o lives in j is given by:

$$\begin{aligned} \pi_j^o &= P(u(\omega)_j^o \geq max\, u(\omega)_j^o \ \ \forall j) \\ &= \frac{T_j(R_j^{1-\alpha})^{-\epsilon}(B_j^o)^{\epsilon}(w^o/d_{jC}^{H,L})^{\epsilon}}{\sum_s T_s(R_s^{1-\alpha})^{-\epsilon}(B_s^o)^{\epsilon}(w^o/d_{sC}^{H,L})^{\epsilon}} = \frac{\Phi_j^o}{\Phi^o} \end{aligned} \tag{7}$$

Because of idiosyncratic shocks to preferences, these residential probabilities imply that individ-

[49] In principle, I could allow idiosyncratic preference distributions to vary by type, both in terms of the average idiosyncratic preferences for amenities in a given location (the scale parameter T_j) and the variance of preferences across locations (captured by the shape parameter ϵ). This would introduce an additional source of heterogeneity in the model. For example, if different types exhibit varying degrees of dispersion in their preferences, they may respond differently to the same commuting cost shock. A group with less dispersed idiosyncratic preferences would be less affected by a change in commuting costs, as their location choices are more rigid. In practice, I assume a common Fréchet distribution for all types, as is standard in the literature. Estimating type-specific shape parameters would require data on type-specific commuting flows, which are often unavailable, including in this setting.

[50] I also assume that firms in the CBD do not occupy any physical space (i.e., land can be only allocated for residential use). These assumptions are necessary since workplace location decisions and land allocation for commercial as opposed to residential use are not observed.

[51] Travel time is measured in minutes and depends on the underlying transport network.

uals of a given type choose different residential locations when faced with the same prices or location characteristics. In particular, individuals are more likely to live in location j, the more attractive its residential amenities are (B_j), the higher its average idiosyncratic utility as determined by T_j, the lower the floorspace prices R_j, and the higher the effective wage (or equivalently, the lower the commuting cost $d_{jC}^{H,L}$). It should be noted that the denominator is type-specific but not location-specific: it indeed measures the expected utility of living in the city.

As an illustration, if we take the ratio between the probability that type o lives in j against the probability of type m living in j, we find:

$$\begin{aligned}\frac{\pi_j^o}{\pi_j^m} &= \frac{T_j(R_j^{1-\alpha})^{-\epsilon}(B_j^o)^{\epsilon}(w^o/d_{jC}^{H,L})^{\epsilon}/\Phi^o}{T_j(R_j^{1-\alpha})^{-\epsilon}(B_j^m)^{\epsilon}(w^m/d_{jC}^{H,L})^{\epsilon}/\Phi^m} \\ &= \left(\frac{B_j^o}{B_j^m}\right)^{\epsilon}\left(\frac{w^o/d_{jC}^{H,L}}{w^m/d_{jC}^{H,L}}\right)^{\epsilon}\left(\frac{\Phi^o}{\Phi^m}\right)^{-1}\end{aligned} \tag{8}$$

That is, a worker of type o is more likely to live in j relative to a worker of type m; the higher her valuation of residential amenities, the higher her effective wage, and the lower the utility she gets from living outside the city. When both types have the same education level, both of the last two terms become constant across locations. This implies that they equally affect the relative probabilities of types o and m of living in each location j. However, since the two types value amenities differently, the relative probability of them living in j depends on the extent of their relative evaluation for amenities in j, captured by the first term on the right-hand side of the expression.

Given that in this setting, the employment location is assumed to be fixed to the CBD, there is no uncertainty concerning the wage that a household receives when choosing to live in location j: the expected worker income conditional on living in j is indeed simply equal to $\mathbb{E}(w^o|j) = w^o$. As a result, effective wage (i.e., net of commuting costs) is higher in residence locations with low commuting costs to the downtown area.

Finally, population mobility for each type implies that the expected utility from moving to the city is equal to the reservation level of utility in the wider economy ($\bar{U}^o$):

$$\mathbb{E}[u^o] = \gamma\left[\sum_s T_s(R_s^{1-\alpha})^{-\epsilon}(B_s^o)^{\epsilon}(w^o/d_{sC}^{H,L})^{\epsilon}\right]^{1/\epsilon} = \bar{U}^o \tag{9}$$

where the expectation is taken over the distribution of the idiosyncratic component of utility; $\gamma = \Gamma(\frac{\epsilon-1}{\epsilon})$ and $\Gamma(.)$ is the Gamma function.

Another implication of the distribution of utility being Fréchet is that residence locations with attractive features attract more residents on the extensive margin until the expected utility in each location is equalized. This means high amenities in a location increase the utility of a resident with a given idiosyncratic realization of utility z, thus increasing the expected utility from living in that location. At the same time, high amenities also attract individuals with lower realizations of the idiosyncratic utility z, reducing the expected utility from living in j. With a Fréchet distribution of utility, these two forces cancel each other.

6.2 Land market clearing

Residential land market clearing requires that the demand for residential land $D(L_j)$ equals the supply of residential land $S(L_j)$ in each location. I follow a standard approach in the literature and assume that floorspace L is supplied by a competitive construction sector that uses land K and capital M as inputs. The production function takes the Cobb-Douglas form $L_j = M_j^{\mu} K_j^{1-\mu}$ (Combes et al., 2014; Epple et al., 2010; Ahlfeldt et al., 2015). Since the price of capital is the same across locations, the relationship between quantities of floorspace and land can be summarized as $L_j = S(L_j) = \phi_j K_j^{1-\mu}$, where $\phi_j = M_j^{\mu}$ is the density of development (Ahlfeldt et al., 2015).

From the households' maximization problem, the demand for residential land in location j is equal to:

$$D(L_j) = \mathbb{E}[L_j] N_j = \frac{(1-\alpha)\bar{W}_j N_j}{R_j} \tag{10}$$

where $\bar{W}_j = (1/N_j) \sum_o (w^o / d_{jC}^{H,L}) N_j^o$.

Land market clearing is hence equal to:

$$\phi_j = \frac{(1-\alpha)\bar{W}_j N_j}{R_j K_j^{1-\mu}} \tag{11}$$

In estimation, the term ϕ_j (unobserved density of development) is a structural residual that guarantees that floorspace market clearing exactly holds in each location, given observed data and recovered amenities.

6.3 Equilibrium

An equilibrium of the model is characterized by the assignment of type-specific residents to neighborhoods and of a vector of floor space prices, such that the land market clears in each location and no individual has the incentive to deviate by moving to a different location. Given the model's parameters $\{\alpha, \epsilon, \kappa, \lambda_{B,W}^o, \rho^o, g^o, \eta, \mu\}$, the vectors of exogenous location characteristics $\{T, B, n, \tau^{H,L}\}$, the reservation level of utility $\bar{U}^o$ and wage w^o for each type $o \in \{WH, WL, BH, BL\}$, the general equilibrium of the model can be referenced by the vectors $\{R_j, \pi_j^o\}$ and by type-specific population scalars N^o.

The following system of nine equations determines the nine elements of the equilibrium vector: population mobility conditions for each type ($\times 4$), the residential choice probability for each type ($\times 4$), and the land market clearing condition. Ahlfeldt et al. (2015) provide proof for the existence and uniqueness of an equilibrium (uniqueness can be established only in the case of exogenous location characteristics). Given the model's parameters, recovered unobserved fundamentals, and starting conditions, I use an iterative procedure to find the equilibrium values of residence and floorspace prices (see Appendix D.3 for details).

7 Estimation

The estimation proceeds in three steps. First, I calibrate the necessary objects (e.g., commuting elasticity, wages).[52] Then, I invert the model in one year to recover overall amenities that perfectly rationalize the observed distribution of population in the city as being an equilibrium of the model. Finally, I use the shock to the urban structure induced by the construction of expressways in the city to estimate the parameters of interest (i.e., racial preferences and disamenity parameters) that best fit the change in the distribution of population within the city (subject to orthogonality conditions).

Before outlining the steps in more detail, I describe the data used for the estimation.

7.1 Data

Three sets of data are required for the quantitative analysis of the model: residence by type, the price of floor space, and the traveling times between locations. I collect this information for Chicago for the periods before and after expressways were built (around the 1940 and 1990 census periods).

For neighborhood demographics, the two primary sources of data are the 1934 Special Census of Chicago and the 1990 Census of Population and Housing.[53] Differently from the censuses administered around those years, the 1934 census of Chicago reports census-tract-level information on race by educational attainment. Hence, employing this data source offers a significant advantage by providing reliable granular information on educational attainment by race before expressways were built. I then classify individuals according to their race (Black versus the residual category of non-Black) and education level (above versus below the city-wide median level in each period).[54] Normalizing the data consistently to 2010 census-tract boundaries, the information is available for the 791 census tracts that cover the city of Chicago.

I complement this information with the 1940 and 1990 land value data from Olcott's Land Value Blue Book (available thanks to the works of Ahlfeldt and McMillen 2014; 2018). I keep only census tracts with non-missing land information, which leads to a final sample of 767 observations.[55]

Finally, I compute commuting times between each pair of locations in the sample. I rely on the 1940 road transportation network, available from the Urban Transition HGIS Project (Shertzer et al., 2016)[56] and on the contemporaneous road transportation network from the US Census Bureau. Travel times are measured in minutes based on the transportation network available in each period.

[52] In the urban models where commuting flows are observed in the data, the commuting elasticity can be estimated from a gravity equation of commuting flows, derived from the household's probability of living and working in different neighborhoods in the city. Exploiting the recursive structure of the model, location-specific wages, usually difficult to observe systematically, are recovered by solving the labor market clearing conditions in one period and obtaining inverted wages (generally up to a transformation).

[53] The 1934 census was conducted by the Chicago Census Commission (and not the US Census Bureau) to know (as reported by the then Major Kelly) "exactly what had been the effects of the depression upon changes of residences, occupancy of dwellings, housing needs, health of the people, etc." (Newcomb and Lang, 1934).

[54] In 1934, the median educational attainment level corresponds to completing grades 5-8. In 1990, it corresponds to completing high school.

[55] The 1990 sample consists of 766 observations, because of the redevelopment in the 1960s of the Midway Airport, located in one of the central census tracts.

[56] https://s4.ad.brown.edu/Projects/UTP2/ncities.htm

I assume average travel speeds for each mode of transport. I allow travel times to differ between low and high-educated individuals by assigning them different modes of transport.[57]

7.2 Step 1: Calibrate commuting elasticity and wages

From the commuting choice probability (the equivalent of the residential choice probability (7) introduced here, if both residence and workplace location choices were observed), it would be possible to recover the commuting elasticity in a gravity equation framework.[58] In this setting, however, "workplace location" is not observed and is hence assumed to be the same for each individual (corresponding to the central business district). As a result, the vector of distances from residence location to workplace location (d_{jC}) varies at the origin level only – it is multicollinear to origin fixed effects. In turn, the commuting elasticity cannot be consistently estimated in a gravity-type framework. As a result, I calibrate the commuting elasticity from the relevant literature. I set the Fréchet parameter to $\epsilon = 6$, corresponding to the central value found in the literature (Miyauchi et al., 2022), and the spatial decay parameter for commuting costs to $\kappa = 0.01$ (Ahlfeldt et al., 2015).

In quantitative spatial urban models with data on both residence and workplace location decisions (like in Ahlfeldt et al., 2015), (unobserved) workplace-specific wages are recovered by exploiting the recursive structure of the model after estimating the commuting elasticity. Using the resulting model's parameters and data, wages can be obtained from the commuting market clearing condition.[59] Lacking information that would allow me to recover wages, I calibrate them using the information at the national level each period, adjusted for the distribution of race by educational attainment observed in the city.[60]

7.3 Step 2: Recover overall amenities

From the expression of the residential choice probabilities (7), after multiplying both sides by the city's total number of residents of type o (N^o), and considering that $\pi_j^o N^o = N_j^o$, I get:

$$N_j^o = \frac{T_j (R_j^{1-\alpha})^{-\epsilon} (B_j^o)^{\epsilon} (w^o / d_{jC}^{H,L})^{\epsilon}}{\sum_s T_s (R_s^{1-\alpha})^{-\epsilon} (B_s^o)^{\epsilon} (w^o / d_{sC}^{H,L})^{\epsilon}} N^o \tag{12}$$

Since T_j enters the model isomorphically (and it cannot be separately identified from the overall residential amenities in the data), I define the following composite variable $\tilde{B}_j^o = T_j^{\frac{1}{\epsilon}} B_j^o$. To simplify the exposition, I also define $W_j^o = (w^o / d_{jC}^{H,L})^{\epsilon}$, so that:

[57] I consider highly educated individuals to move only by car in both periods. Low-educated individuals move exclusively by bus in 1940. In 1990, they are assumed to move by bus with 0.75 probability and by car with 0.25 probability. Further details are reported in Appendix D.1.1.

[58] The estimation follows from regressing commute flows on commute times and origin and destination fixed effects.

[59] The commuting market clearing condition ensures that the total number of workers in a location equals the number of commuting residents that choose to work in that location and commute from every residence location.

[60] The calibrated wages in the pre-period ($2010) are as follows: $12,605 for high-educated Black; $11,020 for low-educated Black; $17,738 for high-educated white; and $12,197 for low-educated white. In the post-period, I find ($2010): $30,766 for high-educated Black; $12,790 for low-educated Black; $38,924 for high-educated white; and $17,153 for low-educated white. For details on the full procedure and data sources, see Appendix D.1.2.

$$N_j^o = \frac{(R_j^{1-\alpha})^{-\epsilon}(\tilde{B}_j^o)^{\epsilon}W_j^o}{\sum_s (R_s^{1-\alpha})^{-\epsilon}(\tilde{B}_s^o)^{\epsilon}W_s^o}N^o \tag{13}$$

The model can then be calibrated to recover unique adjusted location fundamentals ($\tilde{B}_j^o$), given known values of the model's parameters $\{\alpha, \epsilon, \kappa\}$ and the observed data $\{R, N_j^o, \tau^{H,L}\}$. That is, there is a unique mapping between the model parameters and the observed data to the overall residential location characteristics (up to a normalization). Dividing N_j^o by its geometric mean on both sides and noticing that all constant terms cancel out, I get:[61]

$$\frac{N_j^o}{\bar{N}^o} = \left(\frac{R_j}{\bar{R}}\right)^{-(1-\alpha)\epsilon}\left(\frac{\tilde{B}_j^o}{\bar{\tilde{B}}^o}\right)^{\epsilon}\frac{W_j^o}{\bar{W}^o} \tag{14}$$

Inverting the system, I hence find an expression of overall residential amenities only as a function of data, model's parameters, and calibrated values:

$$\frac{\tilde{B}_j^o}{\bar{\tilde{B}}^o} = \left(\frac{N_j^o}{\bar{N}^o}\right)^{\frac{1}{\epsilon}}\left(\frac{R_j}{\bar{R}}\right)^{1-\alpha}\left(\frac{W_j^o}{\bar{W}^o}\right)^{-1/\epsilon} \tag{15}$$

The maps plotting the distribution of recovered overall amenities by type (the left-hand side of equation 15) after inverting the model in each period are reported in Appendix D.2. Darker colors correspond to higher amenity levels.

Total recovered amenities consist of both an exogenous part (that depends on the physical attributes of the place) and an endogenous part (that depends on sorting and residential choices), as described in more detail below. As a result, they appear to be higher in proximity to the shore and lower in the vicinity of industrial corridors. At the same time, there is a high degree of heterogeneity both over time and between races, showing a polarization of the white population towards the northern neighborhoods by 1990 and the Black population favoring locations in the south.[62]

7.4 Step 3: Structural estimation

Before outlining the procedure to structurally estimate the parameters of interest, I allow total residential amenities to depend on residential fundamentals, the disamenity of expressways, and the demographic composition of the neighborhoods. First, residential fundamentals ($\tilde{b}_j^o$) capture features of the physical geography that make a location a more or less attractive place to live (e.g., the presence of green areas or proximity to the shore). Second, residential amenities are affected by proximity to the expressways. Following Brinkman and Lin (2022), I assume that the disamenity of expressways

[61] Denote the geometric mean of a variable X by $\bar{X}$, then: $\bar{X} = (\Pi_{s=1}^{S} X_s)^{\frac{1}{s}}$. It should be noted that the geometric mean of a product equals the product of the geometric means of its terms. Following Ahlfeldt et al. (2015), I choose the unit in which to measure adjusted residential amenities such that the geometric mean of adjusted residential amenities is equal to 1, i.e. $\bar{\tilde{B}}_j^o = (\Pi \tilde{B}_s^o)^{1/S} = 1$ for each type o.

[62] For instance, in 1940, amenities were high for white high-educated residents along most of the southern shore, but by 1990 only the area that constitutes the campus of the University of Chicago is deemed as high-amenity location for this group.

is a function of distance to the road. It is modeled as follows:

$$E_j^o \equiv 1 - g^o e^{-\eta dist_j} \tag{16}$$

where $dist_j$ is the distance to the closest expressway; the parameter g^o governs the size of the disamenity, and η its spatial attenuation.[63]

Third, I capture residential externalities imposing structure on how the amenities in a given location are affected by the demographic characteristics of the other locations. Specifically, I model racial preferences as follows:

$$\Omega_j^o = \left(\sum_{i \neq j} e^{-\rho^o \tau_{ji}^{H,L}} \frac{N_i^B}{K_i} \right)^{\lambda_B^o} \left(\sum_{i \neq j} e^{-\rho^o \tau_{ji}^{H,L}} \frac{N_i^W}{K_i} \right)^{\lambda_W^o} \tag{17}$$

Residential externalities are hence modeled as a power function of distance-weighted exposures to Black and white residents living in every other neighborhood in the city, weighted by a function of the traveling times between locations. ρ^o captures the rate of spatial decay at which racial preferences matter, whereas $\lambda_{B,W}^o$ are the elasticities of amenities with respect to the concentration of nearby Black and white residents.

Finally, I include a term that captures local spillovers as a function of the share of adults with above-median education.[64] The term is intended to capture the value of the endogenous amenities that correlate with higher income (proxied by higher educational attainment), like safety, quality of schools, and the general level of public good provision. It is modeled as follows:

$$H_j = \frac{N_j^H}{N_j} \tag{18}$$

As a result, overall adjusted amenities consist of:

$$\tilde{B}_j^o = \tilde{b}_j^o H_j E_j^o \Omega_j^o \tag{19}$$

Together with equation (15), I get:

$$\frac{\tilde{b}_j^o}{\bar{\tilde{b}}^o} = \left(\frac{N_j^o}{\bar{N}^o} \right)^{\frac{1}{\epsilon}} \left(\frac{R_j}{\bar{R}} \right)^{1-\alpha} \left(\frac{W_j^o}{\bar{W}^o} \right)^{-1/\epsilon} \left(\frac{H_j}{\bar{H}} \right)^{-1} \left(\frac{E_j^o}{\bar{E}^o} \right)^{-1} \left(\frac{\Omega_j^o}{\bar{\Omega}^o} \right)^{-1} \tag{20}$$

where the term on the left-hand side corresponds to location fundamentals (structural residuals of the model). As it is clear from the above expression, they are a function of data and model parameters only. I can then solve for these structural residuals for the entire city before and after the shock induced by expressways. I denote the change over time by Δ. Following Ahlfeldt et al. (2015), I further assume that these structural residuals consist of both a time-invariant fixed component ($\tilde{b}_j^{oF}$)

[63] Note that g^o is type specific, but η is fixed, since it governs the rate at which expressway disamenities (like pollution and noise) decay in space – which depends on environmental factors.

[64] Fogli and Guerrieri (2019) propose the share of college graduates. In this setting, I use the share of adults with above median education since it is an endogenous variable of the model.

and a time-varying stochastic shock ($\tilde{b}^{oV}_{j}$). With a first difference estimator, the time-invariant fixed effects are differenced out so that after taking logs of both sides (and making now time explicit with the subscript t), I get:

$$\Delta \ln\left(\frac{\tilde{b}^{oV}_{jt}}{\bar{\tilde{b}}^{oV}_{t}}\right) = \frac{1}{\epsilon}\Delta \ln\left(\frac{N^{o}_{jt}}{\bar{N}^{o}_{t}}\right) + (1-\alpha)\Delta \ln\left(\frac{R_{jt}}{\bar{R}_{t}}\right) - \frac{1}{\epsilon}\Delta \ln\left(\frac{W^{o}_{jt}}{\bar{W}^{o}_{t}}\right) - \Delta \ln\left(\frac{H_{jt}}{\bar{H}_{t}}\right) - \Delta \ln\left(\frac{E^{o}_{jt}}{\bar{E}^{o}_{t}}\right) - \Delta \ln\left(\frac{\Omega^{o}_{jt}}{\bar{\Omega}^{o}_{t}}\right) \tag{21}$$

These structural residuals correspond to double-difference adjusted residential fundamentals: the first difference is before-after (over time), whereas the second difference is across locations within the city, and it is reflected in the normalization relative to the geometric mean, computed by dividing each term by the geometric mean of the variable in each period before taking logs. This second difference eliminates fixed effects that are common across locations within each period (like the reservation level of utility). Normalizing relative to the geometric mean is also advantageous in that this second difference ensures that the results are invariant to the choice of the units in which residential fundamentals are measured (since the choice is common to all units within one time period). By construction, it hence follows that the mean changes in (log) residential fundamentals are equal to zero.

The parameters of interest are, on the one hand, the ones governing the disamenity of expressways $\{g^{o}\}$ and, on the other hand, the ones governing racial preferences $\{\lambda^{o}_{B,W}, \rho^{o}\}$.

7.4.1 Moment conditions

To estimate the parameters of interest, I use analogous sources of variation as the ones employed in the reduced form analyses. The first set of moment conditions imposes that the changes in adjusted residential fundamentals in (21) are uncorrelated with the exogenous change in exposure to Black areas in the city induced by the construction of urban barriers. Similar to the reduced form results, I capture the exogenous change in the exposure to Black areas as follows. I compute distance grid cells from the centroid of the area with a historically high concentration of African Americans (the so-called Black Belt), interacted with whether or not the location is separated from the Black Belt by the presence of an expressway. The second set of moment conditions, used to isolate the disamenity parameters, imposes orthogonality conditions between changes in adjusted residential fundamentals and distance to the road. I further interact them with distance grid cells from the CBD to capture the idea that locations closer to the CBD tend to be more severely affected by proximity to expressways since the access benefits are lower. The two sets of moment conditions are as follows:

$$\begin{aligned} &\mathbb{E}[\mathbb{I}bb_k \times \mathbb{I}Barrier \times \Delta \ln(\tilde{b}^{o}_{jt}/\bar{\tilde{b}}^{o}_{t})] = 0 \\ &\mathbb{E}[\mathbb{I}cbd_{k'} \times \mathbb{I}exp_{k''} \times \Delta \ln(\tilde{b}^{o}_{jt}/\bar{\tilde{b}}^{o}_{t})] = 0 \end{aligned} \tag{22}$$

where $\mathbb{I}bb_k$ for $k \in \{1, ..., K_{\mathbb{I}bb}\}$ are indicator variables for distance grid cell k from the centroid of the Black Belt; $\mathbb{I}Barrier$ is an indicator for whether or not there is an expressway separating the location from the Black Belt; $\mathbb{I}cbd_{k'}$ for $k' \in \{1, ..., K'_{\mathbb{I}cbd}\}$ are indicator variables for distance grid cell k' from the CBD; $\mathbb{I}exp_{k''}$ are indicator variables for distance to the expressway network. I use 20 indicator variables based on percentiles of distance to the Black Belt; four indicator variables for distance to the CBD, and three for distance to the expressways.[65]

Following the construction of expressways in the city, residential patterns may change. These changes can happen for two reasons. On the one hand, the forces captured in the model. For instance, since expressways increase the effective distance between locations situated on opposite sides of the road, following the reduced-form results, we may expect neighborhood demographics to change in response. On the other hand, residential fundamentals may change over time for reasons unrelated to the construction of expressways in the city. The moment conditions impose restrictions on how residential fundamentals can change – i.e., changes should be mean zero in each distance bin interaction.

7.4.2 Identification

I use the Generalized Method of Moments (GMM) with the moment conditions above to estimate the parameters of interest. Additional details on the procedure are reported in Appendix D.4. Those moment conditions can be used to identify the model's unknown parameters and, simultaneously, to recover the unobserved residential fundamentals (structural residuals). Equation (20) shows closed-form solutions for the structural residuals: they are only functions of the model's parameters and observed data. In principle, the moment conditions need not uniquely identify the model's parameters – there could be multiple local minima that correspond to different combinations of the unknown parameters and structural residuals that are consistent with the observed data. The objective function, however, appears to be well-behaved in the parameter space.[66]

Even though the construction of expressways in Chicago is a single shock, the framework allows separately identifying type-specific racial preferences and disamenity parameters. Overall adjusted amenities can indeed be recovered from the observed data using the equilibrium conditions of the model separately for each type (equation 15). It then remains to separately identify, for each type, the racial preference parameters and the disamenity parameter. The racial preference parameters could be estimated from a regression of changes in total amenities on changes in residential externalities (racial preferences), instrumenting for changes in residential externalities with the road-based change in exposure to Black areas in the city interacted with distance grid cells from the Black Belt. The disamenity parameter could be estimated from a regression of changes in total amenities on changes in the quality of life of the neighborhoods, instrumenting for the change in the quality of life of neighborhoods with distance from the closest expressway interacted with indicator variables for distance grid cells from the CBD (to separate the relative importance of access benefits as opposed to the disamenity). The GMM estimator operates similarly to these IV regressions but allows one to jointly

[65] The choice responds to the trade-off between increasing precision and the risk of picking up noise in the data.

[66] Instead of plotting the results of a grid search over the parameter space – that would be difficult to interpret given the dimensionality of the objective function – I randomly choose the starting values of the iteration procedure and check whether it converges to the same parameter space. It does so in the vast majority of cases (100% for white, 90% for Black). See Appendix D.4.2 for details.

estimate all the parameters of interest in the same system.

Because of the mechanics of the model, any change in residential amenities that is not explained by changes in the composition of neighborhoods (through local spillovers and/or residential externalities) and disamenity effects will be explained by changes in adjusted residential fundamentals. For instance, let's assume that in reality, racial preference elasticities are high for each type (i.e., white people prefer to live close to other white people, Black people choose to live close to other Black people), but I pick wrong (lower) values of these parameters. Then, all else equal, the model will predict high residential fundamentals for Black (white) residents in places with a high concentration of Black (respectively, white) population. Since the demographic composition of neighborhoods changes over time, the model will infer that these changes are due to changes in residential fundamentals (of each type) rather than the underlying racial preferences. Comparing the estimates between the pre-period and the post-period will look as if residential fundamentals have moved as well, together with the relevant population, likely resulting in some of the above moment conditions being different from zero.

In Appendix D.4, I explore the sensitivity of the estimated parameters to the orthogonality conditions following Andrews et al. (2017). The parameters are generally sensitive to all moments in a similar way. The sensitivity measure of the decay parameter ρ is relatively similar in magnitude across moments. The parameter appears quite sensitive to the moments associated with being on the outskirts of the city (those related to longer distances from the Black Belt and the CBD), which could be due to the fact that the estimation does not take into account all the areas surrounding these locations (since they are outside of the city boundaries). The sensitivity measure of the parameter estimating the rate of spatial decay of the disamenity g is mostly sensitive to the set of moments related to distance to the expressways. The racial elasticity parameters $\lambda_{B,W}$ in general are similarly sensitive to all the moments.

7.4.3 GMM estimation results

Table 3 reports the efficient GMM estimation results. I find large and statistically significant racial preferences and disamenity parameters. To better understand how to interpret the estimated results and their magnitudes, I discuss them in order below.

First, the racial preference parameters show large heterogeneity by type, particularly when comparing Black and white individuals. The estimates are consistent with high degrees of homophily, with both Black and white residents exhibiting higher preferences for living close to same-race neighbors. Residential externalities appear to be an important agglomeration force, particularly in relation to the concentration of same-race residents in the vicinity. For white individuals, the estimated elasticity of amenities with respect to the concentration of nearby residents is notably higher for same-race neighbors compared to different-race neighbors: 3.3 times larger for white low-educated individuals and six times as large for white high-educated individuals. Conversely, estimates for Black residents align with substantial agglomeration forces linked to the density of same-race residents nearby while suggesting congestion forces concerning the density of white residents in the surrounding locations.

Second, residential externalities are highly localized and appear more localized for Black households than for white households. The rate of spatial decay of racial preferences is equal to 0.674 (s.e. 0.181) for low-educated Black individuals and 0.747 (s.e. 0.167) for high-educated Black individuals. For low-educated white residents, it equals 0.229 (s.e. 0.047) and 0.291 (s.e. 0.048) for high-educated white residents. The average value of the rate of spatial decay of racial preferences, weighted by the population shares, is $\rho_{whgt} = 0.342$.[67] To give a sense of the magnitudes, other things equal, residential externalities for Black residents fall to zero after around ten minutes of travel time, whereas for white residents after approximately 20 minutes of travel time. Figure 5 reports the proportional reductions in residential externalities with travel time (Table D3 in Appendix D show the same results but in comparison with the model-implied reduction in utility stemming from higher commuting times).

Third, the size of the disamenity, on average, appears larger for white residents than for Black residents, consistent with the reduced-form results. Black residents attach on average 22.0% inferior amenities to neighborhoods in proximity to expressways, whereas white residents 23.9% inferior amenities to the same neighborhoods (attenuating by 95% at 3.8 km from the expressway), comparable to the values recently found in the literature.[68]

7.4.4 Over-identification checks

A limited number of over-identification checks can be run in this setting due to the constraint of data availability. I examine how the model's predictions correlate with other variables not used in the estimation.

First, I start with the number of housing units. In the structural estimation, I recover a measure of the adjusted density of development (the ratio of residential floor space to land area) – a structural residual that ensures that the demand for floor space equals the supply of floorspace in each location.[69] To the extent that the density of development should be higher in dense residential locations, I investigate how it correlates with the reported number of housing units from the census. I find a positive relation (sample correlation of 0.698 in 1940 and 0.318 in 1990) between the (model-derived) density of development measure and the number of housing units reported from the census of the relevant year (both variables expressed in logs). Regression results are reported in the top panel of Table D4 in Appendix D.4. On average, a 1% increase in the number of housing units corresponds to a 0.765% increase in the density of development measure in 1940 and a 0.518% increase in 1990.

Second, I use information about zoning regulations today and check how the density of development measure correlates with the share of land for residential use in each location. The density of development (the model-derived ratio of residential floor space to land area) can be larger than 1

[67] To the best of my knowledge, this is the first estimate of this sort, making it difficult to benchmark its value. Nevertheless, it is reassuring to observe that it falls within the range of the most pertinent estimates found in the literature. On the one hand, the residential externalities parameter is equal to 0.76 in Ahlfeldt et al. (2015). On the other hand, the elasticity of consumption travel cost with travel times is estimated at 0.019 in Miyauchi et al. (2022).

[68] Brinkman and Lin (2022) finds freeway neighborhoods having 18.4% lower amenities.

[69] Figure D9 in Appendix D.4 reports the maps of the deciles of the distribution of this structural residual in 1940 and in 1990. It is reassuring to observe that the two distributions are similar over time.

(which is possible with multistory buildings), but in the data, I can only compute the share of land allowed for residential use (as opposed for instance to commercial and industrial use, or for parks and open spaces). I find a strong and positive log-linear relationship (sample correlation 0.462 in 1990) between (log) ϕ and the share of land for residential use (zoning data).[70] Regression results are reported in the bottom panel of Table D4. Overall, the strength of the results presented in the table provides support for the model's predictions. I find that a 1% increase in the share of land for residential use is associated, on average, with a 0.018% increase in the density of development.

7.5 Counterfactual exercises

I then use the model to run a set of counterfactual exercises. They consist of assuming alternative values of location characteristics or model parameters and solving for the model's counterfactual equilibrium. First, I begin by using counterfactuals to provide further evidence of the model's fit. I analyze the extent to which the effects of the construction of expressways on neighborhood demographics can be explained by the endogenous forces of the model rather than by changes in location fundamentals over time. Second, I run a counterfactual where I set racial preferences to zero. Third, I examine the relative importance of the disamenity as opposed to the barrier effects of expressways by shutting them down one at a time.

In the first counterfactual, I test the model's fit. I simulate the shock induced by the construction of expressways using the estimated parameters, but holding location fundamentals $\{\tilde{b}^o_j, \phi_j\}$ (i.e., residential fundamentals and density of development) fixed to the pre-period. As standard, I use the values of the endogenous variables from the observed equilibrium (in 1990) as the initial guess for the counterfactual equilibrium. I also set the reservation level of utility in the wider economy in the post-period so that the total population (for each type) living in the city is equal to its value in 1990. The correlation between the distribution of the share of Black in 1990 (data) and its counterfactual value is 0.815. The model thus can explain well the observed changes in neighborhoods' demographics (the binned scatter plot of the two variables is reported in Figure D10 in Appendix D).[71]

The counterfactual treatment effects on racial sorting for the remaining counterfactual exercises are reported in Figure 6. It plots the distribution of population in Chicago by neighborhood racial composition. The two counterfactual distributions are reported against the observed equilibrium in 1990 (yellow dots). In 1990, Chicago was largely segregated: 50% of the total population in the city lived in neighborhoods that were at most 10% Black, and 30% of the population lived in neighborhoods that were 90% or more Black.

In a counterfactual exercise, I study the implications of removing the racial bias. I set the elasticity of amenities with respect to the concentration of different-race residents in the surrounding areas equal to the elasticity of amenities with respect to the concentration of one's own race.[72] That is, I

[70]This second over-identification check is computed exclusively for 1990 because I have access to zoning information covering only the present period.

[71]By looking at plot, the model predicts well the distribution of races in Chicago following the construction of expressways. It slightly underestimates the share of Black living in highly Black locations, implying that, for these locations, the change in residential fundamentals must be high (see Figure D10 in Appendix D).

[72]For Black types (both high and low educated), I set $\lambda^*_W = \lambda_B$; for white high and low educated types, I set $\lambda^*_B = \lambda_W$

assume that the elasticity of amenities from having an extra resident of a different race is equivalent to the estimated elasticity from having an additional neighbor of the same race. I find that the counterfactual population of Chicago lives in more integrated neighborhoods. The share of individuals living in neighborhoods with 90% or more Black shares drops to around 5%, and the share of residents living in a neighborhood with at most a 10% Black share drops from 50% to less than 300%. The population appears more evenly distributed across all neighborhood configurations.

I then conduct an additional counterfactual exercise to analyze the implications of mitigating the neighborhood effects of expressways. I investigate the counterfactual treatment effects of simultaneously removing (i) the disamenity ($g^o = 0$) and (ii) the barrier effect of expressways. To set the barrier effect of expressways to zero, I follow Brinkman and Lin (2022). I define pairwise travel time as $\tau_{ji} = \tau^*_{ji} + c_{ji}$ where τ^*_{ji} is the travel time in the absence of the expressway and c_{ji} is the extra cost of crossing after the urban barrier is built. From their estimates, I calibrate $c_{ji} = 2$ minutes for trips less than 5 km that cross the expressway. Before running the counterfactual, I hence recalculate the matrix of pairwise travel times, setting travel times to 2 minutes faster for trips shorter than 5 km crossing an expressway. I find that the distribution of races in the counterfactual equilibrium is characterized by a drop of 10 percentage points in the share of residents of Chicago living in a neighborhood with 90% or more Black individuals. At the same time, almost 20% of Chicago's population would end up residing in perfectly integrated neighborhoods (those characterized by around 30% Black share) – nearly a seven-fold increase relative to the observed equilibrium.

In addition, removing the disamenity alone leads, on average, to a doubling in the population living within 1 km of the expressway (relative to the observed equilibrium) – offsetting the effects evaluated in the reduced-form analysis – and an average increase of 50% in the population living within 2 km of the expressway, with larger effects closer to the city center.

Finally, I calculate how segregated Chicago would be if we removed the adverse neighborhood effects of expressways (canceling both the disamenity and the barrier effect). The index of dissimilarity from the counterfactual experiment is equal to 0.702.[73] In comparison, it is calculated at 0.844 in the observed equilibrium, meaning that in 1990, 84% of Black residents would have had to change location with white residents to achieve full spatial integration in the city. Mitigating the neighborhood effects of expressways is associated with a reduction of 16.8% in racial segregation in Chicago.

8 Conclusion

This paper deepens our understanding of the role of urban structures and urban forms in shaping the allocation of people within cities. Although anecdotal evidence on the link between physical barriers and socio-economic disparities is abundant, to the best of my knowledge, this is the first work to systematically examine this link, providing a setting and an empirical strategy to plausibly make causal estimates of the relationship.

(where the asterisk denotes the counterfactual value).

[73] First proposed by Duncan and Duncan (1955), the index of dissimilarity is a widely used measure of spatial segregation of different racial groups (for an application, see, for instance, Cutler et al., 1999). It is measured as follows: $D = (1/2) \sum_i |\frac{Black_i}{Black_{total}} - \frac{non-Black_i}{non-Black_{total}}|$.

I provide evidence of the importance of the dual nature of expressways in evaluating their impact on racial sorting. First, similar to a policy that creates within-city shocks to residential amenities and hence makes neighborhoods more or less desirable places to live, the impact of expressways operates through a price or disamenity channel. Since Black households are, on average, poorer than white households, they are more likely to live closer to the expressways because of cheap housing. Second, expressways are physical barriers that affect the degree of exposure to the neighborhoods in the city and, hence, to different demographic compositions. I show that higher exposure to Black areas in the city (i) increases the likelihood that a neighborhood becomes more black over time and (ii) reduces its valuation in the long run. This finding suggests that this feature of the expressways provides a second channel of racial sorting, which depends on individual preferences towards more or less integrated places to live in.

Motivated by these findings, I develop a quantitative spatial urban model with racial preferences for residential locations. The setup follows the canonical Alonso (1964), Muth (1969), and Mills (1967) monocentric city model, but it features an internal city structure à la Ahlfeldt et al. (2015). Using the same type of variation of my reduced-form analyses, the model allows me to estimate racial preferences parameters, and then to undertake counterfactual exercises. Mitigating the neighborhood effects of expressways in Chicago would lead to a 16.8% reduction in racial segregation.

The findings shed light on the possible unintended long-lasting neighborhood effects of transportation infrastructure that are pervasive in the landscape of many cities. Future research could tackle the extent to which long-run effects can be ascribed to institutional changes (like modifications to school, police, and administrative boundaries) in response to the construction of urban barriers. In addition, the recent surge in widespread GPS data availability could allow us to explore how urban forms affect the accessibility to different sets of consumption goods and experiences within the city. Given the recent surge in the issue of social justice in the transportation sector, advocated by the Biden administration as one of the main challenges of our time, this research lies in an important space for current public policy.

References

Aaronson, Daniel, Daniel Hartley, and Bhashkar Mazumder (2021). *The Effects of the 1930s HOLC "Redlining" Maps.* American Economic Journal: Economic Policy, 13 (4): 355-392.

Ahlfeldt, Gabriel M., and Arne Feddersen (2018). *From periphery to core: measuring agglomeration effects using high-speed rail.* Journal of Economic Geography, 18 (2): 355-390.

Ahlfeldt, Gabriel M., and Daniel P. McMillen (2014). *Land Values in Chicago, 1913-2010: A City's Spatial History Revealed.* Land Lines, Lincoln Institute of Land Policy, Cambridge MA, 15-21.

Ahlfeldt, Gabriel M., and Daniel P. McMillen (2018). *Tall buildings and land values: Height and construction cost elasticities in Chicago, 1870-2010.* Review of Economics and Statistics, 100 (5): 861-875.

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf (2015). *The economics of density: Evidence from the Berlin Wall.* Econometrica, 83 (6): 2127-2189.

Ahlfeldt, Gabriel M., Volker Nitsch, and Nicolai Wendland (2019). *Ease vs. noise: Long-run changes in the value of transport (dis) amenities.* Journal of Environmental Economics and Management, 98: 102268.

Alonso, William (1964). *Location and land use: toward a general theory of land rent.* Harvard university press.

Ananat, Elizabeth (2011). *The Wrong Side(s) of the Tracks: The Causal Effects of Racial Segregation on Urban Poverty and Inequality.* American Economic Journal: Applied Economics, 3 (2): 34-66.

Anderson, Michael L. (2020). *As the wind blows: The effects of long-term exposure to air pollution on mortality.* Journal of the European Economic Association, 18 (4): 1886-1927.

Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro (2017). *Measuring the sensitivity of parameter estimates to estimation moments.* The Quarterly Journal of Economics, 132 (4): 1553-1592.

Archer, Deborah N. (2020) *"White Men's Roads through Black Men's Homes": Advancing Racial Equity through Highway Reconstruction.* Vanderbilt Law Review, 73 (5): 1259.

Atack, Jeremy (2015). *Historical Geographic Information Systems (GIS) database of US Railroads for 1826-1911.*

Athey, Susan, Billy A. Ferguson, Matthew Gentzkow, and Tobias Schmidt (2020). *Experienced segregation.* NBER Working Paper 27572.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian (2020). *On the Road: Access to Transportation Infrastructure and Economic Growth in China.* Journal of Development Economics, 145 (102442).

Banzhaf, H. Spencer, and Randall P. Walsh (2013). *Segregation and Tiebout sorting: The link between place-based investments and neighborhood tipping.* Journal of Urban Economics, 74: 83-98.

Baum-Snow, Nathaniel (2007). *Did highways cause suburbanization?* Quarterly Journal of Economics, 122 (2): 775-805.

Baum-Snow, Nathaniel (2020). *Urban transport expansions and changes in the spatial structure of us cities: Implications for productivity and welfare.* Review of Economics and Statistics, 102 (5): 929-945.

Baum-Snow, Nathaniel, Loren Brandt, J. Vernon Henderson, Matthew A. Turner, and Qinghua Zhang (2017). *Roads, railroads, and decentralization of Chinese cities.* Review of Economics and Statistics, 99 (3): 435-448.

Baum-Snow, Nathaniel, J. Vernon Henderson, Matthew A. Turner, Qinghua Zhang, and Loren Brandt (2020). *Does investment in national highways help or hurt hinterland city growth?.* Journal of Urban Economics, 115: 103124.

Bayer, Patrick, Fernando Ferreira, and Robert McMillan (2007). *A unified framework for measuring preferences for schools and neighborhoods.* Journal of Political Economy, 115 (4): 588-638.

Bayer, Patrick, Hanming Fang, and Robert McMillan (2014). *Separate when equal? Racial inequality and residential segregation.* Journal of Urban Economics, 82: 32-48.

Bayer, Patrick, Robert McMillan, Alvin Murphy, and Christopher Timmins (2016). *A dynamic model of demand for houses and neighborhoods.* Econometrica, 84 (3): 893-942.

Borusyak, Kirill, and Xavier Jaravel (2017). *Revisiting event study designs.* SSRN 2826228.

Boustan, Leah Platt (2010). *Was postwar suburbanization "white flight"? Evidence from the black migration.* Quarterly Journal of Economics, 125 (1): 417-443.

Brinkman, Jeffrey, and Jeffrey Lin (2022). *Freeway revolts! The Quality of Life Effects of Highways.* Review of Economics and Statistics, 1–45.

Caetano, Gregorio, and Vikram Maheshri (2019). *A Unified Empirical Framework to Study Segregation.* R&R Econometrica.

Callaway, Brantly, and Pedro H. C. Sant'Anna (2021). *Difference-in-Differences with Multiple Time Periods.* Journal of Econometrics, 225 (2) :200-230.

Card, David, Alexandre Mas, and Jesse Rothstein (2008). *Tipping and the Dynamics of Segregation.* Quarterly Journal of Economics, 123 (1): 177-218.

Carter, Chelsea E. (2023). *The Road to the Urban Interstates: A Case Study from Detroit*. R&R Journal of Economic History.

Chandra, Amitabh, and Eric Thompson (2000). *Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system.* Regional science and urban economics, 30 (4): 457-490.

Chetty, Raj, and Nathaniel Hendren (2018a). *The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects.* Quarterly Journal of Economics, 133 (3):1107-1162.

Chetty, Raj, and Nathaniel Hendren (2018b). *The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates.* Quarterly Journal of Economics, 133 (3):1163-1228.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz (2016). *The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment.* American Economic Review, 106 (4): 855-902.

Christensen, Peter, and Christopher Timmins (2023). *The damages and distortions from discrimination in the rental housing market.* The Quarterly Journal of Economics, 138 (4): 2505-2557.

Collins, William, and Margo, Robert (2000). *Residential Segregation and Socioeconomic Outcomes: When Did Ghettos Go Bad?* Economics Letters, 69 (2): 239-43.

Combes, Pierre-Philippe, Gilles Duranton, Laurent Gobillon, and Sébastien Roux (2010). *Estimating agglomeration economies with history, geology, and worker effects.* In Agglomeration economics (pp. 15-66). University of Chicago Press

Conley, Timothy G (1999). *GMM estimation with cross sectional dependence.* Journal of Econometrics, 92 (1): 1-45.

Currarini, Sergio, Matthew O. Jackson, and Paolo Pin (2009). *An economic model of friendship: Homophily, minorities, and segregation.* Econometrica, 77 (4): 1003-1045.

Cutler, David, and Glaeser, Edward (1997). *Are Ghettos Good or Bad?* Quarterly Journal of Economics, 112 (3): 827-72.

Cutler, David, Glaeser, Edward, and Vigdor, Jacob (1999). *The Rise and Decline of the American Ghetto.* Journal of Political Economy, 107 (3): 455-506.

Cutler, David, Glaeser, Edward, and Vigdor, Jacob (2008). *When Are Ghettos Bad? Lessons from Immigrant Segregation in the United States.* Journal of Urban Economics, 63 (3): 759-74.

Davis, Donald R., and Jonathan I. Dingel (2020). *The comparative advantage of cities.* Journal of International Economics, 123: 103291.

Davis, Morris A., Jesse Gregory, and Daniel A. Hartley (2023). *Preferences over the Racial Composition of Neighborhoods: Estimates and Implications.* SSRN 4495735.

de Chaisemartin, Clément, and Xavier D'Haultfoeuille (2020). *Two-way fixed effects estimators with heterogeneous treatment effects.* American Economic Review, 110 (9): 2964-96.

de Chaisemartin, Clément, and Xavier d'Haultfoeuille (2023). *Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey.* The Econometrics Journal, (26) 3: C1-C30.

Derenoncourt, Ellora (2022). *Can you move to opportunity? Evidence from the Great Migration.* American Economic Review, 112 (2): 369-408.

Donaldson, Dave (2018). *Railroads of the Raj: Estimating the Impact of Transportation Infrastructure.* American Economic Review, 108 (4-5): 899-934.

Donaldson, Dave, and Richard Hornbeck (2016). *Railroads and American economic growth: A "market access" approach.* Quarterly Journal of Economics, 131 (2): 799-858.

Dorn, David (2008). *Price and prejudice: The interaction between preferences and incentives in the dynamics of racial segregation.* Technical report, Boston University working paper.

Duncan, Otis Dudley, and Beverly Duncan (1955). *A methodological analysis of segregation indexes.* American sociological review 20 (2): 210-217.

Duranton, Gilles and Matthew A. Turner (2012). *Urban Growth and Transportation.* Review of Economic Studies, 79 (4): 1407-1440.

Duranton, Gilles, Peter Morrow, and Matthew A. Turner (2014). *Roads and Trade: Evidence from the US.* Review of Economic Studies, 81 (2): 681-724.

Echenique, Federico, and Fryer, Roland (2007). *A Measure of Segregation Based on Social Interactions.* Quarterly Journal of Economics, 122 (2): 441-85.

Epple, Dennis, Brett Gordon, and Holger Sieg (2010). *A new approach to estimating the production function for housing.* American Economic Review, 100 (3): 905-924.

Faber, Benjamin (2014). *Trade integration, market size, and industrialization: Evidence from China's National Trunk Highway System.* Review of Economic Studies, 81 (3): 1046-1070.

Fajgelbaum, Pablo D., and Cecile Gaubert (2020). *Optimal spatial policies, geography, and sorting.* The Quarterly Journal of Economics, 135 (2): 959-1036.

Fishback, Price V., Jessica LaVoice, Allison Shertzer, and Randall Walsh (2020). *Race, risk, and the emergence of federal redlining.* NBER Working Paper 28146.

Fogli, Alessandra, and Veronica Guerrieri (2019). *The End of the American Dream? Inequality and Segregation in Us Cities.* NBER Working Paper 26143.

Fretz, Stephan, Raphaël Parchet, and Frédéric Robert-Nicoud (2022). *Highways, market access and spatial sorting.* The Economic Journal 132 (643): 1011-1036.

Gechter, Michael, and Nick Tsivanidis (2023). *Spatial Spillovers from High-Rise Developments: Evidence from the Mumbai Mills.* Mimeo.

Gibbons, Stephen, Teemu Lyytikäinen, Henry G. Overman, and Rosa Sanchis-Guarner (2019). *New road infrastructure: the effects on firms.* Journal of Urban Economics, 110: 35-50.

Goettler, Ronald L., and Ron Shachar (2001). *Spatial competition in the network television industry.* RAND Journal of Economics: 624-656.

Gonzalez-Navarro, Marco, and Matthew A. Turner (2018). *Subways and urban growth: Evidence from earth.* Journal of Urban Economics, 108: 85-106.

Goodman-Bacon, Andrew. *Difference-in-differences with variation in treatment timing.* Journal of Econometrics.

Gregory, Victoria, Julian Kozlowski, and Hannah Rubinton (2022). *The Impact of Racial Segregation on College Attainment in Spatial Equilibrium.* FRB St. Louis Working Paper 2022-36.

Hansen, Walter G. *How accessibility shapes land use.* Journal of the American Institute of planners, 25 (2): 73-76.

Heblich, Stephan, Stephen J. Redding, and Daniel M. Sturm (2020). *The making of the modern metropolis: evidence from London.* Quarterly Journal of Economics, 135 (4): 2059-2133.

Heblich, Stephan, Alex Trew, and Yanos Zylberberg (2021). *East side story: Historical pollution and persistent neighborhood sorting.* Journal of Political Economy, 129 (5): 1508-1552.

Hu, Patricia S., and Timothy R. Reuscher (2004). *Summary of Travel Trends: 2001 National Household Travel Survey.* US Department of Transportation, Federal Highway Administration.

Jedwab, Remi, and Alexander Moradi (2016). *The permanent effects of transportation revolutions in poor countries: evidence from Africa.* Review of economics and statistics, 98 (2): 268-284.

Kain, John F (1968). *Housing segregation, negro employment, and metropolitan decentralization.* Quarterly Journal of Economics, 82 (2): 175-197.

Karner, Alex A., Douglas S. Eisinger, and Deb A. Niemeier (2010). *Near-roadway air quality: synthesizing the findings from real-world data.* Environmental Science and Technology, 44 (14): 5334-5344.

Krimmel, Jacob (2020). *Persistence of Prejudice: Estimating the Long Term Effects of Redlining*. Mimeo.

Lee, Sanghoon, and Jeffrey Lin (2018). *Natural amenities, neighbourhood dynamics, and persistence in the spatial distribution of income.* Review of Economic Studies, 85 (1): 663-694.

Logan, Trevon D., and John M. Parman (2017). *The national rise in residential segregation.* Journal of Economic History, 77 (1): 127-170.

Logan, John R., Zengwang Xu, and Brian J. Stults (2014). *Interpolating US decennial census tract data from as early as 1970 to 2010: A longitudinal tract database.* Professional Geographer, 66 (3): 412-420.

Mahajan, Avichal (2023). *Highways and segregation.* Journal of Urban Economics (in press).

McFadden, Daniel (1974). *The measurement of urban travel demand.* Journal of Public Economics, 3 (4): 303-328.

Michaels, Guy (2008). *The effect of trade on the demand for skill: Evidence from the interstate highway system.* The Review of Economics and Statistics, 90 (4): 683-701.

Mills, Edwin S. (1967). *An aggregative model of resource allocation in a metropolitan area.* American Economic Review, 57 (2): 197–210.

Miyauchi, Yuhei, Kentaro Nakajima, and Stephen J. Redding (2022). *The economics of spatial mobility: Theory and evidence using smartphone Data.* NBER Working Paper 28497.

Mohl, Raymond A. (2001). *Urban Expressways and the Racial Restructuring of Postwar American Cities* Economic History Yearbook, 42 (2): 89-104.

Mohl, Raymond A. (2004). *Stop the road: Freeway revolts in American cities.* Journal of Urban History, 30 (5): 674-706.

Mohl, Raymond A. (2008). *The interstates and the cities: the US Department of Transportation and the freeway revolt, 1966–1973.* Journal of Policy History, 20 (2): 193-226.

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg (2018). *Commuting, migration, and local employment elasticities.* American Economic Review, 108 (12): 3855-90.

Morten, Melanie, and Jaqueline Oliveira (2023). *The Effects of Roads on Trade and Migration: Evidence from a Planned Capital City.* American Economic Journal: Applied Economics (forthcoming).

Muth, Richard F. (1969). *Cities and Housing: The Spatial Pattern of Urban Residential Land Use.* The University of Chicago Press.

Newcomb, Charles S., and Richard O. Lang (1934) *Chicago Census Commission. Census Data of the City of Chicago, 1934.* The University of Chicago Press.

Redding, Stephen J., and Daniel Sturm (2023). *Estimating Neighborhood Effects: Evidence from Wartime Destruction in London.* Mimeo.

Redding, Stephen J., and Matthew A. Turner (2015). *Transportation costs and the spatial organization of economic activity.* Handbook of regional and urban economics, 5: 1339-1398.

Robinson, John (1971). *Highways and Our Environment.* New York: McGraw-Hill.

Regional Transportation Authority, Finance and Performance Management (2013). *2012 Sub-Regional Report - Performance Measures.*

Roberto, Elizabeth, and Jackelyn Hwang (2017). *Barriers to integration: Institutionalized boundaries and the spatial structure of residential segregation.* Annual Meeting of the Population Association of America, Chicago.

Rambachan, Ashesh, and Jonathan Roth (2020). *An Honest Approach to Parallel Trends.* Unpublished manuscript.

Rothstein, Richard (2017). *The color of law: A forgotten history of how our government segregated America.* Liveright Publishing.

Saiz, Albert (2010). *The geographic determinants of housing supply.* The Quarterly Journal of Economics 125 (3): 1253-1296.

Sant'Anna, Pedro H. C., and Jun Zhao (2020). *Doubly Robust Difference-in-Differences Estimators.* Journal of Econometrics, 219 (1): 101–22.

Schelling, Thomas C. (1963). *The Strategy of Conflict.* Cambridge, MA: Harvard University Press.

Schelling, Thomas C. (1971). *Dynamic models of segregation.* Journal of mathematical sociology, 1 (2): 143-186.

Sethi, Rajiv, and Rohini Somanathan (2004). *Inequality and segregation.* Journal of Political Economy, 112 (6): 1296-1321.

Shertzer, Allison, Randall P. Walsh, and John R. Logan (2016). *Segregation and Neighborhood Change in Northern Cities: New Historical GIS data from 1900-1930.* Historical Methods 49 (4): 187-197.

Shertzer, Allison, and Randall P. Walsh (2019). *Racial sorting and the emergence of segregation in American cities.* Review of Economics and Statistics, 101 (3): 415-427.

Shi, Ying, Daniel Hartley, Bhash Mazumder, and Aastha Rajan (2022). *The effects of the Great Migration on urban renewal.* Journal of Public Economics, 209: 104647.

State of Illinois, County of Cook, City of Chicago, and US Department of Commerce (1959). *Chicago Area Transportation Study: Final Report — Volume II: Data Projections.*

Storeygard, Adam (2016). *Farther on down the road: transport costs, trade and urban growth in sub-Saharan Africa.* The Review of Economic Studies, 83 (3): 1263-1295.

Sun, Liyang, and Sarah Abraham (2020). *Estimating dynamic treatment effects in event studies with heterogeneous treatment effects.* Journal of Econometrics.

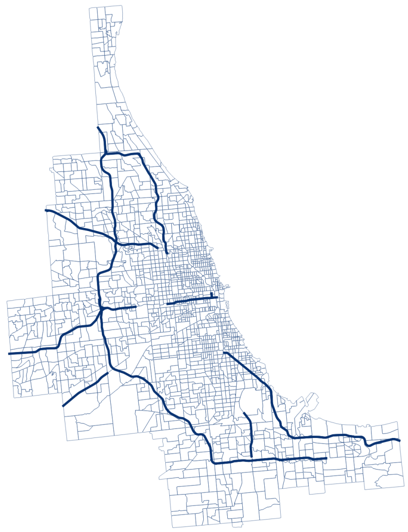
Tsivanidis, Nick (2022). *The Aggregate and Distributional Effects of Urban Transit Infrastructure: Evidence from Bogota's TransMilenio.* R&R American Economic Review.

Weiwu, Laura (2023). *Unequal Access: Racial Segregation and the Distributional Impacts of Interstate Highways in Cities.* Mimeo.

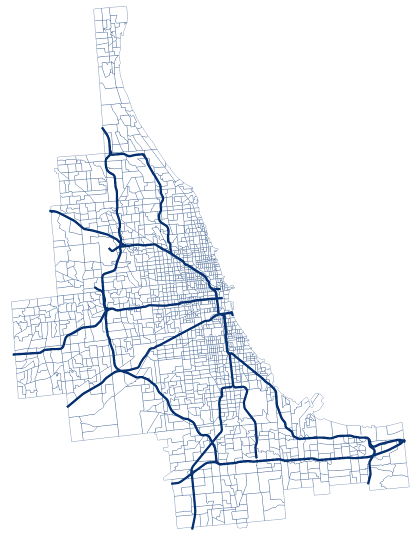
Wooldridge, Jeffrey M. (2015). *Control function methods in applied econometrics.* Journal of Human Resources, 50 (2): 420-445.

Figures

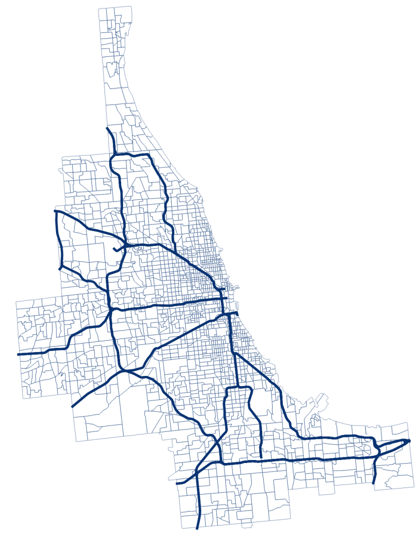
Figure 1: Timeline of expressway construction, 1950-2010



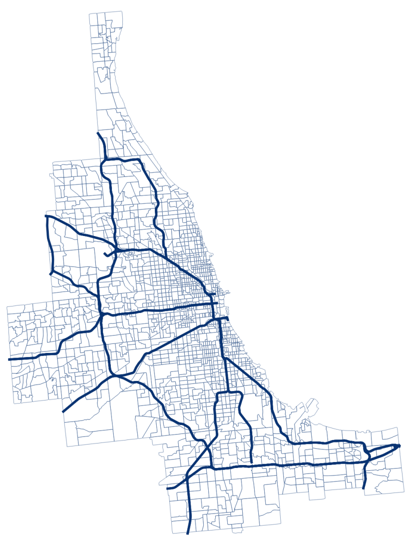
(a) 1960



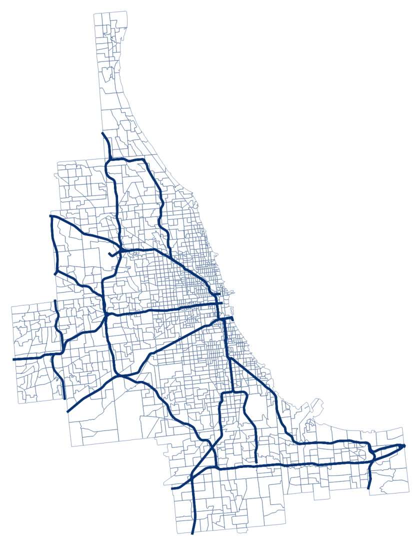
(b) 1970



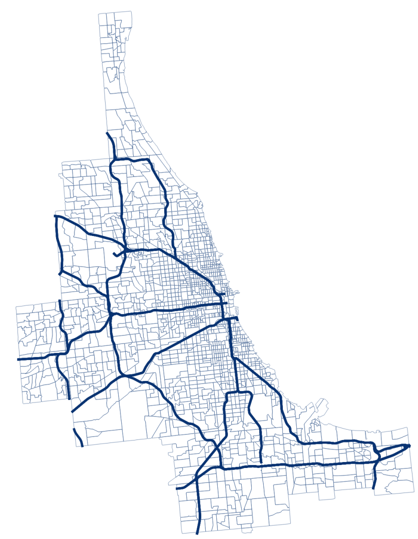
(c) 1980



(d) 1990



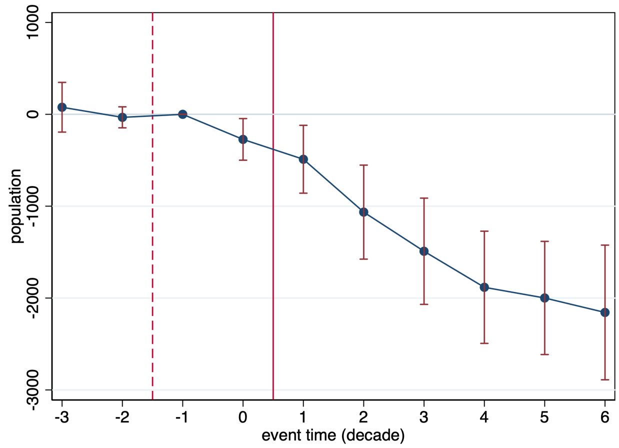
(e) 2000



(f) 2010

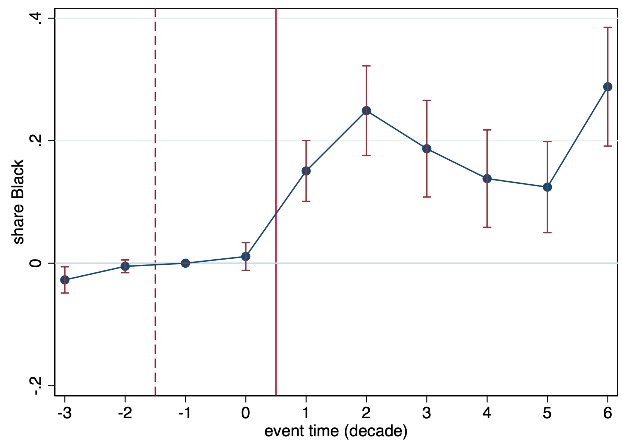
The maps show the rollout of the expressway network within the boundaries of the Chicago Metropolitan Area. The geographic extent of the city is determined by data availability in 1950. Polygons are the 1,511 consistent-boundary census tracts that constitute the units of analysis.

Figure 2: Effect of proximity to expressways on residential population



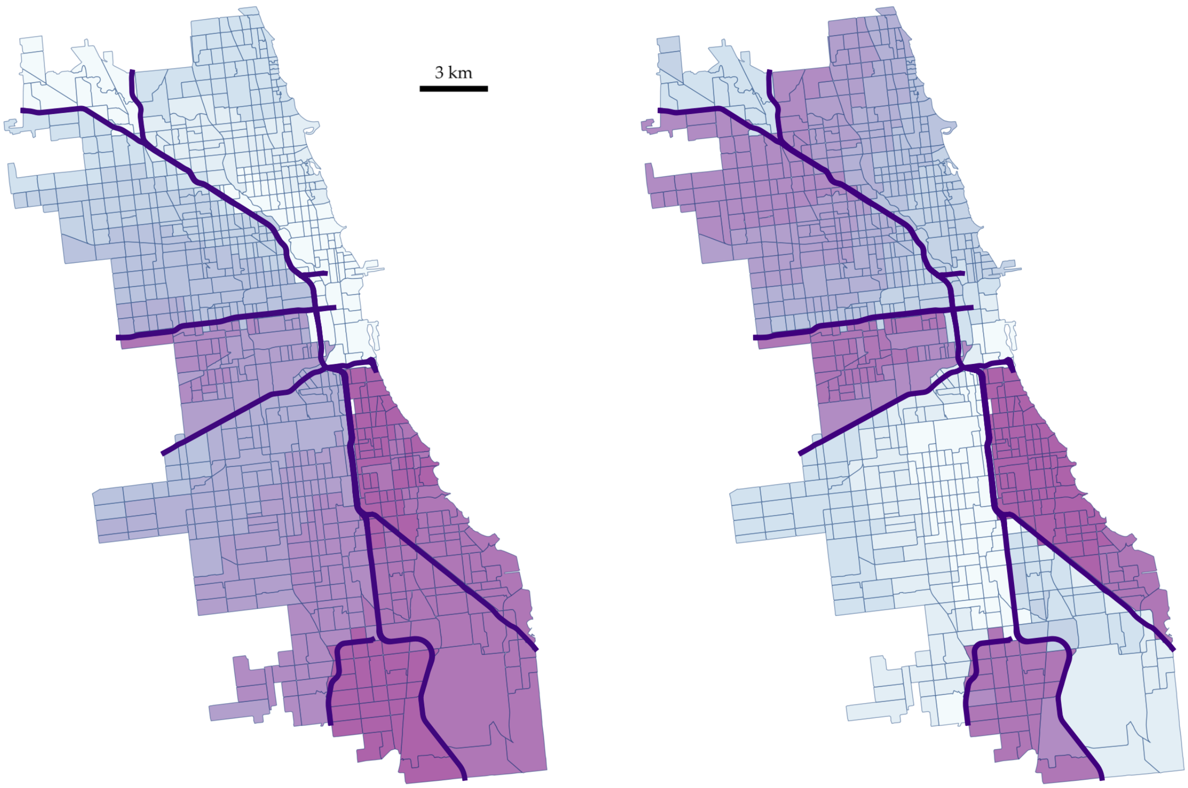
Note: The figure plots the β coefficients estimated from regression (1). using the two-way fixed effects estimator proposed by de Chaisemartin and D'Haultfoeuille (2020). The dependent variable is residential population. The omitted category is event time dummy at $t = -1$. Event time 1 corresponds to the decade in which the expressway is found to be in operation. A census tract belongs to the treatment group if its centroid lies within 1 km from the closest expressway; it is part of the control group if its centroid is farther than 3 km from the closest expressway. Event times to the right of the solid red vertical bar correspond to post-treatment periods. Event times to the left of the dotted red vertical bar correspond to pre-treatment periods. Event times in between correspond to the periods between expressways were first planned and the year in which they were in operation. The full set of controls includes: Dist. to CBD $\times$ Year FE, Quadratic Dist. to CBD $\times$ Year FE, HOLC Grade $\times$ Year FE, 1940 Pop.D. $\times$ Year FE, City $\times$ Year FE. The 95% confidence intervals are based on standard errors clustered at census tract level. Census Tract FE and Year FE are always included.

Figure 3: Effect of proximity to expressways on share of Black residents



Note: The figure plots the β coefficients estimated from regression (1) using the two-way fixed effects estimator proposed by de Chaisemartin and D'Haultfoeuille (2020). The dependent variable is the census tract average share of Black residents. The omitted category is event time dummy at $t = -1$. Event time 1 corresponds to the decade in which the expressway is found to be in operation. A census tract belongs to the treatment group if its centroid lies within 1 km from the closest expressway; it is part of the control group if its centroid is farther than 3 km from the closest expressway. Event times to the right of the solid red vertical bar correspond to post-treatment periods. Event times to the left of the dotted red vertical bar correspond to pre-treatment periods. Event times in between correspond to the periods between expressways were first planned and the year in which they were in operation. The full set of controls includes: Dist. to CBD $\times$ Year FE, Quadratic Dist. to CBD $\times$ Year FE, HOLC Grade $\times$ Year FE, 1940 Pop.D. $\times$ Year FE, City $\times$ Year FE. The 95% confidence intervals are based on standard errors clustered at census tract level. Census Tract FE and Year FE are always included.

Figure 4: Change in exposure to Black residents and its IV, 1950-1990

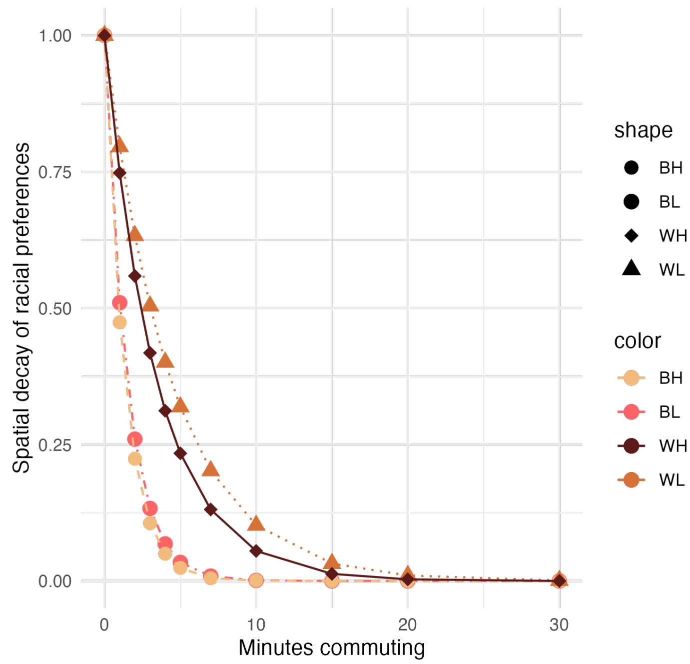


(a) Change in exposure to Black residents (ΔS)

(b) Instrument (ΔSMA)

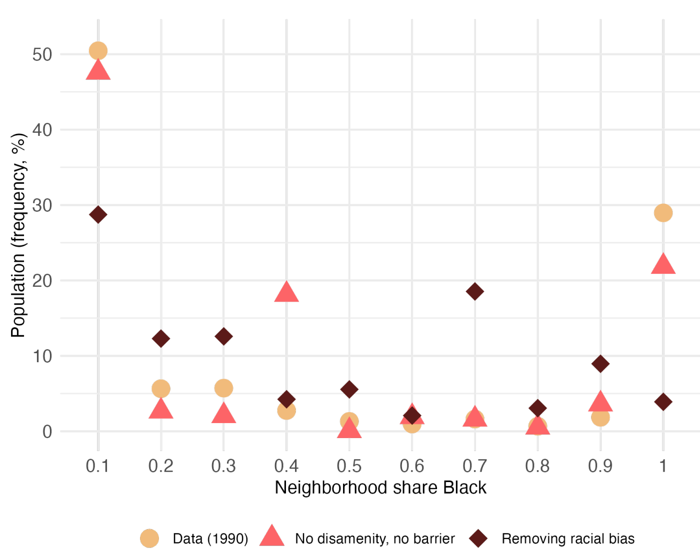
Note: Panel (a) plots the overall change in exposure to Black residents ΔS_i (panel a) and for each neighborhood in the city. The overall change includes both the change in exposure due to the spatial resorting of people between 1950 and 1990 and the change due to the development of the road network. Panel (b) plots the baseline instrument for the change in exposure to Black residents. The instrument isolates the variation that is due only to changes in traveling times through expressway construction. To isolate the barrier effect, I set the costs of crossing the expressway network to infinity. Census tracts are grouped into deciles based on the change in exposure to Black residents, with darker colors indicating larger increases in exposure. The purple lines show the expressways route as of 1990.

Figure 5: Spatial decay of racial preference ρ^o



Note: The figure plots the estimated reductions in residential externalities with travel time (in minutes), $1 - e^{-\rho^o \tau}$ separately for each type. Additional details are reported in Table D3 in Appendix D.

Figure 6: Counterfactual racial distributions



Note: The figure plots the distribution of population in Chicago, by neighborhood racial composition. The counterfactual distributions are reported against the observed equilibrium in 1990 (yellow dots). In 1990, Chicago was largely segregated: 50% of the total population in the city lived in a neighborhood that was at most 10% Black, and 30% of the population lived in a neighborhood that was 90% or more Black. In the two counterfactual scenarios reported here, segregation goes down. Fewer people live in largely segregated neighborhoods, and more people live in mixed neighborhoods.

Tables

Table 1: Dep variable: Δ share Black (standardized)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS All	(6) Grid city se	(7) Conley se	(8) IV All	(9) IV Subs 1	(10) IV Subs 2	(11) IV Δ YMA
ΔS (std)	0.496***	0.500***	0.335***	0.395***	0.411***	0.411***	0.411***	0.158**	0.170**	0.203***	0.156**
	(0.053)	(0.052)	(0.054)	(0.050)	(0.052)	(0.123)	(0.085)	(0.069)	(0.068)	(0.070)	(0.069)
Dist expressway (km)		0.011	-0.179***	-0.204***	-0.202***	-0.202***	-0.202***	-0.231***	-0.233***	-0.241***	-0.231***
		(0.018)	(0.020)	(0.021)	(0.021)	(0.059)	(0.031)	(0.022)	(0.022)	(0.023)	(0.022)
ΔY (std)					0.091*	0.091	0.091	0.033	0.039	-0.079	0.038
					(0.050)	(0.122)	(0.091)	(0.051)	(0.052)	(0.060)	(0.058)
Observations	764	764	764	727	727	727	727	727	722	648	727
Adjusted R^2	0.224	0.223	0.397	0.470	0.471	0.471	0.470				
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tract controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historical Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-stat								1265	1249	1055	573.4

The table reports the estimation results from regression (2) on the change in the share of Black residents. The regressors of interest are: ΔS (standardized) measures the change in exposure to Black areas in the city; *Dist expressway* measures the km distance from the closest expressway; ΔY measures the change in exposure to rich areas in the city. Region FE for being in the north, west, or south side of the city are always included. Tract controls are: a quadratic polynomial of distance to the central business district (CBD), land area, and distance to water. Historical controls include: distance to railroads in 1898; HOLC grade, historical outcomes in levels, and the change in population density between 1920 and 1940. Column (6) shows standard errors assuming that census tract within the same grid cell are spatially correlated (for this exercise, I partitioned the city into 25 squares). Column (7) computes Conley (1999) standard errors to allow arbitrary spatial correlation of errors between tracts within 3 km from each other. Column (8) shows the IV results after instrumenting for the change in exposure to Black residents holding the distribution of races within the city fixed to the pre-period. Column (9) reports the IV results after removing the 5% of the sample that in 1920 already had a large concentration of Black residents (i.e., removes all area with more than 20% share in 1920). Column (10) reports the estimated IV coefficients after removing the 12% of the census tracts in the sample that are within 5 km of the CBD. Column (11) additionally instruments for the change in exposure to rich neighborhoods in the city, with an instrumental variable (ΔYMA) that is constructed similarly to ΔSMA. Supplementary detail is available in the main text. Standard errors clustered at census tract level in parentheses (unless specified differently). *** p<0.01, ** p<0.05, * p<0.1.

Table 2: Dep variable: Δ land value, log (standardized)

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS All	(6) Grid city se	(7) Conley se	(8) IV All	(9) IV Subs 1	(10) IV Subs 2	(11) IV Δ YMA
ΔS (std)	-0.354***	-0.399***	-0.353***	-0.342***	-0.333***	-0.333**	-0.333**	-0.235***	-0.299***	-0.323***	-0.250***
	(0.045)	(0.045)	(0.050)	(0.051)	(0.051)	(0.137)	(0.074)	(0.064)	(0.061)	(0.064)	(0.065)
Dist expressway (km)		-0.130***	0.028	0.042	0.044	0.044	0.044	0.057**	0.052*	0.085***	0.056**
		(0.024)	(0.027)	(0.027)	(0.027)	(0.084)	(0.064)	(0.027)	(0.027)	(0.026)	(0.027)
ΔY (std)					0.057	0.057	0.057	0.077	0.046	0.292***	0.125*
					(0.061)	(0.119)	(0.108)	(0.061)	(0.059)	(0.060)	(0.065)
Observations	742	742	742	720	720	720	720	720	715	641	720
Adjusted R^2	0.284	0.320	0.436	0.436	0.436	0.436	0.435				
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Tract controls	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Historical Controls	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-stat								1341	1283	1104	559.3

The table reports the estimation results from regression (2) on the change in land value. The regressors of interest are: ΔS (standardized) measures the change in exposure to Black areas in the city; *Dist expressway* measures the km distance from the closest expressway; ΔY measures the change in exposure to rich areas in the city. Region FE for being in the north, west, or south side of the city are always included. Tract controls are: a quadratic polynomial of distance to the central business district (CBD), land area, and distance to water. Historical controls include: distance to railroads in 1898; HOLC grade, historical outcomes in levels, and the change in population density between 1920 and 1940. Column (6) shows standard errors assuming that census tract within the same grid cell are spatially correlated (for this exercise, I partitioned the city into 25 squares). Column (7) computes Conley (1999) standard errors to allow arbitrary spatial correlation of errors between tracts within 3 km from each other. Column (8) shows the IV results after instrumenting for the change in exposure to Black residents holding the distribution of races within the city fixed to the pre-period. Column (9) reports the IV results after removing the 5% of the sample that in 1920 already had a large concentration of Black residents (i.e., removes all area with more than 20% share in 1920). Column (10) reports the estimated IV coefficients after removing the 12% of the census tracts in the sample that are within 5 km of the CBD. Column (11) additionally instruments for the change in exposure to rich neighborhoods in the city, with an instrumental variable (ΔYMA) that is constructed similarly to ΔSMA. Supplementary detail is available in the main text. Standard errors clustered at census tract level in parentheses (unless specified differently). *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table 3: GMM estimation results

	BL	BH	WL	WH
Elasticity λ_B^o	0.154	0.180	0.075	0.044
	(0.015)	(0.015)	(0.014)	(0.009)
Elasticity λ_W^o	-0.146	-0.124	0.251	0.268
	(0.042)	(0.040)	(0.030)	(0.017)
Spatial decay of racial pref. ρ^o	0.674	0.747	0.229	0.291
	(0.181)	(0.167)	(0.047)	(0.048)
Size of disamenity g^o	0.215	0.229	0.263	0.204
	(0.101)	(0.099)	(0.057)	(0.044)

The table reports the GMM estimates. Cluster robust standard errors in parentheses.