Quantifying the Welfare Effects of Gentrification on Incumbent Low-Income Renters*

Robert French, Harvard Kennedy School†
Ashvin Gandhi, UCLA Anderson
Valentine Gilbert, Harvard Kennedy School

December 1, 2023

Please click here for the latest draft.

How does gentrification affect the welfare of incumbent residents of low-income neighborhoods? This paper investigates how low-income renters in gentrifying neighborhoods fare relative to renters in neighborhoods in the same metro area that stay poor. We link person-level administrative US Census data to construct an annual panel that tracks the earnings, workplaces, and residential addresses of over 1 million US low-income urban renter households through 2000–2019. We use these data to estimate a dynamic structural model of residential and workplace choice. We identify our model with skill-specific labor demand shocks to potential commuting destinations, constructed with geocoded establishment-level business data. We find that—because low-income renters are highly mobile—gentrification affects incumbent renters primarily by changing the characteristics of other neighborhoods in their choice sets. Our results imply that where low-income renters lived *within* US metro areas mattered comparatively less than which US metro areas they lived in.

*Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2358 (CBDRB-FY24-P2358-R10957, CBDRB-FY24-P2358-R10936). This research uses data from the Census Bureau's Longitudinal Employer Household Dynamics Program, which was partially supported by National Science Foundation grants SES-9978093, SES-0339191 and ITR-0427889; National Institute on Aging Grant AG018854; and grants from the Alfred P. Sloan Foundation. This research was supported by The Horowitz Foundation for Social Policy and a Stone Research Grant from Harvard Kennedy School's James M. and Cathleen D. Stone Program in Wealth Distribution, Inequality, and Social Policy.

†Job Market Paper. Email: robertfrench@g.harvard.edu. French thanks his advisors Ed Glaeser, Gordon Hanson, Mark Shepard, Stefanie Stantcheva, and Matthew Weinzierl for their patience, support, and guidance. We thank Nathaniel Baum-Snow, Tridevi Chakma, Rebecca Diamond, Myrto Kalouptsidi, Mengwei Lin, and Mikko Silliman for helpful discussions and feedback. We thank James Davis, Shahin Davoudpour, Shital Sharma, and Adam Shum for help with the Census data. Oleg Kolev, Kevin Wang, and Li Zhijian provided excellent research assistance. All errors are our own.

1. Introduction

Gentrification is associated with simultaneous increases in housing costs and changes in both public neighborhood amenities (e.g., schools and public safety) and private ones (e.g., restaurants and bars) (Couture and Handbury 2020; Su 2022). Residents may benefit from gentrification if they value the change in amenities by more than the amount that housing costs increase (Vigdor 2010). Conversely, residents may be harmed if the changes in amenities are insufficient to compensate for the rising housing costs. Residents with strong attachments to their home neighborhood are especially vulnerable to rising housing costs as they are less willing to relocate given any cost increase. This paper leverages extensive US Census Bureau data to quantify these trade-offs for incumbent renters in gentrifying neighborhoods throughout 2000–2019.[1]

Beginning in the 1990s and intensifying after the year 2000, gentrification transformed the socioeconomic composition of vast areas within American inner cities (Couture and Handbury 2023).[2] Between 2000 and 2017, the share of residents with a college degree in census tracts near their metro areas' central business districts (CBDs) increased by an average of 15 percentage points from a baseline of 24 percent. This contrasts with a 7-percentage-point average increase in suburban neighborhoods' shares of college graduates, reflecting a secular increase in educational attainment.[3] How this transformation of American inner cities affected the incumbent residents of gentrifying neighborhoods remains an open question.

The gentrification of inner-city America reversed postwar urban decline, in which middle- and upper-class households left inner cities in favor of suburban life (Jackson 1987; Boustan 2010; Mieszkowski and Mills 1993). This postwar suburbanization is considered a major contributor to concentrated inner-city disadvantage (Wilson 1987). As primarily white middle- and upper-class households left for the suburbs, inner-city violent crime increased tenfold (Cullen and Levitt 1999; Curci and Masera 2023), city public finances declined (Derenoncourt 2022), and many employers relocated to the suburbs (Kain 1968; Glaeser and Kahn 2001; Miller 2021). If

[1]We focus on low-income renter households because of their vulnerability to the financial costs of gentrification and the fact that, in 2010, 64 percent of urban housing units occupied by households with incomes below $50,000 were rented (Manson et al. 2022). Recent research does show the potential for homeowners to be harmed from gentrification by rising property taxes (Ding and Hwang 2020; Berry 2021; Fu 2022). Evaluating the welfare effects of gentrification on incumbent homeowners is an interesting question for future research.

[2]The economic forces causing demand for inner-city living to rise among college graduates were multifaceted. Rising top incomes increased demand for local service amenities concentrated in downtown neighborhoods (Couture et al. 2023) and raised the time-cost of commuting (Edlund, Machado, and Sviatschi 2022; Su 2022). Declining urban crime (Ellen and O'Regan 2010; Ellen, Horn, and Reed 2019), shifting preferences for urban amenities (Brueckner, Thisse, and Zenou 1999; Glaeser, Kolko, and Saiz 2001; Baum-Snow and Hartley 2020; Couture and Handbury 2020), evolving transportation infrastructure (LeRoy and Sonstelie 1983; Glaeser, Kahn, and Rappaport 2008), and delayed childbearing (Moreno-Maldonado and Santamaria 2022) all likely contributed to demand for downtown living. These forces were compounded by increases in the valuation of downtown amenities *caused* by the growing presence of college graduates (Berkes and Gaetani 2023; Diamond 2016; Guerrieri, Hartley, and Hurst 2013).

[3]See Appendix A for more details on neighborhood change since the turn of the century.

suburbanization was a significant factor behind these changes, its reversal over the past few decades might have caused environmental shifts favored by incumbent inner-city residents. Residents may have benefited from more local job opportunities, improved public amenities funded by rising land values, and greater access to private consumption amenities such as grocery stores and restaurants.

We conduct our analysis on deidentified person-level data from the US Census Bureau's Master Address File (MAF), which records the near universe of US adults' residential migration histories from 2000 onward. We link these data to persons' earnings and workplace locations from the employer–employee linked Longitudinal Employer–Household Dynamics (LEHD) database, which records the near universe of private sector, state, and local government workers' employment histories between 2000 and 2019. We finally link these data to person-level sociodemographic information from all American Community Survey (ACS) respondents (2005–2021) and property-level data from the Census Bureau's Master Address File Extract (MAF-X) and CoreLogic's residential property databases (2006–2019). Together, these data allow us to observe the residential locations, earnings, and workplaces of over 1 million low-income urban renter households across 50 large metro areas for up to 20 years during the most intense period of recent gentrification. Our panel data give us the distinct ability to study how gentrification affects the welfare of residents living in different neighborhoods within the same metro area—a hitherto underexplored question.

To quantify the welfare effects of gentrification on incumbent residents, we estimate a dynamic model of residential and workplace choice. In our model, heterogeneous agents choose their neighborhood and workplace locations each period to maximize their expected lifetime utility. Agents are subject to a rich set of moving costs, can accumulate neighborhood-specific capital, and are forward looking. They obtain flow utilities comprised of expected housing and nonhousing consumption, neighborhood amenities, and their accumulated neighborhood capital. Conditional on neighborhood rents, agents' expected consumption varies across neighborhoods because of differences in their commute time–discounted proximity to jobs. This feature of our model captures the fact that neighborhoods farther from employment centers are less desirable because of the increased financial cost of commuting (Le Barbanchon, Rathelot, and Roulet 2021).

We model neighborhood amenities as a function of neighborhoods' shares of college graduates. This choice is motivated by a literature documenting a robust positive relationship between the provision of local public and private amenities and the local share of college graduates (Glaeser, Kolko, and Saiz 2001; Diamond 2016; Autor, Palmer, and Pathak 2017; Su 2022; Almagro and Domínguez-Lino 2022; Hoelzlein 2023). We further allow neighborhood amenities to vary over time with unobserved factors not caused by changes in the local share of college graduates. While the evolution of these unobserved exogenous neighborhood amenities is not caused by shifts in the local college graduate share, they may nonetheless be correlated with households' residential location choices, presenting a challenge to identification.

We identify our model parameters by combining establishment-level employment data from the Census Bureau's Longitudinal Business Database (LBD) with advances in the quasi-experimental shift-share literature. Specifically, we construct two sets of instrumental variables (IVs) to disentangle preferences for observed neighborhood characteristics and unobserved exogenous neighborhood amenities. The first set of IVs aggregates skill-specific shocks to potential commuting destinations. The intuition behind these instruments is that as neighborhoods' access to high-skill employment opportunities improve, they become more desirable for college graduates and thus are more likely to gentrify. These sets of IVs build on recent work in Baum-Snow, Hartley, and Lee (2019) and Baum-Snow and Han (2023) that microfound measures of employment access with commuting data in a workplace choice model à la Tsivanidis (2022). Relative to prior work, our establishment-level business data make identification from national industry shocks plausible, allowing the location of business establishments to be correlated with unobserved neighborhood characteristics (Borusyak, Hull, and Jaravel 2022).

Our second set of IVs is motivated by the observation that gentrification tends to occur near neighborhoods with already high shares of college graduates (Guerrieri, Hartley, and Hurst 2013). For each census tract in our data, we construct distance-weighted measures of proximity to other neighborhoods' shares of college graduates. We then exploit the fact that our analysis spans 50 metro areas and interact our neighborhood-level proximity measures with core-based statistical area– (CBSA-) wide Bartik labor demand shocks. Identification then proceeds analogously to that under a difference-in-difference estimator: we compare differences in gentrification between neighborhoods near and far from already gentrified tracts in CBSAs experiencing large labor demand shocks to differences in gentrification between neighborhoods near and far from already gentrified tracts in CBSAs not experiencing large labor demand shocks.

We use our structural model to isolate the impact of changing neighborhood rents and shares of college graduates throughout 2000–2019 on the welfare of low-income incumbent renters. To do so, we use our parameter estimates to compute expected welfare separately for low-income renter households living in each low-income urban census tract in the year 2000. These calculations are based on the observed changes in the distribution of neighborhood rents and shares of college graduates throughout 2000–2019 but hold unobserved neighborhood amenities fixed at their 2000 levels. We then compare these measures to the expected welfare the same households would have obtained if the economy were instead in steady-state in the year 2000. This exercise yields census tract-level measures of how the changing distribution of neighborhood rents and shares of college graduates throughout 2000-2019 affected incumbent renters, as viewed from the standpoint of the year 2000. By finally comparing these measures across census tracts that gentrified and census tracts that stayed poor, we uncover how living in gentrifying neighborhoods affected the welfare of incumbent low-income renters.

Our welfare analysis implies that, *on average*, incumbent renters initially living in census tracts that gentrified after the year 2000 were not made significantly worse off relative to incumbent renters initially living in census tracts that remained poor. As discussed above, incumbent renters will experience welfare losses from gentrification if their willingness to pay for accompanying amenity changes is less than the rise in real rental costs. These welfare losses can become large if households have high moving costs, a strong degree of neighborhood attachment, or there are few desirable alternative neighborhoods in households' choice sets. Our estimates suggest that low-income renter households living in poor neighborhoods both valued living in neighborhoods with a marginally greater share of college graduates and had only moderate moving costs.[4] As a result, we estimate that low-income renters initially living in poor but gentrifying neighborhoods did not experience large welfare losses relative to their counterparts initially living in neighborhoods that stayed poor throughout 2000–2019.

Our modest estimated moving costs underlie a core insight of our welfare analysis. Namely, that where low-income renters lived within US metro areas mattered comparatively less than which US metro area they lived in from the standpoint of the year 2000. Because the average low-income renter household faced only modest costs from relocating to other low-income tracts within their metro area each period, changes to one's home census tract mattered less than changes to their home metro area overall. Policies directed at keeping metro areas broadly affordable for low-income renters may improve welfare for this population more than policies designed to ensure incumbent renters can remain in their home neighborhoods over extensive time horizons.

Relation to Literature. Our paper contributes to three strands of literature. First, we contribute to the literature on the welfare implications of spatial sorting. Grounded in the canonical spatial equilibrium models stemming from Rosen (1974) and Roback (1982), an empirical literature has sought to quantify the implications of urban spatial sorting for households differentiated by their educational attainment. Moretti (2013) studies the implication of cross-metro sorting for real income inequality, while Diamond (2016) incorporates the endogenous supply of citywide amenities to study the implications of cross-metro sorting for welfare inequality. Similarly, Su (2022) and Couture et al. (2023) examine the welfare implications of spatial sorting but focus on

[4]We find that low-income Black (non-Black) renters are willing to pay $1,224 ($312) in annual rents to live in a neighborhood with a 10% higher share of college graduates (all dollar-denominated welfare costs are calculated in year-2010 dollars). We additionally find the fixed cost of moving within one's own CBSA is $3,578 ($1,692) for Black (non-Black) households, and increases by $612 ($237) for households who have lived in the same census tract for at least five years. These moderately-sized structural moving cost estimates are informed by households' baseline residential mobility. We find that low-income renter households are highly mobile, with just 50.4 (51.7) percent of Black (non-Black) households remaining in their home census tracts for at least five years at a time.

within-metro sorting.[5] Other closely related papers focus on understanding the emergence of endogenously provided local amenities (Couture and Handbury 2020; Almagro and Domínguez-Lino 2022; Hoelzlein 2023; Glaeser, Luca, and Moszkowski 2023). With few exceptions, the existing literature quantifies the effect of spatial sorting on the expected welfare of *prospective* city residents.[6] Our paper instead exploits rich panel data to examine the impact of gentrification on the welfare of *incumbent* renters. [7] To the best of our knowledge, we are the first to rigorously examine how the welfare effects of gentrification vary across neighborhoods *within* cities.

Second, we contribute to the empirical residential choice literature that estimates households' willingness to pay for housing and neighborhood characteristics. Earlier static models of residential choice (Brock and Durlauf 2002; Bayer, Ferreira, and McMillan 2007; Vigdor 2010) have given way to dynamic models that account for moving frictions and forward-looking behavior (Kennan and Walker 2011; Bishop 2012; Bayer et al. 2016). Researchers have used these dynamic neighborhood choice models to estimate preferences over the racial composition of neighborhoods (Davis, Gregory, and Hartley 2023), the insurance value of rent control (Diamond, McQuade, and Qian 2018), the willingness to pay to avoid violent crime and air pollution (Bishop and Murphy 2019), and horizontally differentiated consumption amenities (Almagro and Domínguez-Lino 2022), among others. We contribute to this literature by providing estimates on low-income renters' preferences over welfare-relevant neighborhood characteristics, levels of neighborhood attachment, and a rich set of moving costs by combining our detailed census data with advances in the quasi-experimental shift-share literature.[8] We show how Black households appear to place less weight than do non-Black households on access to employment opportunities but more weight on neighborhoods with higher shares of college graduates. We also show how the cost of moving between neighborhoods varies with both the physical distance between neigbhorhoods as well as the *social distance* between neigbhorhoods, where we define social distance as the difference in neighborhoods' shares of college graduates.

Third, our exploratory analyses in Section 3 contribute to empirical research documenting

[5] See Kuminoff, Smith, and Timmins (2013) and Diamond and Gaubert (2022) for a comprehensive review of these and other papers examining the implications of spatial sorting on inequality. This research in turn contributes more broadly to the quantitative spatial equilibrium literature summarized in Redding and Rossi-Hansberg (2017).

[6] Balboni et al. (2020) is a notable exception, using a repeated static commuting model coupled with the exact-hat algebra of Dekle, Eaton, and Kortum (2008) to estimate the welfare impacts of transit infrastructure investments and the resulting sorting of households on the welfare of incumbent residents in Dar es Salaam, Tanzania.

[7] Urban housing policies such as rent control and eviction protections prioritize incumbent renters' welfare over that of landlords and residents unprotected by such policies (Glaeser and Luttmer 2003; Diamond, McQuade, and Qian 2018; Collinson et al. 2023; Abramson 2023). Understanding the difference in welfare effects between prospective and incumbent residents is thus critical to discerning the appropriate set of policy responses to gentrification.

[8] Our instrumental variable construction builds on Baum-Snow, Hartley, and Lee (2019) and Baum-Snow and Han (2023), who are the first to construct and microfound shocks to employment access in a model of workplace choice à la Tsivanidis (2022). Brummet and Reed (2021) and Glaeser, Luca, and Moszkowski (2023) use proximity to already gentrified tracts as an instrument for gentrification.

the effects of gentrification on observable outcomes for low-income residents. Much of this research has focused on gentrification's impact on the propensity of incumbent residents to leave their home neighborhoods (Freeman and Braconi 2004; Freeman 2005; Ellen and O'Regan 2011; Ding, Hwang, and Divringi 2016; Dragan, Ellen, and Glied 2020; Pennington 2021). This research has recently broadened to consider a wider range of outcomes. Baum-Snow, Hartley, and Lee (2019) examine the impact of neighborhood change on children's long-run outcomes. Brummet and Reed (2021) consider effects on employment and experienced neighborhood characteristics alongside effects on residential mobility. Ferreira, Kenney, and Smith (2023) explore minority households' local migration networks. Lester and Hartley (2014) and Meltzer and Ghorbani (2017) focus on the employment impacts of neighborhood change on incumbent residents. This literature broadly finds economically insignificant *average* effects on incumbents' household-level outcomes.[9] Statistical power, however, limits these studies' ability to examine heterogeneity across important dimensions such as the environment in households' origin neighborhoods. We advance this literature by employing our extensive panel data to show that these average results are robust to our accounting for the environment in incumbent households' origin neighborhoods. This is a surprising finding, as one might expect the impacts of gentrification to differ markedly based on local housing supply elasticities and the baseline level of neighborhood amenities, for example.

Roadmap. This paper is structured as follows. Section 2 introduces our data and our sample of low-income households. Section 3 defines our measure of gentrification and presents descriptive evidence on the effects of gentrification on low-income incumbent renters. Sections 4 and 5 detail our dynamic model of neighborhood and workplace choice. Section 6 outlines our identification strategy and reports our parameter estimates. Finally, Section 7 presents our welfare analysis, and Section 8 concludes.

2. Data and Sample Construction

We conduct our analyses using person- and establishment-level administrative micro data from the US Census Bureau spanning 2000–2019. Table 1 provides an overview of these data sources. We postpone to Appendix B the detailed discussion of the raw data and how we use it to construct our analysis samples, presenting only a cursory discussion here.

We conduct our analyses on an annual panel that records the earnings, workplaces, and residential addresses of over 1 million low-income urban renter households through 2000–2019.

[9] Baum-Snow, Hartley, and Lee (2019) find meaningful impacts on incumbent children's future credit outcomes, consistent with the potential for neighborhood environments to effect children's outcomes in adulthood (Chyn and Katz 2021).

TABLE 1. Data Overview

Source	Coverage	Description
A. Household Panel		
Master Address File – Auxiliary Reference File (MAF-ARF)	2000–2019	Annual address-level residential locations
Longitudinal Employer–Household Dynamics (LEHD) database	2000–2019	Annual earnings, workplace locations, basic sociodemographics
American Community Survey (ACS)	2005–19	Detailed sociodemographics
B. Housing Characteristics		
CoreLogic	2006–2017	Address-level housing transactions and multiple listing service entries
Master Address File Extract (MAF-X)	2019	Address-level unit characteristics
C. Business Data		
Longitudinal Business Database (LBD)	2000–2019	Establishment-level employment and payroll aggregates

Notes: The 2019 MAF-X is a continuously updated inventory of all known living quarters in the US. Addresses verified in the past, but that are no longer known living quarters, remain in the MAF-X except in rare circumstances. The MAF-X 2019 therefore contains an inventory of all known addresses spanning our entire sample period.

To construct this panel, we first form an annual panel of persons' residential histories using the MAF-ARF. We then merge to these residential histories persons' annual earnings and workplace locations from the LEHD and additional sociodemographic characteristics from the ACS. These merges are facilitated by a unique person identifier called a protected identity key (PIK), which is assigned to individuals across data sets by the Census Bureau via probabilistic linking (Wagner and Layne 2014). We aggregate earnings by housing unit and designate the highest earner of each unit as the household head for that year.[10] Our household panel is then restricted to persons who have been identified as a household head and who occupy a rental housing unit.[11]

We restrict our sample to household heads that are between 25 and 65 years of age. To focus our analysis on low-income households, we further restrict our sample to household heads earning in the bottom tercile of their respective CBSA and decadal age band in the year that they were first assigned household-head status.[12] We finally limit our study to household heads living in the urban cores of the 50 largest CBSAs, located within the 28 states for which data are

[10]We define housing units by their addresses in the MAF-X. Persons must have positive earnings in the given year to be considered a household head.

[11]We describe how we impute rental unit status in Appendix B. We also detail in Appendix B how we smooth residential histories and handle changes to household formations, out-of-sample migrations, and missing observations.

[12]CBSAs consist of counties associated with an urban core of at least 10,000 persons and adjacent counties deemed integrated through commuting ties. CBSAs are defined by the Office of Management and Budget (OMB).

accessible in the LEHD. We define urban cores using a method similar to the ones in Hwang and Lin (2016) and Couture and Handbury (2020). We denote as urban cores the set of census tracts associated with each CBSA that contain the 50 percent of the CBSA's population that is closest to its central business district (CBD).[13] Table 2 presents descriptive statistics from a 2010 cross-section of our panel.[14]

3. Exploratory Analyses

This section reports findings from descriptive regressions of gentrification on an array of household-level outcomes. Our findings motivate a structural analysis and inform key features of our counterfactual welfare analyses. Before presenting our findings, however, we define our measure of gentrification and discuss our analysis period.

Gentrification. We follow the existing literature and define gentrification at the census tract level (e.g., Ding and Hwang (2020), Dragan, Ellen, and Glied (2020), Brummet and Reed (2021)). Specifically, we define gentrification from period t_0 to t as the increase in the number of college graduates in a census tract, n, during those years. We then normalize this measure by tracts' total adult population in time t_0:

$$\text{Gent}_{n,t_0,t} \equiv \frac{\text{College}_{n,t} - \text{College}_{n,t_0}}{\text{Adult Population}_{n,t_0}}$$

We follow the economics literature by normalizing the change in the number of college graduates by the total adult population in t_0 (Brummet and Reed 2021; Card, Mas, and Rothstein 2008; Böhlmark and Willén 2020). One alternative definition is the change in the *share* of college graduates between periods t and t_0. We nonetheless follow the literature's convention since this measure minimizes the mechanical relationship between gentrification and a primary outcome of interest: the out-migration rates of incumbent renters. This is because the out-migration rates of low-income incumbent renters have little influence on our measure of gentrification; only 11 percent and 16.5 percent of Black and non-Black households in our sample possess a college

[13] Our CBD definitions come from Fee and Hartley (2013). These CBD definitions are with respect to 2008 CBSA delineations. To maintain consistency with these CBD definitions, we therefore use the 2008 delineations of CBSAs throughout our analysis. Moreover, while each CBSA is associated with a primary urban center, some CBSAs additionally contain secondary urban centers called metropolitan divisions that have their own CBD. Although our low-income cutoff is constructed from the earnings distribution of the entire CBSA, our urban core cutoffs are specific to each urban center's population, including metropolitan divisions. We believe that these choices best capture our target population of low-income urban residents.

[14] The 2010 cross-section captures characteristics of the full sample that we use to estimate our descriptive regressions presented in Section 3. We are postponing the release of sample statistics for the complete panel to help streamline the US Census Bureau disclosure review processes.

TABLE 2. Sample Characteristics: Household Heads in 2010

	Black Households	Non-Black Households
Panel A: Household Head Characteristics		
Household Income	21,250	21,660
	(14,220)	(13,570)
Commute Time	27.64	25.6
	(12.43)	(12.82)
College Degree	0.11	0.165
	(0.313)	(0.371)
Immigrant	0.212	0.45
	(0.409)	(0.498)
Age	42.38	43.08
	(10.64)	(10.8)
Female	0.611	0.519
	(0.488)	(0.5)
Household Size	2.544	2.505
	(1.549)	(1.578)
Parent	0.275	0.266
	(0.447)	(0.442)
Panel B: Household Heads' Tract Characteristics		
Median Rents	834.4	934.7
	(232.4)	(286.2)
Median Property Value	247,000	289,200
	(205,100)	(214,600)
Share White	0.435	0.692
	(0.264)	(0.208)
Share College Educated	0.219	0.288
	(0.147)	(0.183)
Share College Educated and White	0.124	0.213
	(0.137)	(0.169)
Distance to CBD	9.552	9.56
	(5.172)	(5.33)
Unique Households	314,000	688,000

Notes: Table 2 reports mean characteristics for household heads with standard errors in parentheses. The sample comprising Table 2 consists of all household heads in the 2010 cross-section of the panel. Panel A reports household-level characteristics during 2010. Panel B reports characteristics of household heads' census tracts, also in 2010. Dollars are deflated to 2010 levels, and census tracts are delineated by 2010 boundaries. Commute time is measured in minutes. The Parent variable is calculated only for household heads present in the ACS 2005–2021 and is inferred by the reported age of the child in the year in which the household head is a survey respondent. The college degree variable is computed only for PIKs for whose education variables are not imputed in the LEHD or for whom we ascertain educational attainment through our ACS surveys. Details on the construction of tract aggregates are in Appendix B. *Sources*: ACS (2005–2021), LEHD (2010), CoreLogic (2006–2017), MAF-X (2019), and MAF-ARF (2010). Sample characteristics were disclosed by the US Census Bureau's Disclosure Review Board. Project Number 2358. Disclosure Clearance Number CBDRB-FY24-P2358-R10936.

degree, and our sample comprises a small fraction of total households in each neighborhood.[15] Our choice to define gentrification on the basis of educational attainment also follows convention in the economics literature (Diamond 2016; Brummet and Reed 2021; Su 2022).[16]

Analysis Period. We choose to focus on the years 2010–2019 to establish our descriptive findings. We do so for two primary reasons. First, this choice mitigates the influence of potentially confounding factors resulting from the great recession, especially since we can control for changing neighborhood-level characteristics prior to 2010. Second, restricting our panel to 2010–2019 allows us to control for household characteristics throughout 2000–2009. Household characteristics such as length of prior residential tenure are likely correlated with households' residential location choices.

3.1. Is Gentrification Associated with Higher Neighborhood Out-Migration Rates?

Gentrification is not associated with higher neighborhood out-migration rates among low-income incumbent renter households. We show that this finding holds across a range of baseline neighborhood environments, including census tracts that are already partially gentrified and that are highly developed.

To understand the relationship between gentrification and incumbent renter households' out-migration rates, we estimate a set of Cox proportional hazard models. These models estimate the impact that gentrification has on the probability an incumbent household leaves its origin neighborhood in any one year (i.e., on incumbents' hazard rates).[17] Our Cox models take the following form:

$$\log(h(t|i)) = \alpha^{Cox} + \beta^{Cox}_{\text{NC}}\text{Gent}_{n(i),10,19} + \gamma^{Cox}X_i + \delta^{Cox}X_{n(i)} + \alpha^{Cox}_{CBSA} + \varepsilon^{Cox}_i \tag{1}$$

where $h(t|i)$ is the hazard in period t for household i. $n(i)$ denotes household i's origin neighborhood, X_i is a vector of household-level controls, $X_{n(i)}$ is a vector of controls characterizing

[15] All the results that we present below are quantitatively similar to those from specifications that exclude college-educated adults from our sample of low-income renters.

[16] Some research defines gentrification on the basis of changes in income (e.g., Dragan, Ellen, and Glied (2020) and Ding and Hwang (2020)). Other research considers gentrification through the lens of changing racial compositions (Baum-Snow and Hartley 2020). We nonetheless believe that focusing on educational composition offers the clearest connection to the existing literature. We have experimented with alternative definitions based on racial composition and income and found quantitatively similar results and will disclose these results in due course.

[17] Our choice to estimate Cox models is motivated by incumbent renters' short unconditional neighborhood tenures. Only 50.4 (51.7) percent of Black (non-Black) incumbent renter households remained in their 2010 origin census tract until at least 2015; these unconditional survival probabilities fall to 31.3 and 32.8 percent for 2019, respectively. Existing research that relies on intermittent sampling of residents loses potential identifying variation from incumbent residents with short unconditional residential tenures. The Cox proportional hazard model allows us to efficiently utilize our annual panel of residential histories to identify the effects of gentrification on incumbent renters' neighborhood tenures.

the origin neighborhood of household i, and α^{Cox}_{CBSA} is a CBSA-level fixed effect. We detail and motivate our choice of controls in Appendix C. To mitigate concerns over sample selection, we often restrict our sample to longtime renters, defined as renter households that have lived in their origin tract for at least five years prior to 2010. We also postpone discussion of this choice and identification more broadly to Appendix C. Throughout our descriptive analyses, we cluster standard errors at the census tract level, which is our treatment unit (Abadie et al. 2023).

Estimates of β^{Cox}_{NC} from equation 1 are reported in Panel A of Table 3. Columns (1) and (2) of Panel A in Table 3 report estimates for our full sample of renter householders separately for Black and non-Black headed households. Columns (3) and (4) report the same estimates but restrict to longtime renters. Finally, columns (5) and (6) further restrict to census tracts with an initial share of college graduates below the sample-weighted median among all tracts in our sample. For all subsets of our data, we document an economically insignificant relationship between gentrification and incumbent renters' hazard rates.

Consider the effect of neighborhood change on non-Black longtime incumbent renters' hazard rates (column (4) in Panel A of Table 3). A 10-percentage-point increase in gentrification corresponds to a 1.87 percent increase in the probability that these renters leave their origin neighborhood in any given year between 2010 and 2019. Since the unconditional probability of leaving one's origin neighborhood in any one year (i.e., baseline hazard) reaches at most 20 percent, these effect sizes are negligible.[18] These results are consistent with the extant literature attributing neighborhood change to changes in in-migration patterns as opposed to increased out-migration among incumbent residents (e.g., Ding, Hwang, and Divringi (2016), Dragan, Ellen, and Glied (2020), and Brummet and Reed (2021)).

We advance our understanding of the impacts of gentrification on out-migration by using our expansive panel data to show that these null results do not mask meaningful underlying heterogeneity. Column (5) shows that, if anything, out-migration rates *decrease* among Black low-income renters in neighborhoods with an initially low share of college graduates. A 10-percentage-point increase in gentrification corresponds to a 7.7 percent decrease in the probability that a longtime incumbent Black renter leaves her origin neighborhood if it has an initially low share of college graduates. Again, however, with a baseline hazard rate of at most 20 percent, these are economically small effect sizes. Table A2 in Appendix C explores additional sources of potential heterogeneity, including the density of urban development. We continue to find economically insignificant effect sizes. Finally, we test the robustness of our Cox proportional hazards models by estimating linear probability models where the dependent variable is an indicator equal to one if an incumbent renter remains in her origin census tract

[18] A ten-percentage-point increase in gentrification is associated with at most a $(20/100) \times 1.87 = 0.374$ percentage point increase in the probability that an incumbent renter household leaves its origin neighborhood in any one year.

until at least 2019.[19] We obtain similar results in this framework.

3.2. Is Gentrification Associated with Changing Economic Outcomes for Incumbent Renters?

Gentrification is not meaningfully associated with changing economic outcomes for incumbent renters. This is true across most baseline neighborhood environments. We do, however, find that gentrification is associated with lower average annual earnings for longtime Black renters living in tracts with an initially low share of college graduates.

To understand the relationship between gentrification and incumbent renters' economic outcomes, we estimate linear regression models regressing changes in incumbent residents' annual earnings and commute distances on gentrification,

$$\Delta y_i = \alpha^{LP} + \beta^{LP}_{\mathrm{NC}} \mathrm{Gent}_{n(i),10,19} + \gamma^{LP} X_i + \delta^{LP} X_{n(i)} + \alpha^{LP}_{CBSA} + \varepsilon^{LP}_i \tag{2}$$

where Δy_i denotes the change in either annual earnings or commute distances between 2010 and 2019. X_i, $X_{n(i)}$, and α^{LP}_{CBSA} are the same set of control variables and fixed effects used in equation 1. Estimates of β^{LP}_{NC} are reported in Panel B of Table 3. To interpret the magnitude of the coefficients in Panel B of Table 3, consider the effect of gentrification on Black incumbent renters' annual earnings and commute times. A 10-percentage-point increase in our measure of gentrification is associated with \$287 higher annual earnings and .2 minute quicker commutes, economically small effect sizes. We do find some evidence, however, that gentrification is associated with lower annual earnings for longtime Black incumbent renters in tracts with a low initial share of college graduates. Here, a 10-percentage-point increase in gentrification is associated with \$914 lower average annual earnings after 10 years.

3.3. Is Gentrification Associated with Changing Neighborhood Conditions for Incumbent Renters?

Gentrification is associated with meaningfully changes in the neighborhood conditions incumbent renters experience. That is, incumbent renters who initially live in gentrifying census tracts experience greater changes in their neighborhood conditions than similar residents in census tracts not gentrifying. This is not a mechanical result, as incumbent renters are free to move across tracts throughout our analysis period.

Panel A in Table 3 reports the relationship between gentrification and changes in incumbent renters' neighborhood conditions. For example, a ten-percentage-point increase in gentrification

[19] Cox proportional hazards models assume that treatment effects are constant over time (i.e., proportional hazards). Our linear probability model specification mimics the existing literature that relies on intermittent sampling of residents (e.g., Brummet and Reed (2021)).

TABLE 3. Effect of Gentrification on Incumbent Renters

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Cox Proportional Hazards Model Outcomes						
Hazard Rate	-0.0688	-0.107*	-0.0793	0.187*	-0.777***	-0.264
	(0.0542)	(0.0439)	(0.132)	(0.0781)	(0.286)	(0.325)
Panel B: Linear Regression Model Outcomes						
Leave Tract	-0.0176	-0.0263	-0.016	0.07*	-0.305***	-0.0888
	(0.0223)	(0.0174)	(0.0509)	(0.0278)	(0.106)	(0.0894)
Annual Earnings	2,872**	2,268**	-932	-86.05	-9,138*	-551.8
	(948.8)	(736.4)	(2,170)	(1,375)	(3,854)	(3,427)
Commute Time	-0.2280	-1.905***	0.637	0.5900	-1.854	3.146
	(0.776)	(0.531)	(1.609)	(1.002)	(3.132)	(2.775)
Rent	0.271***	0.0960***	0.389***	0.222***	0.863***	0.607***
	(0.0472)	(0.0288)	(0.0620)	(0.0329)	(0.145)	(0.096)
College Share	0.726***	0.299***	1.079***	0.567***	3.726***	4.108***
	(0.204)	(0.0598)	(0.131)	(0.0788)	(0.250)	(0.236)
Controls						
Full Controls	✓	✓	✓	✓	✓	✓
CBSA Fixed Effects	✓	✓	✓	✓	✓	✓
Sample Restrictions						
Race	Black	Non-Black	Black	Non-Black	Black	Non-Black
Longtime Renters			✓	✓	✓	✓
Low Initial College Share					✓	✓
N (1,000s)	314	688	56	156	35	67

Notes: We discuss how to interpret all coefficients in the main text. Leave Tract is an indicator equal to 1 if the household leaves its origin tract before 2019. Annual earnings are measured in 2010 dollars. Rent and college share are measured in percent changes. Commute times are measured in minutes. Every specification includes the full set of controls listed and detailed in Appendix C. Standard errors in parentheses are clustered at the census tract level. Longtime renters are renters who have lived in their origin census tract since at least 2005. Tables A3 and A2 in Appendix C report results for a wider range of baseline neighborhood environments. *Sources*: ACS (2005–2021), LEHD (2010), CoreLogic (2006–2017), MAF-X (2019), and MAF-ARF (2010). Estimates were disclosed by the US Census Bureau's Disclosure Review Board. Project Number 2358. Disclosure Clearance Number CBDRB-FY24-P2358-R10936.

is associated with Black incumbent renters living in tracts during 2019 that had on average 2.7 percent higher rents and 7.3 percent higher shares of college graduates. We see in columns (5) and (6) that gentrification was strongly associated with the neighborhood conditions of incumbent renters initially residing in tracts with a low share of college graduates.

Our estimates on experienced neighborhood characteristics are suggestive of moving frictions for at least some low-income renter households in our sample. Without moving frictions, renters would simply reoptimize their location choice each period to ensure proximity to their ideal bundle of neighborhood characteristics, yielding economically insignificant estimates on incumbent renters' neighborhood conditions. The presence of moving frictions, in turn, suggests the potential for differences in welfare effects from gentrification across tracts *within* CBSAs, motivating our paper's focus on incumbent renters. Moving frictions make incumbent renters averse to leaving their home census tract irrespective of changes in its characteristics. Whether incumbent renters then benefit from gentrification depends on their relative valuations of rents, job market access, and amenities vis-à-vis the actual change in these neighborhood characteristics. Our structural analysis in Section 4 assesses whether the implied moving frictions translate into large average moving costs and quantifies their importance for households' welfare.

It is worth noting here that there is limited correspondence between household residential mobility and welfare. Indeed, our estimates documenting no meaningful increase in neighborhood exit rates in response to gentrification are consistent with either positive or negative welfare effects from gentrification, depending on the aforementioned trade-off between rents, job market access, and amenities. Similarly, if we instead observed economically significant increases in neighborhood exit rates in response to gentrification, we would need to understand whether these estimates reflect insignificant moving frictions and dense choice sets or declines in incumbents' relative valuations of their origin neighborhoods (Vigdor 2002). We now turn to estimating a dynamic model of residential and workplace choice to estimate the welfare effects from gentrification on incumbent low-income renter households.

4. A Dynamic Model of Neighborhood and Workplace Choice

To quantify the welfare impact of neighborhood change on low-income incumbent renters, we estimate a single-agent dynamic discrete neighborhood and workplace choice model (Bayer et al. 2016; Diamond, McQuade, and Qian 2018; Davis et al. 2021; Almagro and Domínguez-Lino 2022; Davis, Gregory, and Hartley 2023). Our single-agent framework considers the neighborhood and workplace choice problem that low-income renter households face each period, treating neighborhood and workplace characteristics as exogenous.[20] We use this framework to obtain parameter estimates of low-income renter households' preferences over neighborhood characteristics pertinent to understanding the welfare effects of gentrification.

[20]It is common to treat neighborhood characteristics as exogenous in estimating households' preferences even when these neighborhood characteristics are functions of neighborhoods' socioeconomic composition (Bayer et al. 2016; Davis et al. 2021; Davis, Gregory, and Hartley 2023). We discuss the restrictions that this assumption imposes on our welfare calculations below.

Among papers in the dynamic discrete choice literature, our setup is most similar to the neighborhood demand model of Almagro and Domínguez-Lino (2022), who analyze the endogenous formation of horizontally differentiated private consumption amenities in the context of Amsterdam's 2010–2019 tourism boom. One important departure from their demand model is that we incorporate differences in within-CBSA access to employment opportunities through a first-step workplace choice problem. We combine our LEHD and LBD data to compute microfounded, time-varying, and neighborhood-level measures of job access for our sample population. This addition models an important determinant of households' location choices (e.g., Su (2022), Gu et al. (2021)). Moreover, coupled with our job market access instrument described in Section 6, these measures help facilitate credible identification of households' preferences.

We use our estimated model parameters to quantify the welfare effects of gentrification for incumbent renters in Section 7. To do so, we use our parameter estimates to compute expected welfare separately for renter households living in each low-income urban census tract in the year 2000. We do this using the observed distribution of neighborhood rents and shares of college graduates throughout 2000–2019. We then compare these measures to the expected welfare that the same renter households would have obtained if the economy were instead in steady steady in the year 2000. This exercise yields census tract–level estimates of the welfare effects of gentrification for incumbent renter households. We finally conduct counterfactual experiments to unpack these welfare effects.

4.1. Households and Timing of Choices

In each period, t, household heads, i, must decide which neighborhood in the city, c, they should live in.[21] In addition to choosing their residential neighborhood, $n_{i,t}$, household heads must decide which neighborhood to work in, $m_{i,t}$. Their workplace choice simply maximizes commute time–discounted period income and is decided after neighborhood residence is known. Longer commute times reduce the time that household heads spend working, effectively discounting the wage offered in workplace m. Finally, conditional upon deciding where to live and work, households must then decide how much to spend on housing given the neighborhood-wide and period-specific rental rate.

Households differ ex ante with respect to the household head's race, which we denote by $k \in \{\text{Black, Non-Black}\}$. Households can further differ in their previous neighborhood residences, which inform both their current neighborhood residence, $n_{i,t-1}$, and the how long they have lived there as of period $t-1$, $\tau_{i,t-1}$. We collect these observable household-level state variables in $x_{i,t} \equiv (n_{i,t-1}, \tau_{i,t-1})$.

[21]Households take their city as given.

Just before households make their workplace and neighborhood choices in each period, they receive idiosyncratic productivity and preference shocks, respectively. These preference shocks are unobserved by the econometrician but rationalize observed variation in households' choices within types conditional on $n_{i,t-1}$ and $\tau_{i,t-1}$. The remainder of this section presents the household head's problem, starting with her workplace choice problem.

4.2. Workplace Choice

Upon making their residential neighborhood and housing consumption choices in period t, households receive two independent productivity shocks. The first productivity shock is denoted by b_t^{kc} and is common across all type-k households in city c. The second productivity shock is household- and workplace tract–specific. This second productivity shock is denoted by $z_{m,t}^i$, where m denotes the workplace tract. Conditional on living in neighborhood n, households choose their work location to maximize their commute time–discounted income:

$$I_{n,t}^k \equiv b_t^{kc} \cdot \max_m \frac{z_{m,t}^i}{d_{n,m}} w_{m,t},$$

where $w_{m,t}$ is the wage offered in workplace tract m and period t measured in efficiency units. $d_{n,m} > 0$ is the time that it takes to commute between neighborhood n and m. We assume that households spend a fixed amount of time each day working or commuting, so $d_{n,m}$ effectively discounts the total wage offered in m: $z_{m,t}^i \cdot w_{m,t}$. We assume that $z_{m,t}^i$ is drawn independently from a Frechet distribution with shape parameter $\epsilon^c \, \forall \, i \in c$. These shape parameters are specific to each city, which we make explicit with the superscript c. We further assume the Frechet shocks are independent across years, implying no cost to switching jobs. The expected income for a type-k household living in tract n at time t is therefore

$$\bar{I}_{n,t}^k = \Gamma\left(1 - \frac{1}{\epsilon^c}\right) \cdot b_t^{kc} \cdot RMA_n^{1/\epsilon^c}$$

where $\Gamma(\cdot)$ is the gamma function and $RMA_n \equiv \sum_{m \in \mathcal{N}^c} \left(\frac{w_m}{d_{n,m}}\right)^{\epsilon^c}$ is a summary measure of access to employment, which we follow the literature in terming residential market access. We derive these equations and describe how we construct their empirical analogues in Appendix D.3.

4.3. Neighborhood Choice

Households' Neighborhood Choice Problem. Households choose their residential locations to maximize the sum of their expected discounted utilities,

$$\max_{\{n \in \mathcal{N}^c\}_t^\infty} \quad \mathbb{E}\left[\sum_{t'=t}^{\infty} \delta^{t'-t} \cdot u_n^k\left(s_{i,t'}^c\right) \,\middle|\, \mathcal{I}_{i,t}\right] \tag{3}$$

where δ is a known discount factor and $s_{i,t'}^c$ is a vector of state variables that determine household i's flow utility u_n^k from choosing neighborhood n. $s_{i,t'}^c$ includes the measures of expected income, $\bar{I}_{n,t}^k$, derived in Section 4.2. $\mathbb{E}\left[\cdot|\mathcal{I}_{i,t}\right]$ denotes the expectation operator conditioned on household i's information set at time t. $\mathcal{N}^c \equiv \{OO^c, 1^c, \ldots, N^c\}$ is the city-specific choice set, where OO^c denotes the outside option of leaving the city entirely. In each period, households observe the state variables s_{it}^c before choosing their residential location. Flow utilities are then realized, and states evolve. Each household's information set $\mathcal{I}_{it}$ therefore includes all current and past state variables that households may use to form expectations over their future evolution. We specify households' belief formation in Section 5.1.

State Variables. Households' flow utilities depend on the vector of state variables $s_{it}^c \equiv \left(x_{it}, \varepsilon_{int}, \omega_t^c, \xi_t^{kc}\right)$, where $\left(x_{it}, \varepsilon_{int}\right)$ are household-level observable and unobservable state variables, respectively. By contrast, $\left(\omega_t^c, \xi_t^{kc}\right)$ are city-specific observable and unobservable state variables, respectively. The household-level observable state variables, x_{it}, are comprised of households' residential tenure and neighborhood choice in the previous period, $x_{it} = (n_{it-1}, \tau_{it-1})$. The evolution of these observable household-level state variables is determined by household i's residential choices. ε_{int} is the household's unobservable state, which we assume is i.i.d. across households, neighborhoods, and time. We conceptualize ε_{int} as an unobserved-to-the-econometrician time-varying household and neighborhood-specific preference shock. As is common, we assume that ε_{int} is distributed according to a type I extreme value distribution.

Observable city-specific state variables are denoted by ω_t^c. The collection of city-specific state variables includes vectors for each neighborhood's housing costs, r_{nt}^k, the share of college graduates in the neighborhood, $\frac{Coll_{nt}}{Pop_{nt}}$, each neighborhood's commute time–discounted expected income, $\bar{I}_{nt}^k$, and an index for the time period t[22]:

$$\omega_t^c = \left(\{r_{nt}\}_{n \in \mathcal{N}^c}, \left\{\frac{Coll_{nt}}{Pop_{nt}}\right\}_{n \in \mathcal{N}^c}, \left\{\bar{I}_{nt}^k\right\}_{n \in \mathcal{N}^c}, t\right)$$

[22]We include a time index in the set of observable neighborhood-level state variables to explicitly incorporate nonstationarity, so that the remaining observed state variables' evolutions can depend on time.

Finally, ξ_t^{kc} is a city-specific vector of unobservable time-varying neighborhood-level amenity valuations among type k households. For example, ξ_t^{kc} could include time-varying valuations among type-k residents for suburban life, independent of gentrification. To facilitate exposition, we define $\bar{\omega}_t^{kc} \equiv \left(\omega_t^c, \xi_t^{kc}\right)$ as the vector containing both observable and unobservable city-specific state variables.

Flow Utility. Preferences over neighborhood characteristics net of moving costs for a type-k household can be represented by

$$A_{n,t}^k \, Q_{n,t}^k \, \tau_{it}^{\beta_\tau^k} \, \exp(\varepsilon_{int})$$

where $A_{n,t}^k$ is a type-k's valuation of amenities in neighborhood n and $Q_{n,t}^k$ is a consumption composite that is Cobb–Douglas over nonhousing consumption, $C_{n,t}^k$, and housing consumption, $H_{n,t}^k$:

$$Q_{n,t}^k \equiv \left(C_{n,t}^k\right)^{\beta_C^k} \left(H_{n,t}^k\right)^{1-\beta_C^k}$$

Households' expected period- and neighborhood-specific budget constraint is given by

$$C_{n,t}^k \geq \bar{I}_{n,t}^k - r_{n,t} \cdot H_{n,t}^k$$

Type-specific neighborhood amenities are

$$A_{n,t}^k \equiv \left(\frac{Coll_{nt}}{Pop_{nt}}\right)^{\beta_A^k} \exp\left(\xi_{nt}^k\right)$$

We can decompose unobserved neighborhood- and period-specific amenities, ξ_{nt}^k, into time-invariant neighborhood-specific components, time-varying city-level components, and neighborhood-specific time-varying components:

$$\xi_{nt}^k \equiv \alpha_n^k + \alpha_t^{kc} + \tilde{\xi}_{nt}^k$$

Taking logs, incorporating moving costs (defined below), solving for expected optimal housing consumption, and substituting in the amenity specification yields the following expected flow utility specification for a type-k household choosing neighborhood n with state s_{it}^c that is consistent with households' preferences over neighborhood characteristics[23]:

$$u_n^k\left(s_{it}^c\right) = \alpha_n^k + \alpha_t^{kc} + \beta_w^k \ln\left(\bar{I}_{n,t}\right) - \beta_r^k \log\left(r_{n,t}\right) + \beta_A^k \ln\left(\frac{Coll_{nt}}{Pop_{nt}}\right)$$
$$+ \beta_\tau^k \ln(\tau_{it}) - MC_t^k(n_t, n_{it-1}) + \tilde{\xi}_{nt}^k + \varepsilon_{int}$$

Moving Costs. If a type-k household decides to leave its current neighborhood for another neighborhood in the same city, it incurs a nonmonetary moving cost, $MC_t^k(n_{it}, n_{it-1})$, that is composed of a fixed disutility from moving, the physical straight-line distance between the household's origin and destination neighborhoods, and the "social distance" (defined below) between these two neighborhoods. Conversely, if a type-k household decides to leave its city entirely, it incurs a single, city-specific fixed cost. Specifically,

$$MC_t^k(n_t, n_{t-1}) = \begin{cases} 0 & \text{if } n_t = n_{t-1} \\ MC^k + \beta_d^{k\prime} \boldsymbol{d}(n_t, n_{t-1}) + \beta_s^{k\prime} \boldsymbol{s}(n_t, n_{t-1}) & \text{if } n_t \neq n_{t-1} \text{ and } n_t, n_{t-1} \neq OO^c \\ MC^{kc} & \text{if } n_{it} \neq n_{t-1} \text{ and } n_t \text{ or } n_{t-1} = OO^c \end{cases}$$

where MC^k and MC^{kc} are the fixed intensive- and extensive-margin moving costs, respectively. $\boldsymbol{d}(n_t, n_{t-1})$ is a vector describing the physical distance between n_t, n_{t-1} and $\boldsymbol{s}(n_t, n_{t-1})$ is a vector describing the social distance between n_t, n_{t-1} in period t:

$$\boldsymbol{d}(n_t, n_{t-1}) \equiv \begin{bmatrix} |Dist(n_t, n_{t-1})| \\ |Dist(n_t, n_{t-1})|^2 \end{bmatrix} \quad \boldsymbol{s}_t(n_t, n_{t-1}) \equiv \begin{bmatrix} |(S(n_t) - S(n_{t-1}))/(S(n_t) + S(n_{t-1})| \\ |(S(n_t) - S(n_{t-1}))/((S(n_t) + S(n_{t-1}))|^2 \end{bmatrix}$$

where $Dist(n_t, n_{t-1})$ is the straight-line distance between the centroids of neighborhood n_t and n_{t-1} and $S(n_t)$ is the share of college graduates in neighborhood n at time t.[24] Our measure of social distance captures the fact that, while low-income renter households may value residing in neighborhoods with a high share of college graduates, it may be costly to assimilate to neighborhood environments different from one's own (Gans 1982; Jargowsky 2009). Indeed, recent experimental research suggests that low-income households' moving costs are poorly approximated by the physical distance of residents' potential moves but strongly predicted by differences in the sociodemographic composition of households' origin and potential destination neighborhoods (Bergman et al. 2023).

[23] Flow utility is expected in that it is the value that households expect to obtain before their period-specific workplace tract productivity shocks are realized. It is this flow utility specification that is relevant for households' neighborhood choice and thus for our structural estimation.

[24] We normalize our social distance measure by the sum of shares so that the percentage changes are invariant to the direction of the residential move.

Value Functions, Choice Probabilities, and Expectational Errors. We denote $V^k(s^c_{it})$ as the value function of the dynamic programming problem associated with equation 3. By Bellman's principle of optimality,[25]

$$V^k\left(s^c_{it}\right) = \max_{n\in\{OO^c,1^c,\ldots,N^c\}}\left\{\mathbb{E}_{x'|x,n}\left[u^k_n\left(s^c_{it}\right)\right] + \delta\mathbb{E}_t\left[V^k\left(s^c_{it+1}\right)|n, s^c_{it}\right]\right\}$$

We define household i's ex ante continuation value function as the expectation of the value function with respect to ε_{int}:

$$\bar{V}^k\left(x_{it}, \bar{\omega}^{kc}_t\right) \equiv \int V^k\left(s^c_{it}\right) dF^{\varepsilon}(\varepsilon_{int}) \tag{4}$$

and define household i's conditional value function as

$$\begin{aligned} v^k_n\left(x_{it}, \bar{\omega}^{kc}_t\right) &\equiv \mathbb{E}_{x'|x,n}\left[u^k_n\left(s^c_{it}\right)\right] - \varepsilon_{int} + \delta\mathbb{E}_t\left[\bar{V}^k\left(x_{it+1}, \bar{\omega}^{kc}_{t+1}\right)|n, x_{it}, \bar{\omega}^{kc}_t\right] \\ &\equiv \bar{u}^k_n\left(x_{it}, \bar{\omega}^{kc}_t\right) + \delta\mathbb{E}_t\left[\bar{V}^k\left(x_{it+1}, \bar{\omega}^{kc}_{t+1}\right)|n, x_{it}, \bar{\omega}^{kc}_t\right] \end{aligned} \tag{5}$$

Then, given our assumption that ε_{int} are distributed i.i.d type I extreme value, the probability that a type-k household with state variables $\left(x_{it}, \bar{\omega}^{kc}_t\right)$ chooses neighborhood n in period t is given by

$$p^k_n\left(x_{it}, \bar{\omega}^{kc}_t\right) = \frac{\exp\left(v^k_n\left(x_{it}, \bar{\omega}^{kc}_t\right)\right)}{\sum\limits_{n'\in\mathcal{N}^c}\exp\left(v^k_{n'}\left(x_{it}, \bar{\omega}^{kc}_t\right)\right)}, \tag{6}$$

and the ex ante value function in 4 has the value

$$\bar{V}^k\left(x_{it}, \bar{\omega}^{kc}_t\right) = \ln\left(\sum_{n\in\mathcal{N}^c}\exp\left(v^k_n\left(x_{it}, \bar{\omega}^{kc}_t\right)\right)\right) + \gamma$$

where γ is Euler's constant. Combining these two expressions yields the following well-known result, which is critical to deriving our estimating equations (Hotz and Miller 1993):

$$\bar{V}^k\left(x_{it}, \bar{\omega}^{kc}_t\right) = v^k_n\left(x_{it}, \bar{\omega}^{kc}_t\right) - \ln\left(p^k_n\left(x_{it}, \bar{\omega}^{kc}_t\right)\right) + \gamma \tag{7}$$

Another expression critical for deriving our estimating equations is the difference between

[25]The expectation operator $\mathbb{E}_{x'|x,n}[\cdot]$ is with respect to the future value of households' observed household-level state variables, x', conditional on households' current state and on their neighborhood choice. While the current deterministic setup renders this operator redundant, we include it here to be consistent with our empirical application that models the evolution of households' residential tenure stochastically conditional on their neighborhood choice.

households' expected ex ante continuation values and their realized counterparts:

$$e^{\bar{V}}(x', \bar{\omega}_t^{kc}, \bar{\omega}_{t+1}^{kc}) \equiv \underbrace{\bar{V}(x', \bar{\omega}_{t+1}^{kc})}_{\text{realized}} - \mathbb{E}_{\bar{\omega}'|\bar{\omega}_t^{kc}}\left[\bar{V}(x', \bar{\omega}')\middle|\bar{\omega}_t^{kc}\right] \tag{8}$$

We follow Kalouptsidi, Scott, and Souza-Rodrigues (2021) and term the differences *expectational errors*. These expectational errors allow us to discard households' actual expectations in estimation. Having to solve for households' expectations would be prohibitively costly given the high-dimensional nature of a household's state space (some urban cores have over a thousand 2010-delineated census tracts).

Now that the function dependencies are clear, going forward, we suppress their arguments and remove the city superscripts unless they are required for explicative purposes: $\bar{V}_{xnt}^k \equiv \bar{V}^k\left(x_{it}, \bar{\omega}_t^{kc}\right)$, $V_{int}^k \equiv V^k(s_{it}^c)$, $\bar{u}_{xnt}^k \equiv \bar{u}_n^k\left(x_{it}, \bar{\omega}_t^{kc}\right)$, $v_{xnt}^k \equiv v_n^k\left(x_{it}, \bar{\omega}_t^{kc}\right)$, and $p_{xnt}^k \equiv p_n^k\left(x_{it}, \bar{\omega}_t^{kc}\right)$

5. Structural Estimation

We estimate our neighborhood demand parameters with what is termed in the dynamic discrete choice literature an Euler equations in conditional choice probabilities (ECCP) estimator.[26] Similarly to other conditional choice probability estimators, ECCP estimation involves two steps. In the first step, we estimate households' conditional choice probabilities, the soon-to-be-introduced transition distributions for the household-level state variables, and households' variable moving costs, $\boldsymbol{\beta}_d^{k\prime}$ and $\boldsymbol{\beta}_s^{k\prime}$. We then estimate the remaining model parameters in a second step, conditional on our first-step estimates. Estimation in the second step is based on moment restrictions implied by the dynamic optimization of households, which we derive and explain in Section 5.2. With flow utilities linear in the model's parameters, we can evaluate these moment restrictions in a standard linear generalized method of moments (GMM) framework.[27]

The ECCP estimator has many advantages in our setting. First, the ECCP estimator is computationally light. Since our analysis covers the residential history of low-income households for ten years in 50 large US metropolitan areas at the census tract level, traditional dynamic discrete choice estimation procedures that explicitly solve for households' value functions are

[26] ECCP estimators are called such since they can be viewed as discrete choice analogues to Euler equations in models with continuous choice variables (Aguirregabiria and Magesan 2013; Kalouptsidi, Scott, and Souza-Rodrigues 2021).

[27] ECCP estimation has been used in a variety of applied settings, from choices over agricultural land use (Scott 2013) and new technology adoption (Groote and Verboven 2019) to occupational choice (Traiberman 2019; Gendron-Carrier 2023) and, most relevant to our setting, residential neighborhoods (Diamond, McQuade, and Qian 2018; Davis et al. 2021; Almagro and Domínguez-Lino 2022). See Kalouptsidi, Scott, and Souza-Rodrigues (2021) for a comprehensive econometric treatment of linear regression techniques with ECCP estimators.

infeasible (e.g., Rust (1987)). Second, our focus on gentrification implies an inherently nonstationary environment, making modeling the evolution of neighborhood change conceptually and computationally challenging. As we demonstrate in the derivation of our moment restrictions, ECCP estimation requires complete specification of neither households' information sets nor, therefore, the evolution of our city-specific state variables. Third, by providing moment conditions that we can evaluate in a linear GMM framework, we can exploit our instrumental variables detailed in Section 6 to estimate our model parameters in a manner consistent with the recent literature on identification by means of Bartik shift-share instruments (Goldsmith-Pinkham, Sorkin, and Swift 2020; Borusyak, Hull, and Jaravel 2022).

5.1. Estimation Assumptions

To identify our neighborhood demand parameters, we must make the following set of assumptions:

(a) *State Transitions*: The state variables s^c_{it} evolve according to a controlled first-order Markov process with a transition distribution that factors as[28]

$$f\left(s^c_{it+1}|n_{it}, s^c_{it}\right) = f^x(x_{it+1}|n_{it}, x_{it}) \cdot f^{\bar{\omega}}\left(\bar{\omega}^{kc}_{t+1}|\bar{\omega}^{kc}_t\right) \cdot f^{\varepsilon}(\varepsilon_{int+1})$$

(b) *Utility Normalization*: The utility offered by the outside option in every city is normalized to α^{ck} for each time period:

$$\alpha^k_{OO^c} + \beta^k_w \ln(\bar{I}_{OO^c,t}) - \beta^k_r \ln(r^k_{OO^c,t}) + \beta^k_A \ln\left(\frac{Coll_{nt}}{Pop_{nt}}\right) + \tilde{\xi}^k_{OO^c t} = \alpha^{ck} \quad \forall\, t$$

(c) *Rational Expectations*: Households' expectations over the evolution of the CBSA-level state variables conditional on their information set $\mathcal{I}_{it}$ correspond to the conditional expectations of the true data generating process given $\mathcal{I}_{it}$:

$$\mathbb{E}\left[e^{\bar{V}}(x', \bar{\omega}^{kc}_t, \bar{\omega}^{kc}_{t+1})|\mathcal{I}_{it}\right] = 0$$

where $e^{\bar{V}}(x', \bar{\omega}^{kc}_t, \bar{\omega}^{kc}_{t+1})$ are the expectational errors defined in equation 8.

[28]The evolution of the individual-level state variables, x_{it}, are "controlled" in that their evolution is influenced by the household's choices. While the current setup ensures that the evolution of x_{it} is fully determined by the household's neighborhood choice, our empirical implementation assumes that τ_{it} evolves stochastically conditional on the household's choice to reduce the dimensionality of our problem. Whether τ_{it} evolves stochastically or not, our discussion surrounding households' choice sets and moving costs in Section 4.3 should make clear that x_{it+1} evolves independently of $\bar{\omega}^{kc}_t$ conditional on n_{it} and x_{it}.

An important implication of assumption (a) is that the market-level state variables $\bar{\omega}_t^{kc}$ are *perceived* as exogenous by individual households; a household cannot expect to individually affect the evolution of $\bar{\omega}_t^{kc}$ with its own neighborhood choice (cf. Assumption (1) in Kalouptsidi, Scott, and Souza-Rodrigues (2021)). Given that the typical 2010-delineated US census tract contains around 4,000 residents, we believe that this assumption is plausible.[29] Note that Assumption (a) does not require the observed and unobserved city-specific state variables to evolve independently of one another. We highlight this to foreshadow the econometric challenge we face when attempting to identify preferences over functions of ω_t^c.

Assumption (b) says that residents who choose to reside outside of their respective CBSA's urban core obtain a time-invariant and CBSA-specific mean utility. [30] This assumption normalizes each CBSA's neighborhood mean utilities to a constant and time-invariant level, which is necessary to compare welfare across households within CBSAs given that logit models identify only differences in mean utilities. The assumption moreover facilitates exposition and, because each α^{ck} is unobserved, highlights the incommensurability of expected welfare both across different k-types and across CBSAs.

Last, Assumption (c) says that, on average, households correctly anticipate the evolution of $\bar{\omega}_t^{kc}$. An important corollary of Assumption (c) is that the contents of households' information sets in time t are mean independent of their expectational errors at time t as well (cf. Lemma 1 in Kalouptsidi, Scott, and Souza-Rodrigues (2021)). The importance of this corollary will become clear in Section 5.3 when we discuss our choosing among our set of IVs to estimate households' preferences.

5.2. Deriving Our Estimating Equations

Our goal now is to take our model setup and show how one can derive estimating equations that are linear in households' demand parameters. With such estimating equations, we can estimate our model with standard linear GMM techniques. To derive these equations, however, we must first introduce the concept of renewal actions.

Renewal Actions. To derive our estimating equations, we make use of what are termed renewal actions in the dynamic discrete choice literature (Hotz and Miller 1993; Arcidiacono and Miller 2011). Renewal actions are actions, that when taken in period t, lead to the same distribution of

[29] While our estimating equations are implied by households' dynamic optimization, they are not informed by any equilibrium conditions, allowing us to remain agnostic over how households' individual choices influence the evolution of the CBSA-wide state variables.

[30] Recall that *all* residents of a given CBSA (i.e., including those outside the urban core) additionally receive a time-varying but neighborhood-invariant utility shock, α_t^k. The value of the outside option can therefore shift over time, albeit always in proportion to the mean utilities in the urban cores.

states at the beginning of period $t+1$, regardless of the household's state in period t. In our setting, simply moving to a new neighborhood is a renewal action; moving to a new neighborhood resets a household's residential tenure to 0 regardless of the household's origin neighborhood or its current residential tenure. Moreover, because the city-specific state variables, $\bar{\omega}^{kc}_{t+1}$, and unobserved idiosyncratic preference shocks, ε_{int}, are independent of the household's state in period t, *all* the remaining state variables are reset to a common value upon moving to a new neighborhood.[31]

We exploit such renewal actions when deriving our estimating equation. To see how, consider the residential choices of two hypothetical type-k households between periods $t-1$ and $t+1$. Assume that, in period $t-1$, these households reside in the same neighborhood, n_{t-1}, but not in period t (i.e., at least one household chooses a new neighborhood in period t). Further, assume that, in period $t+1$, both households move to the same neighborhood, n_{t+1}. Our estimation procedure involves relating the difference in the expected discounted utilities of the two households' neighborhood choices to the difference in the probability that these neighborhood choices are actually made. Critically, because moving to neighborhood n_{t+1} in period $t+1$ constitutes the same renewal action for both households, their state variables are reset to a common value, which in turn equalizes their continuation values. Differences in the expected discounted utilities associated with the two sets of neighborhood choices are thus a function only of households' flow utilities. Relating such differences in expected discounted utilities to households' choice probabilities in this way helps disentangle the observed variation in households' flow utilities from households' unobserved continuation values.

To ease exposition, moving forward, we term these consecutive residential location choices *residential paths*.

Our Estimating Equation. Consider the set of residential paths that we just described for type-k households but with an additional requirement that one of the households chooses the outside option in period t:

(a) In period $t-1$, both households reside in the same neighborhood, n_{t-1}.

(b) In period t, one household chooses neighborhood n, while the other household chooses n'. n' also happens to be the outside option. While it must be the case that $n' \neq n$, it may be that $n_{t-1} = n$ or $n_{t-1} = n'$.

(c) In period $t+1$, both households convene at $\tilde{n}$, where $\tilde{n} \neq n'$ and $\tilde{n} \neq n$.

[31]Note that renewal actions depend on our construction of households' neighborhood tenure in Section 4.3. This construction assumes that the length of households' prior residential tenures has no impact on the value to future residential tenure. While this is a strong assumption, it is necessary to keep the dimension of the household-level state space manageable.

Given these residential paths, we can derive equation 9, which is linear in households' preference parameters. It is this equation that we use to construct the moment conditions that identify households' preferences. We derive it by equating the difference in the expected discounted utilities associated with the two residential paths to the difference in the probability that households actually take these paths:

$$(9)\quad Y^k_{xnn'\tilde{n}t} = \tilde{\alpha}^k_n + \alpha^k_t + \beta^k_w \ln(\bar{I}_{OO^c,t}) - \beta^k_r \ln(r^k_{OO^c,t}) + \beta^k_A \ln\left(\frac{Coll_{nt}}{Pop_{nt}}\right) + \beta^k_\tau \tilde{\tau}_x - \widetilde{MC}^k_t + \nu^k_{xnn'\tilde{n}t}$$

where

$$\begin{aligned}
Y^k_{xnn'\tilde{n}t} &\equiv \ln\left(\frac{p^k_{xnt}}{p^k_{xn't}}\right) - \delta\left(\sum_{x'} \ln\left(p^k_{x\tilde{n}t}\right) f^x\left(x'|n, x_t, \bar{\omega}^{kc}_t\right) - \sum_{x'} \ln\left(p^k_{x\tilde{n}t}\right) f^x\left(x'|n', x_t, \bar{\omega}^{kc}_t\right)\right) \\
\tilde{\alpha}^k_n &\equiv \alpha^k_n - \alpha^{ck} \\
\tilde{\tau}_x &\equiv \sum_{x'} \ln(\tau_{xt}(x')) f^x\left(x'|n, x_t, \bar{\omega}^{kc}_t\right) - \sum_{x'} \ln(\tau_{xt}(x')) f^x\left(x'|n', x_t, \bar{\omega}^{kc}_t\right) \\
\widetilde{MC}^k_t &\equiv MC^k_t(n, n_{xt-1}) - MC^k_t(n', n_{xt-1}) - \delta\left(MC^k_t(\tilde{n}, n) - MC^k_t(\tilde{n}, n')\right) \\
\nu^k_{xnn'\tilde{n}t} &\equiv \tilde{e}(x_t, \bar{\omega}^{kc}_t, \bar{\omega}^{kc}_{t+1}) + \tilde{\xi}^k_{nt}
\end{aligned}$$

Since the full derivation of this estimating equation is becoming well known, we relegate it to Appendix D.1. With estimates of households' conditional choice probabilities, $\hat{p}^k$, and estimates of the household-level transition distributions, $\hat{f}^x$, we can estimate equation 9 using linear GMM. The next subsection details our two-step estimation procedure.

5.3. Two-Step Estimation Procedure

In the first step of the estimation procedure, we estimate i) transition probabilities for the household-level state variables, ii) households' conditional choice probabilities, and iii) households' variable moving cost parameters. With these estimates in hand, we can estimate equation 9 using linear GMM.

Household Transition Distributions. To keep the dimension of the household state space manageable, we follow the literature stemming from Rust (1987) and discretize our household-level residential tenure measure into two buckets.[32] Specifically, we aggregate tenure into two

[32]The remaining household-level state variable is the household's residential location in the previous year. This state variable evolves deterministically depending upon the residential path under consideration. We therefore do not need to specify any transition probability for this component of x_t.

buckets:

$$\bar{\tau} = \begin{cases} 1 & \text{if } \tau \leq 4 \\ 2 & \text{otherwise} \end{cases}$$

We assume that this aggregated location tenure variable evolves stochastically according to the following distribution function[33]:

$$f^{\bar{\tau}}(\bar{\tau}_t = 1 | n_t, x(n_{t-1}, \bar{\tau}_{t-1}), \bar{\omega}^{kc}) = 1 \text{ if } n_t \neq n_{t-1}$$
$$f^{\bar{\tau}}(\bar{\tau}_t = 2 | n_t, x(n_{t-1}, \bar{\tau}_{t-1}), \bar{\omega}^{kc}) = \begin{cases} 1, & \text{if } n_t = n_{t-1} \text{ and } \bar{\tau}_{t-1} = 2 \\ g^k_{n_{t-1}}, & \text{if } n_t = n_{t-1}, \bar{\tau}_{t-1} = 1, \text{and } i \in k \end{cases}$$

where we estimate $g^k_{n_{t-1}}$ directly from the data:

$$\hat{g}^k_{n_{t-1}} = \frac{\sum\limits_{i \in n_t, k} \mathbb{1}\{\tau_{xt} = 5\}}{\sum\limits_{i \in n_{t-1}, k} \mathbb{1}\{\tau_{xt-1} \leq 4\}}$$

Given that our analysis is at the census tract level, estimated in this way, $\hat{g}^k_{n_{t-1}}$ is a sparse measure of $g^k_{n_{t-1}}$. In practice, we therefore take a weighted average of $\hat{g}^k_{n_{t-1}}$ across census tracts in each county and each year.

Conditional Choice Probabilities. Researchers typically face a trade-off between sparsity and flexibility when estimating first-step conditional choice probabilities, $\hat{p}^k_{xnt}$. Our setting is no different. On the one hand, we may estimate $\hat{p}^k_{xnt}$ directly from the data by calculating the probability that a type-k household with state x_{it} moves to each neighborhood $n_t \in \mathcal{N}^c$.[34] While this approach does not impose any restrictions on the implied data generating process, it leads to very sparse estimates of $\hat{p}^k_{xnt}$ in our setting given the number of census tracts in our largest CBSAs. On the other hand, we may impose some structure on the implied data generating process to smooth $\hat{p}^k_{xnt}$. Given that our setting yields particularly sparse empirical choice probabilities, we choose this latter option.

We model the count of type-k households in neighborhood $n \in \mathcal{N}^c$ choosing neighborhood $n' \in \mathcal{N}^c$ between periods t and $t+1$ as being derived from a Poisson distribution.[35] We parameterize the mean of the Poisson distribution in a way that does not impose additional restrictions

[33]Our notation here corresponds to the marginal transition distribution of τ_{it}, with n_{it} taken as given. This approach follows Almagro and Domínguez-Lino (2022).

[34]We could similarly employ any nonparameteric method to compute $\hat{p}^k_{xnt}$ directly from the data.

[35]We choose to model the data as a Poisson distribution because of its ability to account for sparse data and its computational efficiency (Correia, Guimarães, and Zylkin 2020).

on the data generating process implied by our dynamic model.[36] Specifically, we estimate the following flexible Poisson regression separately for each type-k household:

$$\log\left(\mathbb{E}\left[\left|n_{t-1}^{k\bar{\tau}} \to n_t^{\prime k\bar{\tau}}\right|\right]\right) = \gamma_{n't}^k + \mu_{\bar{\tau}} \cdot \mathbb{1}\{n' = n_{t-1})\} + \gamma_{n't}^k \cdot \lambda_{\bar{\tau}} \cdot \mathbb{1}\{n' = n_{t-1})\} - MC_t^k(n', n_{t-1}) \tag{10}$$

where $\left|n_{t-1}^{k\bar{\tau}} \to n_t^{\prime k\bar{\tau}}\right|$ is the count of type-k households with aggregated tenure status $\bar{\tau}$ in neighborhood n that choose neighborhood n' between periods $t-1$ and t. $\gamma_{n't}^k$ is a fixed effect that captures the neighborhood- and period-specific component of utility associated with choosing neighborhood n' in period t, $\mu_{\bar{\tau}}$ is a fixed effect that captures the additional utility that residents obtain from staying in their origin neighborhood given their tenure status, $\bar{\tau}$, and $\lambda_{\bar{\tau}}$ is a parameter capturing how the additional value one obtains from staying in her origin neighborhood varies with neighborhood mean utilities. $\mathbb{1}\{n' = n_{t-1})\}$ is an indicator variable that equals 1 if the household stays in its origin tract, and $MC_t^k(n', n_{t-1})$ are the same moving costs described in Section 4.3. We use our estimates from the Poisson model to predict the probability a type-k household with aggregated tenure status $\bar{\tau}$ living in neighborhood n chooses neighborhood n' in each year: $\hat{p}_{xn't}^k$.[37] As expected, the predicted probabilities are strongly correlated with their empirical counterparts; the coefficients of correlation are 0.951 and 0.984 for Black and non-Black households, respectively.

Equation 10 additionally identifies our variable moving cost parameters, $\beta_d^{k\prime}$ and $\beta_s^{k\prime}$. Since the cost of moving to neighborhood n differs for each type-k household depending on its origin neighborhood, we can separately identify the parameters governing variable moving costs from γ_{nt}^k, $\mu_{\bar{\tau}}$, and $\lambda_{\bar{\tau}}$. How the cost of moving varies with the physical distance of the move, $\beta_d^{k\prime}$, is identified with variation in the distance that households move within their urban core, conditional on moving. Similar variation but with respect to social distance identifies the cost of moving to neighborhoods socially different to one's origin neighborhood, $\beta_s^{k\prime}$.[38] We report our variable moving cost estimates with the rest of our parameter estimates in Table 4.

[36]Note that the independence of households' neighborhood moves in any one period implied by the Poisson distribution is embedded in Assumption (a); f^x for household i is independent of all other households' actions. Appendix D.2 shows how our Poisson regression specification does not impose additional restrictions on the neighborhood choice problem of households in our dynamic model.

[37]Many census tracts lack type-k households or in-migrants in a given year. Consequently, we cannot compute their conditional choice probabilities and exclude them from type-k households' choice sets. These tracts tend to be suburban or affluent, and their removal is unlikely to affect our estimates.

[38]Note that, in this repeated cross-sectional framework, we are unable to separately identify the fixed cost of moving from the value of residential tenure. We must instead estimate households' fixed moving costs in a second step using equation 9. Note also that our assumption of a Poisson distribution in equation 10 does not impact the variable moving cost estimates that we obtain here. This is because of the isomorphism between the score of the Poisson distribution and of the conditional logit (represented in equation 6) for these continuous variables, yielding identical maximum likelihood estimation (MLE) estimates (Guimarães, Figueirdo, and Woodward 2003).

Step Two. Given our estimated variable moving cost parameters, our estimated conditional choice probabilities, and our estimated transition probabilities from step 1, we may now construct the empirical analogue of equation 9:

$$\hat{Y}^k_{xnn'\tilde{n}t} = \tilde{\alpha}^k_n + \alpha^k_t + \beta^k_w \ln(\bar{I}_{n,t}) - \beta^k_r \ln(r_{n,t}) + \beta^k_A \ln\left(\frac{Coll_{nt}}{Pop_{nt}}\right) + \beta^k_\tau \hat{\tilde{\tau}}_{xt} - \widetilde{\widehat{MC^F}}^k_t + \nu^k_{xnn'\tilde{n}t} \tag{11}$$

where

$$\begin{aligned}
\hat{Y}^k_{xnn'\tilde{n}t} &\equiv \ln\left(\frac{\hat{p}^k_{xnt}}{\hat{p}^k_{xn't}}\right) - \delta\Bigg(\sum_{x'} \ln\left(\hat{p}^k_{x\tilde{n}t}\right)\hat{f}^x\left(x'|n, x_t, \bar{\omega}^{kc}_t\right) \\
&\qquad - \sum_{x'} \ln\left(\hat{p}^k_{x\tilde{n}t}\right)\hat{f}^x\left(x'|n', x_t, \bar{\omega}^{kc}_t\right)\Bigg) + \widetilde{\widehat{MC^V}}^k_t \\
\hat{\tilde{\tau}}_{xt} &\equiv \sum_{x'} \ln(\bar{\tau}_{xt}(x'))\hat{f}^x\left(x'|n, x_t, \bar{\omega}^{kc}_t\right) - \sum_{x'} \ln(\bar{\tau}_{xt}(x'))\hat{f}^x\left(x'|n', x_t, \bar{\omega}^{kc}_t\right) \\
\widetilde{\widehat{MC^\upsilon}}^k_t &\equiv \widehat{MC^\upsilon}^k_t(n, n_{t-1}) - \widehat{MC^\upsilon}^k_t(n', n_{t-1}) - \delta\left(\widehat{MC^\upsilon}^k_t(\tilde{n}, n) - \widehat{MC^\upsilon}^k_t(\tilde{n}, n')\right) \\
\nu^k_{xnn'\tilde{n}t} &\equiv \tilde{e}(x_t, \bar{\omega}^{kc}_t, \bar{\omega}^{kc}_{t+1}) + \tilde{\xi}^k_{nt}
\end{aligned}$$

and where $\hat{\cdot}$ represents estimates from the first step. $\widetilde{\widehat{MC^\upsilon}}^k_t$ is the difference in either the fixed, $\upsilon = F$, or variable, $\upsilon = V$, portion of moving costs. We detail the difference in expectational errors, $\tilde{e}(x_t, \bar{\omega}^{kc}_t, \bar{\omega}^{kc}_{t+1})$, in Appendix D.1.

To be precise about identification, it is worth unpacking the error term, $\nu^k_{xnn'\tilde{n}t}$, in equation 11. $\nu^k_{xnn'\tilde{n}t}$ is comprised of both unobserved neighborhood-specific amenities, $\tilde{\xi}^{kc}_{nt}$, and expectational errors, $\tilde{e}(x_t, \bar{\omega}^{kc}_t, \bar{\omega}^{kc}_{t+1})$. We consider each of these in turn starting with the unobserved neighborhood-specific amenities, $\tilde{\xi}^{kc}_{nt}$. Since we place no restriction on the relationship between $\tilde{\xi}^{kc}_{nt}$ and the remaining time-varying observable neighborhood characteristics, ordinary least squares (OLS) estimates of 11 would be biased. Expected income, neighborhood-level housing costs, and the share of college graduates will invariably be correlated with many unobserved neighborhood-level factors such as proximity to natural amenities that we do not observe as econometricians.

To distinguish between preferences for observed versus unobserved neighborhood amenities, we start by differencing equation 11 using residential paths starting in 2017 (i.e., $t-1 = 2017$) and residential paths starting in 2010 (i.e., $t-1 = 2010$), obtaining,

$$\Delta\hat{Y}^k_{xnn'\tilde{n}} = \Delta\alpha^k_t + \beta^k_w \Delta\ln(\bar{I}_{n,t}) - \beta^k_r \Delta\ln(r_{n,t}) + \beta^k_A \Delta\ln\left(\frac{Coll_n}{Pop_n}\right) + \Delta\nu^k_{xnn'\tilde{n}} \tag{12}$$

where the Δs correspond to the difference in the associated variables between $t = 2011$ and $t = 2018$. Differencing equation 11 removes the time-invariant component of exogenous neighborhood amenities, $\tilde{\alpha}_n^k$, the measures of residential tenure, $\beta_\tau^k \tilde{\tau}_x$, and the time-invariant components of the moving costs variables, MC^k, MC^{kc}, and $\boldsymbol{\beta}_d^{k\prime} \boldsymbol{d}(n_t, n_{t-1})$.[39] Our main concern now is that *changes* in the observed components of households' flow utilities are correlated with *changes* in unobserved neighborhood amenities and household expectational errors. We must therefore construct neighborhood-level instruments, z_n, for our endogenous regressors that are orthogonal to both of these components:

$$
\begin{aligned}
(13) \qquad 0 &= \mathbb{E}\left[z_n \, \Delta\nu_{xnn'\tilde{n}}^k\right] \\
&= \mathbb{E}\left[z_n \, \left(\Delta\xi_n^k + \Delta\, \tilde{e}(x, \bar{\omega}_t^{kc}, \bar{\omega}_{t'}^{kc})\right)\right] \\
&= \mathbb{E}\left[z_n \, \left(\Delta\xi_n^k + \Delta\left(e^{\bar{V}}(n, x, \bar{\omega}_t^{kc}, \bar{\omega}_{t'}^{kc}) - e^{\bar{V}}(n', x, \bar{\omega}_t^{kc}, \bar{\omega}_{t'}^{kc})\right)\right)\right]
\end{aligned}
$$

where $e^{\bar{V}}(n, x, \bar{\omega}_t^{kc}, \bar{\omega}_{t'}^{kc})$ is the difference between the realized type-k ex ante continuation value and type-k households' expectations of these continuation values, integrated over the potential realizations of the household-level states:

$$
e^{\bar{V}}(n, x, \bar{\omega}_t^{kc}, \bar{\omega}_{t'}^{kc}) \equiv \sum_{x'} \left(\bar{V}(x', \bar{\omega}_{t'}^{kc}) - \mathbb{E}_{\bar{\omega}'|\bar{\omega}_t^{kc}}\left[\bar{V}(x', \bar{\omega}')|\bar{\omega}_t^{kc}\right]\right) f^x\left(x'|n, x_n \bar{\omega}_t^{kc}\right).
$$

In addition to being orthogonal to changes in unobserved neighborhood amenities, equation 13 shows that our instruments must also be mean independent of changes in households' expectational errors—the second component of $\nu_{xnn'\tilde{n}t}^k$. Recall Assumption (c), which states that households have rational expectations over the evolution of the model's state variables. A corollary of this assumption is that the contents of households' information sets at time t are mean independent of their expectational errors (Kalouptsidi, Scott, and Souza-Rodrigues 2021). Conversely, elements of households' future information sets that cannot be predicted from their period-t information sets will be correlated with their expectational errors. For this reason, our instruments must not predict future values of $\bar{\omega}_t^{kc}$ in a way that cannot simultaneously be predicted from the information in households' period-t information sets. To see why, consider an instrument that shocks the neighborhood-n elements of $\omega_{t'}^{kc}$ for any $t \leq 2017$.[40] Assume that this shock is uncorrelated with changes in unobserved neighborhood amenities but cannot

[39]We estimate the remaining time-invariant parameters in a final stage. Specifically, we estimate equation 11 conditional on the estimates from our differenced regressions and our first-step multinomial choice model. We assume that the effect of residential tenure and moving costs on the likelihood of different residential paths is uncorrelated with unobserved neighborhood amenities.

[40]If the instrument shocks elements of $\omega_{t'}^{kc}$ for $t' > 2017$, we must also consider how the instrument affects the difference in expectational errors over time. The current example is sufficient to show that z_n must be constructed from variation that can be predicted from households' 2010 information sets.

be predicted with the information in households' contemporaneous information sets, $\mathcal{I}_{it}$. If this instrument is relevant, it will be mechanically correlated with the realized values of households' time-t ex ante continuation values, $\bar{V}(x', \bar{\omega}_{t'}^{kc})$, but uncorrelated with households' time-t expectations, $\mathbb{E}_{\bar{\omega}'|\bar{\omega}_t^{kc}}\left[\bar{V}(x', \bar{\omega}')\middle|\bar{\omega}_t^{kc}\right]$, violating the exclusion restriction embodied in 13. For this reason, the instruments that we detail immediately below are designed to predict changes in the CBSA-level state variables through 2011–2018 using only variation that can be predicted from households' 2010 information sets.

6. Identification Strategy

We now present the IVs that we use to identify low-income renters' preferences for our three endogenous variables, β_w^k, β_r^k, and β_A^k. Our first set of instruments predicts changes in neighborhood-level job market access for our target population of low-income renters, helping to identify their preference for job market access, β_w^k. Our second set of instruments is similar but predicts changes in neighborhood-level job market access for college graduates. By predicting neighborhood demand among college graduates, these instruments help identify preferences for endogenous amenities, β_A^k. Our third set of instruments interacts the predicted changes in neighborhood-level job market access for college graduates with the intensity of neighborhoods' ex ante urban development.[41] Neighborhoods with an ex ante high share of urban development tend to fall on the inelastic segment of local housing supply curves (Baum-Snow and Han 2023). This positioning makes it more probable that demand shocks will lead to increased rents. These interaction terms are thus useful for identifying households' distaste for paying rent, β_r^k.

We find that, conditional on our controls, the job market access instruments predict changes in rents and job market access for low-income workers well. They are less predictive of changes in neighborhoods' college shares, however. To increase first-stage power, we therefore include a fourth set of instruments. These instruments use the proximity to other neighborhoods' shares of college graduates to help predict neighborhood demand among college graduates. These instruments are particularly helpful in identifying households' preferences for endogenous amenities, β_A^k. We now discuss each instrument in turn.

6.1. Job Market Access IV

Neighborhoods differ in their access to employment opportunities. Neighborhoods located near establishments with a high demand for skilled labor will be attractive to college graduates

[41] Specifically, we interact the changes in job market access with the share of land in the census tract that is covered by urban development in 2011. We obtain these measures from Baum-Snow and Han (2023), who, in turn, construct them using data from the National Land Cover Database.

because of shorter expected commute times among this group, all else equal. A similar argument holds for low-income households and neighborhoods near establishments employing these workers. Our job market access instruments predict changes in the desirability of neighborhoods based on changes to their expected commute times to employment opportunities. We construct these instruments in two steps. First, we define industry- and neighborhood-specific measures of job market access in a baseline year separately for college graduates and low-income households. Second, we interact these baseline measures with national industry employment trends to predict changes in job market access that are plausibly uncorrelated with underlying trends in neighborhoods' exogenous amenities.

Our job market access instruments are based on the instruments constructed in Baum-Snow, Hartley, and Lee (2019) and Baum-Snow and Han (2023). We build on these instruments by using our employer–employee linked (LEHD) and business establishment (LBD) data to construct more precise and granular measures of job market access.[42] Moreover, because our business data are disaggregated at the 6-digit NAICS level, we can appeal to the exogeneity of national industry shocks to identify our model parameters. This allows the location of business establishments to be correlated with unobserved neighborhood characteristics, $\xi^{k}_{n,t}$ (Borusyak, Hull, and Jaravel 2022). We discuss how we construct these instruments and threats to identification below.

Instrument Construction. Our neighborhood-level measures of job market access can be formally defined in terms of each neighborhood n's access to employment in industry d at time t:

$$JMA_{ndt} = \sum_{m \in \mathcal{N}^c \setminus n} e^{-\eta^c \tau_{nm}} l_{mdt} \tag{14}$$

where l_{mdt} is the number of jobs in workplace tract m and 6-digit NAICS industry d at time t, $\tau_{n,m}$ is the travel time between tracts n and m, and η^c is a spatial decay parameter governing the importance of faraway jobs relative to closer jobs in determining a tract's employment access. We derive equation 14 from households' workplace choice problem in Appendix D.3.

We obtain measures of $\tau_{n,m}$ for college graduates using reported commute destinations and times from our ACS data. However, since the number of neighborhood–pairs in each CBSA is large relative to the number of college graduates surveyed in the ACS, we follow Baum-Snow, Hartley, and Lee (2019) and estimate a simple forecasting model to predict $\tau_{n,m}$ for all neighborhood pairs in each CBSA. We obtain CBSA-specific measures of η^c by using our employer–employee linked data to estimate gravity equations derived from a workplace choice model à la Tsivanidis (2022). We detail both our forecasting model and gravity equations in

[42] See Chow et al. (2021) and Giroud and Rauh (2019) for details on the LBD and Abowd et al. (2009) for details on the LEHD.

Appendix D.3.[43]

We use these industry-specific measures of employment access to construct our job market access IVs:

$$\Delta\,\widetilde{JMA}_{n,t_0,t} = \sum_{d\in\mathcal{T}} \underbrace{\frac{JMA_{n,d,t_0}\theta^c_d}{\sum_{d'} JMA_{n,d',t_0}\theta^c_{d'}}}_{Share} \underbrace{\frac{L^c_{d,t} - L^c_{d,t_0}}{L^c_{d,t_0}}}_{Shift} \tag{15}$$

where $L^c_{d,t}$ is national employment in industry d less employment in industry d located in neighborhood n's CBSA. $\mathcal{T}$ denotes the set of tradable industries that we use to predict changes in job market access.[44] As we do not observe the educational level or race of workers in our establishment-level data, we scale our industry-level measures of job market access by the share of workers employed in each industry and in each state who have a college degree, θ^c_d. Remember that analogous IVs are constructed for low-income renter households.

For all our job market access instruments, we select $t_0 = 2002$ and $t = 2007$. As our discussion around households' expectational errors in Section 5.3 highlights, our instruments may not use information outside of households' information sets to shift the endogenous variables. If the instruments predict changes in the endogenous variables that households do not—on average—expect, then the instruments will be correlated with their expectational errors. By setting $t_0 = 2002$ and $t = 2007$, we rely on serial correlation in the endogenous variables to ensure that our instruments remain relevant. Our choice of $t_0 = 2002$ is motivated by the fact that the US Census Bureau's Economic Census occurs on years ending in 2 and 7 (Chow et al. 2021). The allocation of firm employment data across establishments is most accurate in these years, increasing the precision of our baseline shares.[45] We find that first-stage power for our job market access instruments are then maximized when we choose $t = 2007$; local labor demand shocks induced by the great recession do not appear to influence households' within-CBSA location choices throughout 2011–2018.

Identifying Assumptions. Recent studies on shift-share instruments show how the exclusion restriction can hold with either conditionally exogenous shares (Goldsmith-Pinkham, Sorkin,

[43]The travel times and spatial decay parameters are defined separately for college graduates and low-income renters.

[44]We define our set of tradable industries using trade costs for 6-digit NAICS manufacturing and service industries, as estimated in Gervais and Jensen (2019). We label an industry as tradable if its estimated trade costs are in the bottom three quartiles of the manufacturing and service industries analyzed by Gervais and Jensen (2019). We find that this threshold ensures sufficient first-stage power while excluding industries whose establishment locations are likely endogenous to the spatial sorting within cities, such as local retailers. We discuss below our choice to focus on tradable industries.

[45]We considered using 1997 as our baseline year, but this would have required manually geocoding establishments' addresses since the LBD had not started reporting establishments' census tracts at that time.

and Swift 2020) or with conditionally exogenous shifts (Borusyak, Hull, and Jaravel 2022). The shares in equation 15 correspond to neighborhoods' baseline commute time–discounted exposure to employment in each industry. The shifts correspond to the national employment growth rate in each industry. As it is less plausible that establishments' baseline neighborhood locations are unrelated to underlying trends in nearby neighborhoods' exogenous amenities, we argue that identification comes from the conditional exogeneity of national industry employment shifts.

Borusyak, Hull, and Jaravel (2022) show that three conditions are together sufficient to ensure that our employment "shifts" are conditionally exogenous. First, establishments in industries with nationwide employment shocks (positive or negative) must not be concentrated near neighborhoods experiencing trends (positive or negative) in their unobserved exogenous amenities. Second, no small subset of industries may comprise a large portion of the baseline shares. Third, industries' national employment shifts must be mutually uncorrelated given trends in unobserved amenities and baseline shares.[46]

We consider a number of threats to identification. First, researchers have argued that, throughout our analysis period, there was a general trend toward suburbanization among low-income households irrespective of gentrification (Bartik and Mast 2023). If establishments concentrated near suburban neighborhoods that primarily employed low-skilled workers were overrepresented in industries experiencing negative nationwide employment shocks, then we may mistake a secular migration trend for a distaste for market access. To account for this possibility, we residualize $\Delta\,\widetilde{JMA}_{n,t_0,t}$ on measures of proximity to the metro division's CBD. These measures include a quadratic in the physical distance between the CBD and neighborhood n's centroid, a quadratic in the population-weighted distance between the CBD and neighborhood n's centroid, and fixed effects for five equally sized concentric rings centered on the CBD. The concentric rings are measured in population-weighted distance. Together, these measures ensure that our job market access instruments induce variation in the endogenous variables among neighborhoods that are equidistant from each metro division's CBD.

A second threat to identification is that changing consumer preferences may jointly influence households' location choices *and* industry employment trends. For example, changing preferences for different types of nontradable services may simultaneously influence employment in those industries and households' within-CBSA residential location decisions. To address this concern, we construct $\Delta\,\widetilde{JMA}_{n,t_0,t}$ using only employment shifts in *tradable* 6-digit NAICS

[46] The first condition can be represented formally as $\mathbb{E}[g_d|\{\bar{\Delta\nu}_d\}_d, \{s_d\}_d] = \mu$, the second condition as $\mathbb{E}\left[\sum_d s_d^2\right] \rightarrow 0$, and the third condition as $\text{Cov}\left(g_d, g_{d'}|\{\bar{\Delta\nu}_d\}_d, \{s_d\}_d\right) = 0\,\forall(d, d')$, where $s_d = \sum\limits_n s_{n,d} = \sum\limits_n \frac{JMA_{n,d,t_0}\theta^c_d}{\sum\limits_{d'} JMA_{n,d',t_0}\theta^c_{d'}}$, $g_d = \frac{L^c_{d,t} - L^c_{d,t_0}}{L^c_{d,t_0}}$, and $\bar{\Delta\nu}_d = \frac{\sum_{nn'\tilde{n}} s_{n,d}\Delta\nu_{nn'\tilde{n}}}{s_d}$.

industries.[47] This ensures that employment trends are not caused by households' residential location choices. We also construct our national employment shifters excluding employment in the CBSA for which we are predicting the changes in job market access.

A third threat to identification relates to the secular decline in manufacturing (NAICS 31–33) employment throughout 2002–2007 (Autor, Dorn, and Hanson 2013). Correlated employment shocks to 6-digit NAICS industries within the manufacturing sector threaten the consistency of our estimates. To account for correlated employment shocks within the manufacturing sector, we residualize $\Delta\,\widetilde{JMA}_{n,t_0,t}$ on neighborhoods' baseline exposure to manufacturing employment. Identification then requires that industry employment shifts *within* the manufacturing sector be mutually uncorrelated conditional on baseline shares and unobserved amenities, a significantly less stringent requirement (Borusyak, Hull, and Jaravel 2022).

We now move on to detailing our final set of instruments. These final instruments are particularly helpful in identifying households' preferences for observed neighborhood amenities, β_A^k.

6.2. Proximity IV

Guerrieri, Hartley, and Hurst (2013) show that during a positive city-wide employment demand shock, among low-income neighborhoods, it is those closest to other high-income neighborhoods that experience the greatest home price appreciation—a proxy for gentrification. Motivated by these findings, we construct a neighborhood-level and distance-weighted measure of proximity to other neighborhoods' share of college graduates. We then interact this measure with CBSA-wide Bartik labor demand shocks constructed from the initial CBSA-wide shares of college graduates in tradable 6-digit NAICS industries. Again, though our instrument construction is novel, we are not the first to use proximity to other high-income neighborhoods as an instrument for gentrification (Brummet and Reed 2021; Glaeser, Luca, and Moszkowski 2023).

Instrument Construction. We first form neighborhood-level distance-weighted measures of proximity to other neighborhoods' shares of college graduates in the same CBSA. We then interact these neighborhood-level measures with shocks to CBSA-wide tradable industry em-

[47] By restricting to tradable industries, the effective "shares" in $\Delta\,\widetilde{JMA}_{n,t_0,t}$ do not sum to 1 across industries and within neighborhoods. To control for the possibility that neighborhoods near high concentrations of tradable industry establishments are systematically exposed to increasing/decreasing exogenous neighborhood amenities, we additionally control for neighborhoods' exposure to the total share of tradable employment (Borusyak, Hull, and Jaravel 2022).

ployment:

$$\Delta\, \widetilde{\text{Prox}}_{n,t_0,t',t} = \sum_{d\in\mathcal{I}} \underbrace{\sum_{m\in\mathcal{N}^c\setminus n} e^{\rho\tau_{n,m}} \frac{Coll_{m,t_0}}{Pop_{m,t_0}}}_{Proximity} \cdot \underbrace{\frac{l_{d,t'}\cdot\theta^c_d}{\sum_{d'} l_{d',t'}\cdot\theta^c_{d'}}}_{Share} \underbrace{\frac{L^c_{d,t}-L^c_{d,t'}}{L^c_{d,t'}}}_{Shift}$$

where $\frac{Coll_{m,t_0}}{Pop_{m,t_0}}$ is the share of residents in neighborhood m who are college graduates and where $l_{d,t'} = \sum_n l_{d,n,t'}$. The variables $l_{d,n,t'}$, $L^c_{d,t}$ and the parameters $\theta^c_{d'}$ are defined as before. ρ is a spatial decay parameter governing the importance of faraway neighborhood college graduate shares relative to closer neighborhood shares. As $\rho \to \infty$, only the neighborhoods immediately adjacent to neighborhood n matter for determining the value of $\Delta\, \widetilde{\text{Prox}}_{n,t_0,t}$. Conversely, as $\rho \to 0$, the instrument loses all relevance, as every neighborhood matters equally in determining $\Delta\, \widetilde{\text{Prox}}_{n,t_0,t}$, ensuring that $\Delta\, \widetilde{\text{Prox}}_{n,t_0,t}$ is constant within each CBSA. Since we do not have a good prior for ρ, we calibrate its value using k-fold cross-validation in a set of first-stage regressions, regressing $\text{Gent}_{n,t_0,t}$ on $\Delta\, \widetilde{\text{Prox}}_{n,t_0,t}$ and selecting the value of ρ that best predicts the changes in neighborhood share of college graduates.

In contrast to local labor demand, industry-wide employment shocks throughout the great recession are very predictive of future *CBSA-wide* labor demand. For this reason, we use employment shifts between $t' = 2007$ and $t = 2010$ to construct our Bartik shift share, but construct the proximity shares in $t_0 = 2002$ for reasons that we discuss below.

Identification. Identification here proceeds differently from that under our job market access instruments. Because our final estimating equations include CBSA-wide fixed effects, the identifying variation must come from within CBSAs, and so identification from the CBSA-wide shift shares alone is ruled out. Identification instead proceeds from the interaction between the "proximity" term and the Bartik shift share.

Note that neighborhoods' proximity to other neighborhoods with a high share of college graduates is likely correlated with underlying trends in exogenous amenities. To account for this possibility, we take two steps. First, we use lagged shares of neighborhoods' proximity to other neighborhoods with a high share of college graduates. Second, we residualize $\Delta\, \widetilde{\text{Prox}}_{n,t_0,t',t}$ on the proximity terms themselves, so that identification comes solely from the *interaction* between the shift shares and the proximity terms. Identification thus proceeds analogously to that under a difference-in-difference estimator: we compare differences in gentrification between neighborhoods near and far from already gentrified tracts in CBSAs experiencing large labor demand shocks to differences in gentrification between neighborhoods near and far from already gentrified tracts in CBSAs not experiencing large labor demand shocks.

6.3. Moment Conditions

The final collection of instruments is:

$$z_{1,n} = \Delta\, \widetilde{JMA}^{Coll}_{n,02,07} \tag{16}$$

$$z_{2,n} = \Delta\, \widetilde{JMA}^{Low\text{-}Income}_{n,02,07} \tag{17}$$

$$z_{3,n} = \Delta\, \widetilde{JMA}^{Coll}_{n,02,07} \cdot \varphi_n, \tag{18}$$

$$z_{4,n} = \Delta\, \widetilde{\text{Prox}}_{n,02,07,10} \tag{19}$$

where φ_n are neighborhood-level measures of the share of land in the census tract covered by urban development in 2011. The superscripts *Coll* and *Low-Income* refer to the group for whom the instrument is constructed.[48] We estimate the moment restriction in equation 13 via linear GMM separately for each type-k household:

$$\mathbb{E}\left[z_n\, \Delta\nu^k_{xnn'\tilde{n}}\right] = 0$$

A corresponding observation in this moment restriction is a set of residential paths for a type-k household with individual state x. Each type-k household of city c has $2 \cdot (N^c + 1)$ initial states in period $t-1$, N^c possible neighborhood choices in that same period, and $N^c - 1$ possible neighborhood choices in period t, implying $\sum_c 2 \cdot (N^c + 1) \cdot N^c \cdot (N^c - 1)$ observations for for both Black and non-Black low-income renter households. Since the urban cores of our largest CBSAs contain approximately two thousand census tracts, the number of potential residential paths reaches into the tens of billions. To ease the computational burden, we restrict the set of residential paths that we analyze. Specifically, we restrict the set of residential paths to those that start from neighborhoods with the highest shares of type-k low-income renters in their CBSA. The cutoff for "highest shares" varies across CBSAs. For CBSAs with over 500 census tracts in their urban core, we select the top ten percent of tracts in terms of their share of type-k low-income renters. For CBSAs with under 500 census tracts in their urban core, we select the 50 highest tracts in terms of their share of type-k low-income renters. This choice ensures that a minimum number of census tracts from each CBSA informs our estimates. We report our estimates in Table 4.

[48] We estimate the commute elasticities and the employment share parameters for Black and non-Black low-income households jointly when constructing $\Delta\, \widetilde{JMA}^{Low\text{-}Income}_{n,02,07}$. This ensures that the same variation identifies both type-k households' preferences.

6.4. Parameter Estimates

The leftmost columns of Panel A in Table 4 report parameter estimates for households' valuations of neighborhood characteristics, while the same columns in Panel B report estimates on households' moving frictions. Since logit models identify only relative changes in welfare, the rightmost columns in Table 4 translate the estimates into households' willingness to pay measured in annual rents. The annual rent values for college share (β_A) and market access (β_w) reflect the extra annual rent payments that households would incur to reside in a neighborhood with a 10 percent higher share of college graduates or market access, respectively.[49]

The interpretation of the annual rent valuations differs across the moving friction parameter estimates. The annual rent values for physical move distance (β_d and β_d^{Sq}) represent the amount in annual rents that households would be willing to pay to move to a census tract one mile closer to their origin neighborhood, conditional on moving. The annual rent values for social distance (β_s and β_s^{Sq}) represent the amount in annual rents that households would be willing to pay to move to a census tract 10 percentage points more similar to their current census tract in terms of neighborhood college shares, again conditional on moving.[50] The fixed moving cost annual rent valuations represent how much in annual rents households would be willing to pay to avoid moving in any one year. The high residential tenure annual rent valuations represent how much the fixed cost of moving increases for each additional year that the household remains in its current neighborhood.

Our parameter estimates all have the expected sign, though Black households have a surprisingly low valuation for market access. One explanation for Black households' low valuation for market access is that the variation in market access induced by our job market access instruments is concentrated in industries primarily employing non-Black workers; such variation would not induce migration responses from households not employed in these industries, suggesting no preference for market access. Future iterations of the paper will test the robustness of our parameter estimates to skill- and race-specific measures of job market access.

Our parameter estimates reveal that it is moderately costly to move across neighborhoods. Low-income Black renter households' within-CBSA migration decisions suggest that the fixed cost of moving neighborhoods is over \$3,000 and that this cost increases by approximately \$600 for residents with a high amount of accumulated neighborhood capital ($\bar{\tau} = 2$). Surprisingly, the social distance between neighborhoods plays little role in households' residential migration

[49] We use the average neighborhood rent levels for Black and non-Black households in Table 2 as the basis of the percentage change in rents.

[50] The physical distance value of rent measures is calculated based on an initial move of 0 miles. The social distance value of rent measures is calculated assuming the average baseline share of college graduates for both Black and non-Black households as reported in Table 2.

decisions.[51] One explanation for this result is that low-income households initially residing in gentrifying tracts move to tracts with a lower share of college graduates, conditional on moving. Our model would interpret these moves as revealing a preference for social distance. A different model specification may incorporate asymmetric moving costs, capturing the distinct cost of moving into a neighborhood with a *higher* share of college graduates than one's origin tract.

7. Welfare Analysis

In this section, we examine what our parameter estimates imply about the welfare effects of gentrification for incumbent renters in the year 2000. Our welfare calculations are based on the following counterfactual analysis. Consider a representative low-income renter living in a gentrifying urban-core neighborhood, n, in the year 2000. To this renter, the distribution of neighborhood-level rents and shares of college graduates across her CBSA are exogenous. Given the exogenous distribution of these market-level state variables, $\bar{\omega}_t^k$, and her expectations over their evolution, we compute her expected welfare in 2000. Then, to estimate the welfare impact of gentrification, we compare this measure to the expected welfare that the same renter in neighborhood n would obtain in the year 2000 if the economy were instead in a steady state. If her expected welfare is lower in the steady-state equilibrium, we say that the changing distribution of neighborhood rents and college shares across the CBSA throughout 2000–2019 increased her welfare. The opposite is true if her expected welfare is higher in the steady-state equilibrium. We make these calculations separately for representative low-income renters in every low-income neighborhood in the year 2000.[52]

A steady-state equilibrium in this framework holds the distribution of market-level state variables, $\bar{\omega}_t^k$, fixed at their 2000 levels for every period t in the future. It further fixes the value of exogenous neighborhood amenities, $\{\xi_n\}_n$, at levels that induce a stationary distribution among low-income renter households. A stationary distribution implies that the share of each type-k household of tenure level $\bar{\tau}$ in each neighborhood remains constant over time. We define these concepts formally in Appendix D.4.

For every low-income neighborhood across each of our 50 large CBSAs, we make the follow-

[51]While Black households appear to value some social distance conditional on moving, they have a distaste for moves that involve significant differences in neighborhood educational composition.

[52]For all these welfare calculations, we keep the observed distribution of job market access at its 2000 levels. This choice is motivated by the reduced-form evidence in Section 3 showing that gentrification has negligible impacts on the employment outcomes of incumbent renter households. While job market access factors into households' location choices, gentrification does not appear to affect market access in a way that impacts incumbent low-income renter households. This could result, for example, from gentrification skewing the composition of nontradable employment opportunities toward higher-skilled workers even within 6-digit NAICS industries.

TABLE 4. Parameter Estimates

	Estimates		Value in Annual Rent	
	Black	Non-Black	Black	Non-Black
Panel A: Neighborhood Characteristics				
Rents (β_r)	-10.31***	-26.40***	–	–
	(1.494)	(3.341)		
College Share (β_A)	12.61***	7.343***	$1,224	$312
	(0.390)	(1.437)		
Market Access (β_w)	0.122	20.04***	$12	$851.4
	(1.635)	(3.487)		
Panel B: Moving Frictions				
Distance (β_d)	-0.2782***	-.3091***		
	(0.0004)	(0.0003)	-$270	-$130
Distance Squared (β_d^{Sq})	0.0039***	0.0044***		
	(0.0000)	(0.0000)		
Social Distance (β_s)	0.1335***	-0.1751***		
	(0.0108)	(0.0084)	$5	-$15
Social Distance Squared (β_s^{Sq})	-0.5598***	-0.4096***		
	(0.0074)	(0.0060)		
Within-CBSA Fixed Moving Cost (MC)	-4.558***	-4.318***	-$3,578	-$1,692
	(0.000)	(0.0000)		
Outside-CBSA Fixed Moving Cost (MC^c)	-4.458***	-4.621***	-$3,515	-$1,801
	(0.000)	(0.0000)		
High Residential Tenure ($\bar{\tau}$)	0.612***	0.551***	$612	$237
	(0.0002)	(0.0001)		
Controls				
Origin-Tract Fixed Effects	✓	✓		
Exposure to Tradable Industry JMA	✓	✓		
Exposure to Manufacturing JMA	✓	✓		
Proximity To CBD	✓	✓		
Fraction Tract Developed	✓	✓		
F-Stat	63.35	27.92		
N (1,000s)	2,573,000	10,630,000		

Notes: An observation is a set of residential paths described in Section D.1. We include fixed effects for the origin tract (i.e., n_{t-1}) of each residential path. Details on remaining controls are in the main text. Standard errors are in parentheses. The number of residential paths in our analysis differs across Black and non-Black low-income renter households because we predict fewer conditional choice probabilities for Black households in the first step of the analysis, as discussed in footnote 37. *Sources*: ACS (2005–2021), LEHD (2010–2012, 2017–2019), CoreLogic (2006–2017), MAF-X (2019), and MAF-ARF (2010–2012, 2017–2019). Estimates were disclosed by the US Census Bureau's Disclosure Review Board. Project Number 2358. Disclosure Clearance Number CBDRB-FY24-P2358-R10957.

ing calculation for a representative type-k household with tenure $\bar{\tau}$:

$$\Delta W^k(n, \bar{\tau}) = \frac{\bar{V}^k(x = (n, \bar{\tau}), \bar{\omega}^{kc}_{2000}) - V^k_{ss}(n, \bar{\tau})}{39\, \beta^k_r} \quad \forall\, k\, n \tag{20}$$

where $\bar{V}^k(x = (n, \bar{\tau}), \bar{\omega}^{kc}_{2000})$ is the expected welfare of a type-k incumbent renter living in neighborhood n with tenure status $\bar{\tau}$ in the year 2000. $V^k_{ss}(n, \bar{\tau})$ is the same calculation but computed under the assumption that the economy is in steady state. To make comparisons across type-k households, we normalize $\Delta W^k(n, \bar{\tau})$ by households' marginal utility of log rent. We can therefore interpret $\Delta W^k(n, \bar{\tau})$ as the welfare impact of neighborhood change starting in 2000 on incumbent renters initially living in neighborhood n. We provide further details on our welfare calculations in Appendix D.4.

$\Delta W^k(n, \bar{\tau})$ differs across neighborhoods to the extent that moving frictions mediate the welfare impacts of gentrification. If it were costless to move across neighborhoods and households did not accumulate neighborhood-level capital, households' individual states, $x(n_{i,t-1}, \tau_{i,t-1})$, would become irrelevant, and differences in welfare effects across neighborhoods within CBSAs disappear. This happens because, in each period, households simply choose to live in the neighborhood that offers the most ideal bundle of neighborhood characteristics regardless of where they are currently located. By contrast, with high moving costs and neighborhood capital accumulation, incumbent renters are averse to leaving their current neighborhood irrespective of changes in its characteristics. In this sense, incumbent renters bear the incidence of their changing neighborhood environment when there are substantial moving frictions.

Before we present our results, it is worth discussing two nuances surrounding our welfare analysis. First, when computing $\bar{V}^k(x = (n, \bar{\tau}), \bar{\omega}^{kc}_t)$, we hold the distribution of exogenous neighborhood amenities fixed at their 2000 levels. As a result, the distribution of low-income renters across neighborhoods predicted by our model does not match its observed counterpart after the year 2000. While fixing exogenous neighborhood amenities at their 2000 levels helps isolate the welfare effects of gentrification, the discrepancy between the model's predicted choice probabilities and their observed counterparts highlights a limitation of our single-agent framework—namely, that we cannot use our framework to assess counterfactual *policies* such as rent control, as these policies will affect the distribution of $\bar{\omega}^{kc}_t$ in ways that we cannot predict. All counterfactual analyses that we run therefore either directly manipulate households' preferences or elements in $\bar{\omega}^{kc}_t$.

Second, we do not have a good measure of *absolute* changes in amenities during our analysis period. Recall that we model observed neighborhood amenities as a function of the share of college graduates in each neighborhood. The share of college graduates increased nationwide from 26.8 percent to 36.5 percent between 2000 and 2017.[53] While changes in the distribution of college graduates across neighborhoods within a city likely reflects changes in the relative distribution of neighborhood amenities (Su 2022), it is unlikely that the absolute level of amenities increased in proportion to nationwide educational attainment. For this reason, our welfare analysis focuses on differences in welfare effects across renters living in different neighborhoods

[53]These are our own calculations based on publicly available IPUMS census data (Manson et al. 2022).

in the year 2000.

Minimal Variation in Welfare Effects Across Neighborhoods. Table 5 reports results from regressions of $\Delta W^k(n, \bar{\tau})$ on our measure of gentrification defined in Section 3.[54] The purpose of these regressions is to understand the welfare effects from gentrification for incumbent renters living in different neighborhoods within the same CBSA. The unit of analysis in these regressions is a 2010-delineated census tract. The sample includes the poorest 50 percent of all census tracts among the 50 largest urban cores in the US. These regressions include measures of neighborhoods' proximity to their CBSA's CBD. Neighborhoods closer to the periphery of the urban core have an artificially lower degree of access to other neighborhoods than do tracts closer to the CBD.[55] Tracts closer to the periphery of the urban core also experience less gentrification. The CBD proximity controls are designed to account for this correlation. The regressions also contain CBSA fixed effects and a set of baseline neighborhood controls that include baseline rents and college shares. These controls ensure that our regression results compare the impact of gentrification across neighborhoods within the same CBSA and across neighborhoods with the same baseline levels of rents and college shares. Our results are not sensitive to our excluding any of these controls.

The coefficients on gentrification in Table 5 reflect how gentrification impacts incumbent renters' welfare. The economically insignificant coefficients in Table 5 suggest that the economic impact of gentrification in one's current neighborhood is negligible. A 10-percentage-point increase in gentrification is associated with only a 0.0211 (0.00333) increase in log consumption units for Black (non-Black) renters with high neighborhood tenure (only \$17.69 (\$3.12) in annual rents). Results are similar across households with high and low levels of accumulated neighborhood capital.

There are at least three explanations for the null results reported in Table 5. First, it may be that the benefits of improving amenities exactly offset the costs of higher rents as incumbent renters stay in their home neighborhoods. Under this scenario, gentrification is welfare neutral regardless of one's home census tract. Second, gentrification may sometimes benefit incumbent renters and sometimes harm them, depending on the characteristics of their baseline neighborhood environment. In this scenario, the results in Table 5 mask underlying heterogeneity. Third, it may be that incumbent renters are sufficiently mobile to render changes in their current neighborhood relatively unimportant for expected utility. Here, it is not that gentrification is

[54] Here, we define gentrification with the end period, t, set to 2006. We found that gentrification defined over 2000–2006 explained the most variation in outcomes across a range of exploratory analyses.

[55] We define urban cores in Section 2. There are neighborhoods outside each urban core (and, hence, inside each CBSA's outside option) that residents of neighborhoods on the periphery of their urban core could in practice easily move to. Since, in our model, it is equally costly to move to the outside option from anywhere in the urban core, our setup artificially lowers the level of access to neighboring tracts for incumbent residents' in the urban periphery. Our CBD proximity controls attempt to account for these facts.

welfare neutral but that *where* incumbent renters initially live within their city is unimportant. We show that this third scenario drives our results.

TABLE 5. Welfare Effects of Gentrification on Incumbent Renters

$\Delta W^k(n, \bar{\tau})$	(1)	(2)	(3)	(4)
Gentrification	0.126***	-0.0032	0.2111***	0.0333
	(0.0200)	(0.0174)	(0.0303)	(0.0191)
	[1.613]	[1.249]	[1.591]	[1.225]
Controls				
CBD Proximity Controls	✓	✓	✓	✓
Baseline Neighborhood Controls	✓	✓	✓	✓
CBSA Fixed Effects	✓	✓	✓	✓
Sample Restrictions				
Race	Black	Non-Black	Black	Non-Black
Tenure Status	$\bar{\tau} = 1$	$\bar{\tau} = 1$	$\bar{\tau} = 2$	$\bar{\tau} = 2$
Number of Tracts	1,539	1,560	1,539	1,560

Notes: Table 5 reports results from regressions of $\Delta W^k(n, \bar{\tau})$ on $\text{Gent}_{n,2000,2006}$. Standard errors in parentheses are clustered by CBSA. The average change in log consumption across all neighborhoods is reported in square brackets. For computational reasons, we compute $\Delta W^k(n, \bar{\tau})$ for only five CBSAs: Boston, MA; Chicago, IL; Philadelphia, PA; Portland, OR; and San Francisco, CA. We will update Table 5 with the full sample of the largest 50 CBSAs very soon. We do not expect our results to change qualitatively. *Sources*: Publicly available ACS (2005–2021) and Decennial Census (2000) (Manson et al. 2022).

Changes in Neighborhood Choice Sets Govern Welfare Impacts. We show here that the null results in Table 5 are not because gentrification is welfare neutral or because they mask underlying heterogeneity. Rather, we show that gentrification's effects operate primarily through changing the characteristics of *other* neighborhoods in renters' choice sets. We show this by conducting the following counterfactual analysis. Consider a representative incumbent low-income renter residing in neighborhood n in the year 2000. Assume that neighborhood n's observable characteristics evolve just as observed in the data before. However, also assume that all other neighborhoods' characteristics remain fixed at their 2000 levels. The question is then: do the changing characteristics of neighborhood n benefit or harm the incumbent renter in neighborhood n? By holding the characteristics of other neighborhoods fixed, this counterfactual analysis separates the role of changing neighborhood characteristics in one's home neighborhood and the changing characteristics in other neighborhoods across the CBSA.

We perform this counterfactual analysis by constructing a modified measure of $\Delta W^k(n, \bar{\tau})$:

$$\Delta\widetilde{W}^k(n, \bar{\tau}) = \frac{\widetilde{V}^k(x = (n, \bar{\tau}), \bar{\omega}^{kc}_{2000}) - V^k_{ss}(n, \bar{\tau})}{\beta^k_r} \quad \forall\, k\, n \tag{21}$$

where $\widetilde{V}^k(x = (n, \bar{\tau}), \bar{\omega}^{kc}_{2000})$ is constructed exactly like $\bar{V}^k(x = (n, \bar{\tau}), \bar{\omega}^{kc}_{2000})$ except that the market-level state variables for $n' \neq n$ are fixed at their 2000 values for every year from 2000 onward.

Table 6 reports results from regressions of $\widetilde{V}^k(x = (n, \bar{\tau}), \bar{\omega}^{kc}_{2000})$ on our measure of gentrification. We observe economically and statistically significant positive effects of gentrification on incumbent renters' welfare. This is true for both Black and non-Black households (though the results are stronger for Black households), which have meaningfully different valuations of neighborhood characteristics, and for households with high and low levels of accumulated neighborhood capital. For example, a 10-percentage-point increase in gentrification increases Black (non-Black) renters' welfare by 1.1408 (.08023) log consumption units (\$1,776.66 (\$78.08) in annual rents) when these renters have a high level of accumulated neighborhood capital (Table 6, columns (3) and (4)).

Table 6 shows that, when we hold other neighborhoods' characteristics constant, gentrification meaningfully affects incumbent renters' welfare. Conversely, as we let other neighborhoods' characteristics vary, where an incumbent renter initially lives has little impact on her expected welfare. Tables 5 and 6 together suggest that how cities as a whole changed throughout 2000–2019 was more important for incumbent renters than how their own neighborhood alone changed throughout this time period.

Did Gentrification Benefit Incumbent Residents? The results reported in Table 5 suggest that gentrification benefited incumbent renters on average from the perspective of the year 2000. Despite negligible variation in welfare across tracts within CBSAs, the average tract-level increase in presented-discounted expected welfare over 2000–2019 for an incumbent Black (non-Black) resident with a low level of accumulated capital was \$40,230 (\$27,894).[56] There are a number of reasons to interpret these average effect sizes with caution, however. First, as mentioned above, we do not have a good measure of *absolute* changes in amenities during our analysis period. It is unlikely that the secular increase in nationwide educational attainment led to proportional increases in absolute levels of neighborhood amenities. Second, our model assumes that households can continuously adjust their housing consumption in response to rising rents. A model that assumes unit housing demand and therefore income effects may yield different average welfare effects in response to gentrification. Third, our analysis assumes

[56] These estimates are obtained based on the average increase in log consumption units reported in the square brackets in Table 5 and the average annual rents in Table 2.

TABLE 6. Mediation of Welfare Impacts by Neighborhood Choice Sets

$\Delta\widetilde{W}^k(n, \bar{\tau})$	(1)	(2)	(3)	(4)
Gentrification	11.9395***	0.5645**	11.4080***	.8023***
	(0.4420)	(0.2407)	(0.3969)	(0.1950)
	[0.9793]	[0.0398]	[0.9284]	[-0.0200]
Controls				
CBD Proximity Controls	✓	✓	✓	✓
Baseline Neighborhood Controls	✓	✓	✓	✓
CBSA Fixed Effects	✓	✓	✓	✓
Sample Restrictions				
Race	Black	Non-Black	Black	Non-Black
Tenure Status	$\bar{\tau} = 1$	$\bar{\tau} = 1$	$\bar{\tau} = 2$	$\bar{\tau} = 2$
Number of Tracts	1,539	1,560	1,539	1,560

Notes: Table 6 reports results from regressions of $\Delta\widetilde{W}^k(n, \bar{\tau})$ on $\text{Gent}_{n,2000,2006}$. The table is otherwise identical to Table 5.

constant preferences within each set of type-k households. Our results show that, with moderate moving costs, the welfare effects are governed by the changing neighborhood choice sets. For households with higher-than-average moving costs, however, gentrification may affect incumbent residents through changes in their initial neighborhood's characteristics. Future iterations of this paper will test the robustness of our results to these caveats.

8. Discussion

Beginning in the 1990s and intensifying after the year 2000, gentrification transformed the socioeconomic composition of vast areas within American inner cities. In this paper, we show how this transformation affected incumbent renters in these cities. We first show that gentrification affected neither the employment outcomes of renters living in gentrifying neighborhoods nor the length of time that incumbent renters stay in their home neighborhoods. We show that these results are robust to heterogeneity in neighborhoods' baseline environments, an a priori surprising finding. Gentrification did, however, impact the characteristics of the neighborhoods that incumbent renters lived in throughout our analysis period. This last finding suggests the possibility that gentrification meaningfully affected incumbent renters' welfare. To test whether and how gentrification affected incumbent renters' welfare, we estimate a dynamic model of residential and workplace choice. We use our parameter estimates from this model to quantify the welfare effects of gentrification. We show using our framework that gentrification affected incumbent renters primarily by changing the characteristics of *other* neighborhoods in renters' choice sets. We finally conclude with some caveats on interpreting our welfare estimates.

References

Abadie, Alberto, Susan Athey, Guido Imbens, and Jeffrey Wooldridge. 2023. "When should you adjust your standard errors for clustering?" *The Quarterly Journal of Economics* 138 (1): 1–35.

Abowd, John, Bryce Stephens, Lars Vilhuber, Fredrik Andersson, Kevin McKinney, Marc Roemer, and Simon Woodcock. 2009. "The LEHD Infrastructure Files and the Creation of the Quarterly Workforce Indicators." In *Producer Dynamics: New Evidence from Micro Data*, edited by Timothy Dunne, Bradford Jensen, and Mark Roberts, chap. 5, 149–230: University of Chicago Press.

Abramson, Boaz. 2023. "The Equilibrium Effects of Eviction and Homelessness Policies." *Working Paper*.

Aguirregabiria, Victor, and Arvind Magesan. 2013. "Euler Equations for the Dynamic Estimation of Discrete Choice Structural Models." *Structural Econometric Methods: Advances in Econometrics* 31: 3–44.

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.

Almagro, Milena, and Tomás Domínguez-Lino. 2022. "Location Sorting and Endogenous Amenities: Evidence from Amsterdam." *Working Paper*.

Angrist, Joshua D, and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.*: Princeton university press.

Arcidiacono, Peter, and Robert Miller. 2011. "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models with Unobserved Heterogeneity." *Econometrica* 79 (6): 1823–1867.

Autor, David H, David Dorn, and Gordon H Hanson. 2013. "The China syndrome: Local labor market effects of import competition in the United States." *American economic review* 103 (6): 2121–2168.

Autor, David H, Christopher J Palmer, and Parag A Pathak. 2017. "Gentrification and the amenity value of crime reductions: Evidence from rent deregulation."Technical report.

Balboni, Clare, Gharad Bryan, Melanie Morten, and Bilal Siddiqi. 2020. "Transportation, gentrification, and urban mobility: The inequality effects of place-based policies." *Working Paper*.

Bartik, Alexander, and Evan Mast. 2023. "Black Suburbanization: Causes and Consequences of a Transformation of American Cities." *Working Paper*.

Baum-Snow, Nathaniel, and Lu Han. 2023. "The Microgeography of Housing Supply." *Working Paper*.

Baum-Snow, Nathaniel, and Daniel Hartley. 2020. "Accounting for central neighborhood change, 1980-2010." *Journal of Urban Economics* 117: 103228.

Baum-Snow, Nathaniel, Daniel Hartley, and Kwan Ok Lee. 2019. "The Long-Run Effects of Neighborhood Change on Incumbent Families." *Working Paper*.

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, Robert McMillan, Alvin Murphy, and Christopher Timmins. 2016. "A Dynamic Model of Demand for Houses and Neighborhoods." *Econometrica* 84 (3): 893–942.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence Katz, and Christopher Palmer. 2023. "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice." *Working Paper*.

Berkes, Enrico, and Ruben Gaetani. 2023. "Income Segregation and the Rise of the Knowledge Economy." *American Economic Journal: Applied Economics* 15 (2): 69–102.

Berry, Christopher. 2021. "Reassessing the Property Tax." *Working Paper*.

Bishop, Kelly C. 2012. "A Dynamic Model of Location Choice and Hedonic Valuation." *Mimeo, Washington University in St. Louis*: 1–38.

Bishop, Kelly C, and Alvin D Murphy. 2019. "Valuing Time-Varying Attributes Using the Hedonic Model:

When is a Dynamic Approach Necessary?" *Review of Economics and Statistics* 101 (1): 134–145.
Böhlmark, Anders, and Alexander Willén. 2020. "Tipping and the Effects of Segregation." *American Economic Journal: Applied Economics* 12 (1): 318–47.
Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2022. "Quasi-Experimental Shift-Share Research Designs." *Review of Economic Studies* 89 (1): 181–213.
Boustan, Leah Platt. 2010. "Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration." *The Quarterly Journal of Economics* 125 (1): 417–443.
Brock, William A, and Steven N Durlauf. 2002. "A Multinomial-Choice Model of Neighborhood Effects." *American Economic Review, P&P* 92 (2): 298–303.
Brueckner, Jan. 1987. "The structure of urban equilibria: A unified treatment of the Muth-Mills model." *Handbook of Regional and Urban Economics* 2: 821–845.
Brueckner, Jan K, Jacques-François Thisse, and Yves Zenou. 1999. "Why is central Paris rich and downtown Detroit poor?: An amenity-based theory." *European Economic Review* 43 (1): 91–107.
Brummet, Quentin, and Davin Reed. 2021. "The Effects of Gentrification on Incumbent Residents." *Working Paper*.
Card, David, Alexandre Mas, and Jesse Rothstein. 2008. "Tipping and the Dynamics of Segregation." *Quarterly Journal of Economics* 123 (1): 177–218.
Chow, Melissa, Teresa Fort, Christopher Goetz, Nathan Goldschlag, James Lawrence, Elisabeth Perlman, Marther Stinson, and Kirk White. 2021. "Redesigning the Longitudinal Business Database." *NBER Working Paper* (28839).
Chyn, Eric, and Lawrence F Katz. 2021. "Neighborhoods Matter: Assessing the Evidence for Place Effects." *Journal of Economic Perspectives* 35 (4): 197–222.
Collinson, Robert, John Eric Humphries, Nicholas Mader, Davin Reed, Daniel Tannenbaum, and Winnie van Dijk. 2023. "Eviction and Poverty in American Cities." *Quarterly Journal of Economics*.
Correia, Sergio, Paulo Guimarães, and Tom Zylkin. 2020. "Fast Poisson estimation with high-dimensional fixed effects." *The Stata Journal* 20 (1): 95–115.
Couture, Victor, Cécile Gaubert, Jessie Handbury, and Erik Hurst. 2023. "Income Growth and the Distributional Effects of Urban Spatial Sorting." *The Review of Economic Studies*.
Couture, Victor, and Jessie Handbury. 2020. "Urban Revival in America." *Journal of Urban Economics* 119: 103267.
Couture, Victor, and Jessie Handbury. 2023. "Neighborhood Change, Gentrification, and the Urbanization of College Graduates." *Journal of Economic Perspectives* 37 (2): 29–52.
Cullen, Julie Berry, and Steven D Levitt. 1999. "Crime, Urban Flight, and the Consequences for Cities." *The Review of Economics and Statistics* 81 (2): 159–169.
Curci, Federico, and Federico Masera. 2023. "Flight from Urban Blight: Lead Poisoning, Crime, and Suburbanization." *The Review of Economics and Statistics*: 1–45.
Curci, Federico, and Hasin Yousef. 2022. "Gentrifying Cities, Amenities and Income Segregation: Evidence from San Francisco." *Working Paper*.
Davis, Morris, Jesse Gregory, and Daniel Hartley. 2023. "Preferences over the Racial Composition of Neighborhoods: Estimates and Implications." *Working Paper*.
Davis, Morris, Jesse Gregory, Daniel Hartley, and Kegon Tan. 2021. "Neighborhood Effects and Housing Vouchers." *Quantitative Economics* 12 (4): 1307–1346.
Dekle, Robert, Jonathan Eaton, and Samuel Kortum. 2008. "Global Rebalancing with Gravity: Measuring the Burden of Adjustment." *IMF Staff Papers* 55 (3): 511–540.
Derenoncourt, Ellora. 2022. "Can You Move to Opportunity? Evidence from the Great Migration." *American*

Economic Review 112 (2): 369–408.
Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review* 106 (3).
Diamond, Rebecca, and Cecile Gaubert. 2022. "Spatial Sorting and Inequality." *Annual Review of Economics* 14: 795–819.
Diamond, Rebecca, Timoth McQuade, and Franklin Qian. 2018. "The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco." *NBER Working Paper* (24181).
Ding, Lei, and Jackelyn Hwang. 2020. "Effects of Gentrification on Homeowners: Evidence from a Natural Experiment." *Regional Science and Urban Economics* 83: 103536.
Ding, Lei, Jackelyn Hwang, and Eileen Divringi. 2016. "Gentrification and Residential Mobility in Philadelphia." *Regional Science and Urban Economics* 61: 38–51.
Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A "Market Access" Approach." *The Quarterly Journal of Economics* 131 (2): 799–858.
Dragan, Kacie, Ingrid Gould Ellen, and Sherry Glied. 2020. "Does Gentrification Displace Poor Children and their Families? New evidence from Medicaid Data in New York City." *Regional Science and Urban Economics* 83: 103481.
Edlund, Lena, Cecilia Machado, and Maria Sviatschi. 2022. "Gentrification and the Rising Returns to Skill." *Economica* 89 (354): 258–292.
Ellen, Ingrid Gould, Keren Mertens Horn, and Davin Reed. 2019. "Has Falling Crime Invited Gentrification?" *Journal of Housing Economics* 46: 101636.
Ellen, Ingrid Gould, and Katherine O'Regan. 2010. "Crime and urban flight revisited: The effect of the 1990s drop in crime on cities." *Journal of Urban Economics* 68 (3): 247–259.
Ellen, Ingrid Gould, and Katherine M O'Regan. 2011. "How Low Income Neighborhoods Change: Entry, Exit, and Enhancement." *Regional Science and Urban Economics* 41 (2): 89–97.
Fee, Klye, and Daniel Hartley. 2013. "The Relationship Between City Center Density and Urban Growth or Decline." In *Revitalizing American Cities*, edited by Kimberly Zeuli and Susan Wachter, chap. 3, 45–64: University of Pennsylvania Press.
Ferreira, Fernando V, Jeanna H Kenney, and Benjamin Smith. 2023. "Household Mobility, Networks, and Gentrification of Minority Neighborhoods in the US."Technical report, National Bureau of Economic Research.
Freeman, Lance. 2005. "Displacement or Succession? Residential Mobility in Gentrifying Nneighborhoods." *Urban Affairs Review* 40 (4): 463–491.
Freeman, Lance, and Frank Braconi. 2004. "Gentrification and Displacement New York City in the 1990s." *Journal of the American planning association* 70 (1): 39–52.
Fu, Ellen. 2022. "The Financial Burdens of Property Taxes: Evidence from Philadelphia." *Working Paper*.
Gans, Herbert J. 1982. *Urban Villagers*.: Simon and Schuster.
Gendron-Carrier, Nicolas. 2023. "Prior Work Experience and Entrepreneurship: The Careers of Young Entrepreneurs." *IZA Discussion Paper* (16145).
Gervais, Antoine, and J. Bradford Jensen. 2019. "The tradability of services: Geographic concentration and trade costs." *Journal of International Economics* 118: 331–350.
Giroud, Xavier, and Joshua Rauh. 2019. "State Taxation and the Reallocation of Business Activity: Evidence from Establishment-Level Data." *Journal of Political Economy* 127 (3): 1262–1316.
Glaeser, Edward, Matthew Kahn, and Jordan Rappaport. 2008. "Why do the poor live in cities? The role of public transportation." *Journal of Urban Economics* 63: 1–24.
Glaeser, Edward, Jed Kolko, and Albert Saiz. 2001. "Consumer City." *Journal of Economic Geography* 1:

27–50.

Glaeser, Edward L, and Matthew E Kahn. 2001. "Decentralized Employment and the Transformation of the American City.": 1–63.

Glaeser, Edward L, and Erzo F P Luttmer. 2003. "The Misallocation of Housing Under Rent Control." *American Economic Review* 93 (4): 1027–1046.

Glaeser, Edward, Michael Luca, and Eric Moszkowski. 2023. "Gentrification and retail churn: Theory and evidence." *Regional Science and Urban Economics* 100: 103879.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review* 110 (8): 2586–2624.

Groote, Olivier De, and Frank Verboven. 2019. "Subsidies and Time Discounting in New Technology Adoption: Evidence from Solar Photovoltaic Systems." *American Economic Review* 109 (6): 2137–72.

Gu, Yizhen, Naijia Guo, Jing Wu, and Ben Zou. 2021. "Home Location Choices and the Gender Commute Gap." *Journal of Human Resources*.

Guerrieri, Veronica, Daniel Hartley, and Erik Hurst. 2013. "Endogenous Gentrification and Housing Price Dynamics." *Journal of Public Economics* 100: 45–60.

Guimarães, Paulo, Octávio Figueirdo, and Douglas Woodward. 2003. "A Tractable Approach to the Firm Location Decision Problem." *Review of Economics and Statistics* 85 (1): 201–204.

Hanson, Gordon H. 2005. "Market potential, increasing returns and geographic concentration." *Journal of International Economics* 67 (1): 1–24.

Harris, Chauncy D. 1954. "The Market as a Factor in the Localization of Industry in the United States." *Annals of the Association of American Geographers* 44 (4): 315–348.

Herzog, Ian. 2021. "National Transportation Networks, Market Access, and Regional Economic Growth." *Journal of Urban Economics* 122: 103316.

Hoelzlein, Mattias. 2023. "Two-Sided Sorting and Spatial Inequality in Cities." *Working Paper*.

Hotz, Joseph, and Robert Miller. 1993. "Conditional Choice Probabilities and the Estimation of Dynamic Models." *The Review of Economic Studies* 60 (3): 497–529.

Hwang, Jackelyn, and Jeffrey Lin. 2016. "What Have We Learned About the Causes of Recent Gentrification?" *Cityscape* 18 (3): 9–26.

Jackson, Kenneth T. 1987. *Crabgrass Frontier: The Suburbanization of the United States.*: Oxford University Press.

Jargowsky, Paul A. 2009. "Immigrants and Neighbourhoods of Concentrated Poverty: Assimilation or Stagnation?" *Journal of Ethnic and Migration Studies* 35 (7): 1129–1151.

Kain, John. 1962. "The journey-to-work as a determinant of residential location." *Papers of the Regional Science Association* 9 (1): 137–160.

Kain, John F. 1968. "Housing Segregation, Negro Employment, and Metropolitan Decentralization." *The Quarterly Journal of Economics* 82 (2): 175–197.

Kalouptsidi, Myrto, Paul Scott, and Eduardo Souza-Rodrigues. 2021. "Linear IV regression estimators for structural dynamic discrete choice models." *Journal of Econometrics* 222: 778–804.

Kennan, John, and James R Walker. 2011. "The Effect of Expected Income on Individual Migration Decisions." *Econometrica* 79 (1): 211–251.

Kleinbaum, David G, and Mitchel Klein. 1996. *Survival Analysis: A Self-Learning Text.*: Springer.

Kuminoff, Nicolai V, V Kerry Smith, and Christopher Timmins. 2013. "The New Economics of Equilibrium Sorting and Policy Evaluation Using Housing Markets." *Journal of Economic Literature* 51 (4): 1007–1062.

Le Barbanchon, Thomas, Roland Rathelot, and Alexandra Roulet. 2021. "Gender Differences in Job Search: Trading off Commute against Wage." *The Quarterly Journal of Economics* 136 (1): 381–426.

LeRoy, Stephen, and Jon Sonstelie. 1983. "Paradise lost and regained: Transportation innovation, income, and residential location." *Journal of Urban Economics* 13: 67–89.
Lester, T William, and Daniel A Hartley. 2014. "The Long Term Employment Impacts of Gentrification in the 1990s." *Regional Science and Urban Economics* 45: 80–89.
Manson, Steven, Jonathan Schroeder, David Van Riper, Tracy Kugler, and Steven Ruggles. 2022. "IPUMS National Historical Geographic Information System: Version 17.0 [dataset]. Minneapolis, MN: IPUMS.."
Meltzer, Rachel, and Pooya Ghorbani. 2017. "Does Gentrification Increase Employment Opportunities in Low-Income Neighborhoods?" *Regional Science and Urban Economics* 66: 52–73.
Mieszkowski, Peter, and Edwin S Mills. 1993. "The Causes of Metropolitan Suburbanization." *Journal of Economic Perspectives* 7 (3): 135–147.
Miller, Conrad. 2021. "When Work Moves: Job Suburbanization and Black Employment." *Working Paper*.
Moreno-Maldonado, Ana, and Clara Santamaria. 2022. "Delayed Childbearing and Urban Revival." *Working Paper*.
Moretti, Enrico. 2013. "Real Wage Inequality." *American Economic Journal: Applied Economics* 5 (1): 65–103.
Notowidigdo, Matthew J. 2020. "The Incidence of Local Labor Demand Shocks." *Journal of Labor Economics* 38 (3): 687–725.
Pennington, Kate. 2021. "Does Building New Housing Cause Displacement?: The Supply and Demand Effects of Construction in San Francisco." *Working Paper*.
Redding, Stephen J, and Esteban Rossi-Hansberg. 2017. "Quantitative spatial economics." *Annual Review of Economics* 9: 21–58.
Roback, Jennifer. 1982. "Wages, Rents, and the Quality of Life." *Journal of Political Economy* 90 (6): 1257–1278.
Rosen, Sherwin. 1974. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82 (1): 34–55.
Rust, John. 1987. "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher." *Econometrica* 55 (5): 999–1033.
Scott, Paul. 2013. "Dynamic Discrete Choice Estimation of Agricultural Land Use." *Discussion Paper, New York University*.
Silva, JMC Santos, and Silvana Tenreyro. 2006. "The Log of Gravity." *The Review of Economics and statistics* 88 (4): 641–658.
Su, Yichen. 2022. "The Rising Value of Time and the Origin of Urban Gentrification." *American Economic Journal: Economic Policy* 14 (1): 402–439.
Traiberman, Sharon. 2019. "Occupations and Import Competition: Evidence from Denmark." *American Economic Review* 109 (12): 4260–4301.
Tsivanidis, Nick. 2022. "Evaluating the Impact of Urban Transit Infrastructure: Evidence from Bogotá's TransMilenio." *Working Paper*.
Vigdor, Jacob L. 2002. "Does Gentrification Harm the Poor?" *Brookings-Wharton Papers on Urban Affairs*: 133–182.
Vigdor, Jacob L. 2010. "Is urban decay bad? Is urban revitalization bad too?" *Journal of Urban Economics* 68 (3): 277–289.
Wagner, Deborah, and Mary Layne. 2014. "The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software." *CARRA Working Paper Series*.
Wilson, William Julius. 1987. *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy.*: University of Chicago Press.

Appendix A. Aggregate Neighborhood Change Since 2000

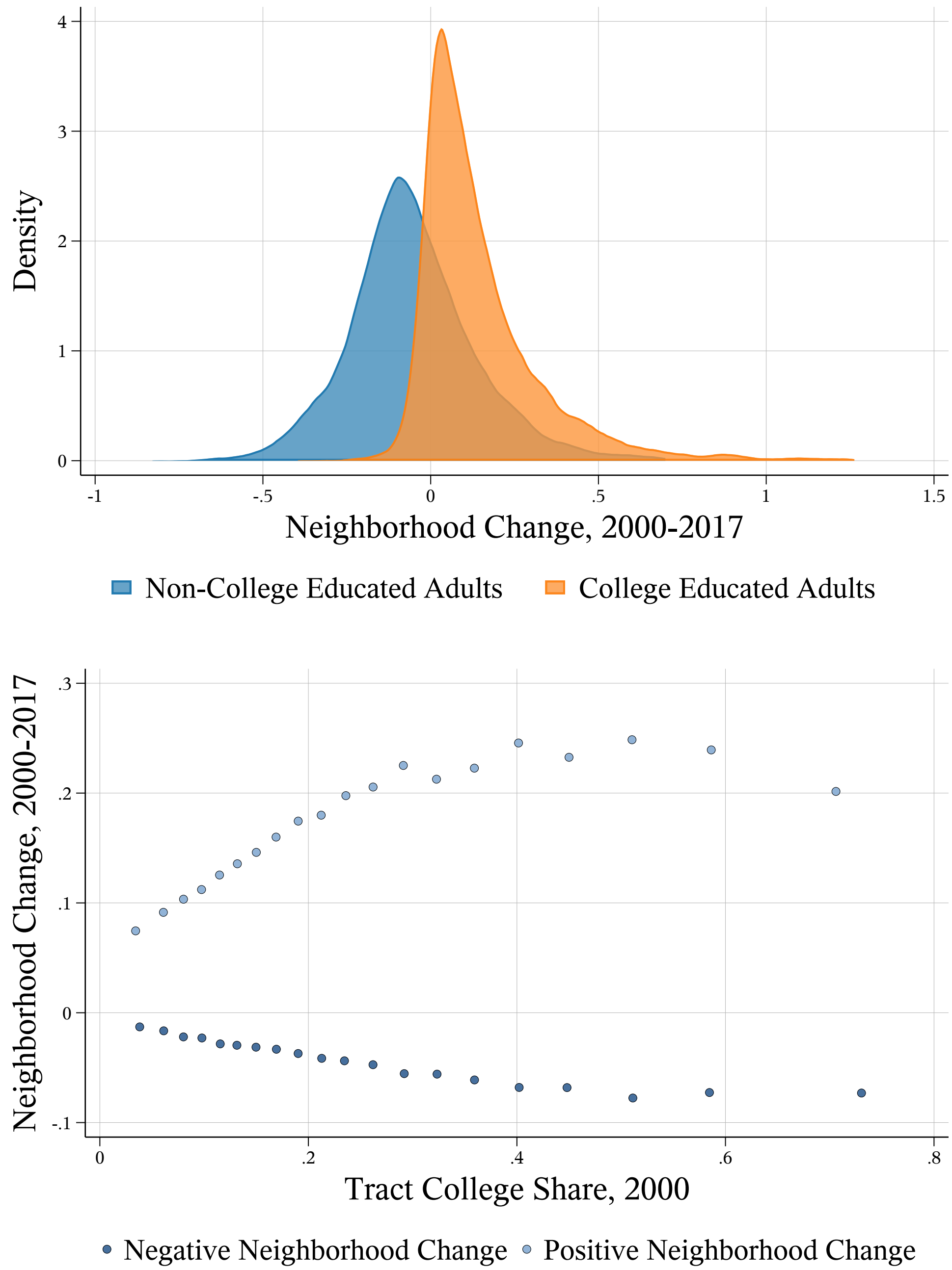

FIGURE A1. Neighborhood Change, 2000 - 2017

These figures are constructed using 2010 delineated census tract population counts for the 100 largest (ranked by population) US CBSAs from the publicly available 2000 NHGIS Census and the American Community Survey 2015-2019 5-year aggregates (Manson et al. 2022). We exclude tracts with fewer than 1,000 adult residents in 2000. Neighborhood change for non-college educated and for college-educated households is defined analogously to equation X. The top panel plots kernel densities of neighborhood change between 2000-2017 among Census tracts that contain the 20 percent of the CBSA's population that is closest to its CBD. The bottom panel plots our tract-level measure of neighborhood change (y-axis) against the tract's initial share of college-educated workers in 2000, separately for tracts that experienced positive and negative changes in their share of college-educated households.

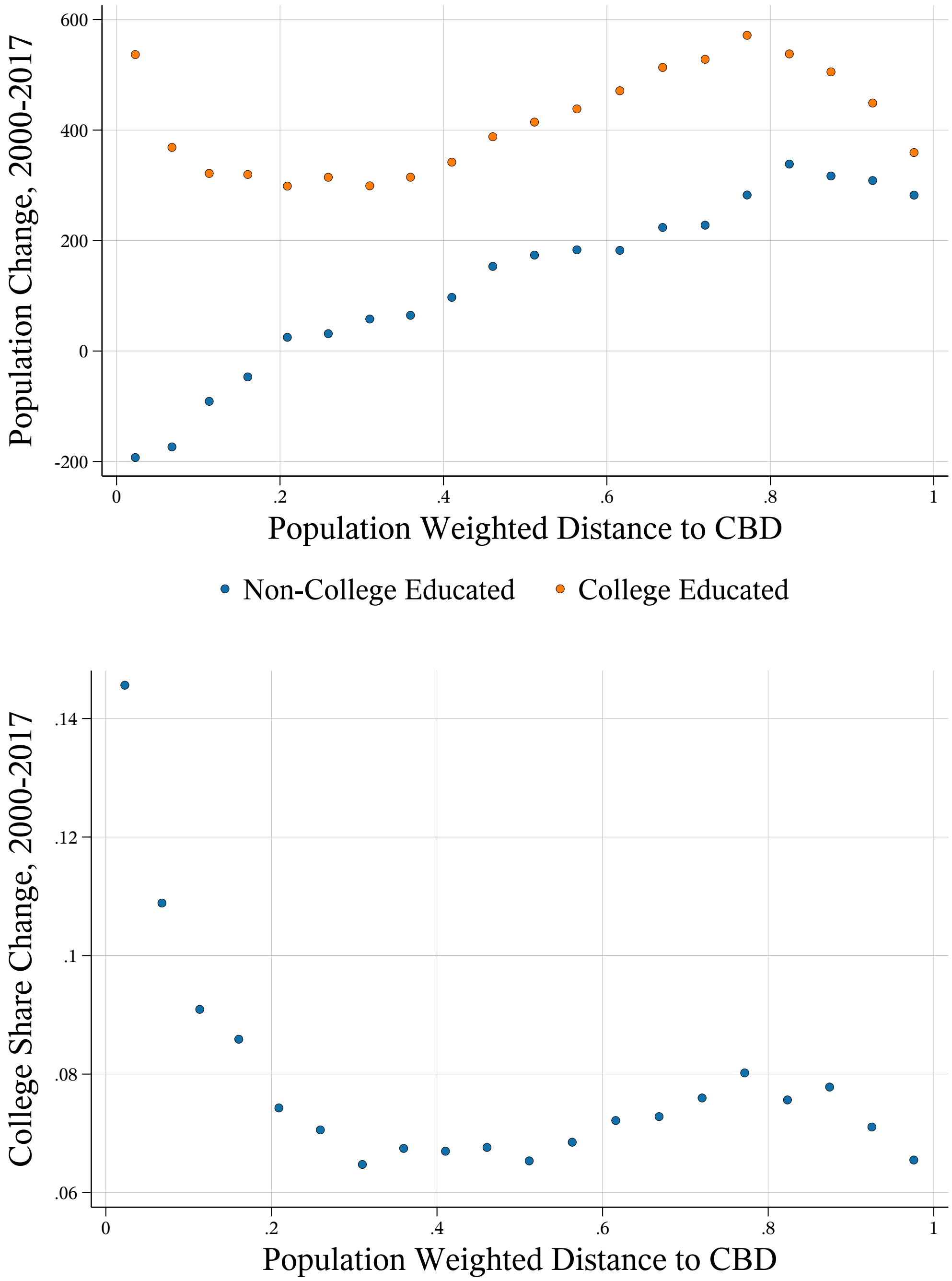

FIGURE A2. Neighborhood Change and Initial College Share

The figure is constructed using 2010 census tract population counts for the 100 largest (ranked by population size) US metropolitan divisions from the publicly available 2000 NHGIS Census and the American Community Survey 2015-2019 5-year aggregates (Manson et al. 2022). We exclude tracts with fewer than 1,000 adult residents in 2000. Neighborhood change is defined by equation X. The top panel plots the degree of population change for college and non-college educated households between 2000-2017 by the population-weighted distance to each metropolitan division's respective CBD. The bottom panel plots the change in the share of college educated households, also by the population-weighted distance to each metropolitan division's respective CBD.

Appendix B. Data Appendix

Coming very soon.

Appendix C. Reduced-Form Appendix

C.1. Control Variables

Household-Level Controls. Our vector of household-level controls, X_i, is designed to capture household characteristics that jointly determine one's origin location and our set of outcome variables, Δy_i and $h(t|i)$. These controls include i) average household income in defined as the mean income of all adults residing at the same address during 2010; ii) household size defined as the number of adults residing at the same address;[57] iii) a second-order polynomial in the household head's date of birth; iv) indicators for the household head's race, immigrant status, sex, and college degree attainment;[58] and v) the length of the household head's prior residential tenure in neighborhood $n(i)$. All variables are defined in our base year, 2010.

Neighborhood-Level Controls. Our vector of neighborhood-level controls, $X_{n(i)}$, is designed to capture characteristics of the household's origin neighborhood that are potentially correlated with our measure of neighborhood change and our set of outcome variables, Δy_i and $h(t|i)$. We first include neighborhood-level controls defined in 2010 which include i) the share of adults in the neighborhood with a college degree; ii) the share of adults in the neighborhood that identify as non-hispanic and white; iii) the median income among working age adults residing in the neighborhood; iv) the median property value and rent payment in the neighborhood; v) the degree of neighborhood churn which we define as the share of adults residing in the neighborhood during 2008 who also remain in the neighborhood during 2009; vi) second-order polynomials in the distance to the metropolitan division's CBD, measured in both the physical distance as well as in the cumulative share of the metropolitan division's residents residing closer to the CBD than those in tract $n(i)$; vii) fixed effects capturing five bands equidistant from the metropolitan division's CBD; and viii) the 5-year lag in our measure of neighborhood change.

In addition to our neighborhood-level controls defined in 2010, we also include a few contemporaneous neighborhood-level controls that are similarly designed to capture *changing* characteristics of the household's origin neighborhood that are potentially correlated with our measure of neighborhood change and our set of outcome variables, Δy_i and $h(t|i)$. In particular, we control for changes in job market access to *tradable* industries among low-skilled workers. In terms of equation 14, these changes in job market access among tradable industries for low-skilled workers are defined as $\Delta \sum_{d \in \mathcal{T}} JMA_{ndt} \tilde{\theta}_d^c$, where Δ corresponds to the difference in our *JMA* measure across 2010 and 2019, and $\tilde{\theta}_d^c$ is the share of non-college educated workers

[57] We topcode this value to 10, as a few addresses record an infeasibly high number of adult residents.

[58] While we run our analysis separately for Black and non-Black headed households, we include finer racial distinctions as part of our controls.

employed in industry d in CBSA c's state.[59] We finally include in $X_{n(i)}$ a third-order polynomial in the change in neighborhood $n(i)$'s total population.

Recall that all specifications additionally include CBSA-level fixed effects α_{CBSA}, ensuring that variation in gentrification comes from across neighborhoods in the same CBSA.

C.2. Identification

A causal interpretation of our coefficients of interest, β_{NC}^{Cox} and β_{NC}^{LP}, is based on the conditionally random assignment of neighborhood change across neighborhoods between 2010 and 2019. That is, conditional on our control variables, we assume there are no unobserved neighborhood- or household-level characteristics that are correlated with our measure of neighborhood change. Given our setup, to interpret β_{NC}^{Cox} and β_{NC}^{LP} causally, we much ensure two conditions are satisfied:

(a) First, residents in observably similar neighborhoods in 2010 must not differ in unobservable ways that correlate with our outcome variables, Δy_i and $h(t|i)$. If residents moving into gentrifying neighborhoods prior to 2010 are observably similar to their incumbent residents but different in unobservable ways that affect Δy_i and $h(t|i)$ (access to familial wealth, for example), β_{NC}^{Cox} and β_{NC}^{LP} will partly reflect sample selection.

(b) Second, neighborhoods experiencing gentrification throughout our analysis period must not be undergoing changes in their unobserved characteristics which independently predict incumbents' outcomes. While changes to unobserved public and private amenities caused by increased neighborhood demand among college-educated residents constitute part of our treatment, our identifying variation must be purged of shocks to neighborhood characteristics that affect incumbent residents' outcomes independently of gentrification. These unobserved shocks may include changes in tradable job market access (Kain 1962; Miller 2021), changes in transportation infrastructure that precedes gentrification (LeRoy and Sonstelie 1983; Glaeser, Kahn, and Rappaport 2008; Curci and Yousef 2022), or changes in neighborhood valuations resulting from secular trends in preferences or from shifts to within-city income distributions (Brueckner 1987; Couture et al. 2023). That suburbanization has continued unabated throughout our analysis periods—particularly for Black residents—makes these concerns especially acute (Bartik and Mast 2023; Couture and Handbury 2023).

We take steps to help ensure conditions (a) and (b) are met. First, in addition to our rich household-level controls, we often subset our sample to those who have lived in their origin neighborhood for at least 5 years. While this decision focuses our analysis on longtime renters,

[59] Note that including non-tradable industries into this measure of job market access would induce a classic "bad controls" problem, as changes in non-tradable employment can easily be conceptualized as an outcome of neighborhood change (Angrist and Pischke 2009).

it helps disambiguate the outcomes of gentrifiers and the outcomes of our target population of incumbent low-income renters, reducing the potential for sample bias. This is especially true when considering we control for the five-year lag in our measure of gentrification. Indeed, for sample selection to influence our results, it must be that low-income renter households predict with some accuracy how trends in neighborhood change will vary five years in the future and base their current residential choices on these predictions in a way that is uncorrelated with both their own observable characteristics (measured in 2010) and their chosen neighborhoods' observable characteristics (also measured in 2010).

Second, we are careful to control for changes in neighborhoods' unobserved characteristics that may independently predict incumbents' outcomes. By controlling for changes in access to low-skill tradable employment opportunities, we help ensure aggregate economic conditions that may independently affect neighborhood composition are not driving incumbents' outcomes. This concern is relevant given stark evidence of differential mobility responses to aggregate labor demand shocks across low- and high-skill workers (Notowidigdo 2020). By controlling for a cubic in neighborhood-level population change over our analysis period, we further control for unobserved neighborhood-level shocks that similarly affect incumbent residents and potential gentrifiers. To take an extreme example, consider a natural disaster during our analysis period that leads to a large depopulation of affected census tracts. In this scenario, we will observe low rates residential tenure among incumbent residents which we - without our population controls - would falsely attribute to a decline in our measure of neighborhood change. Note that the inclusion of our population controls implies that the coefficients on our measure of neighborhood change should be interpreted as the effect of changing neighborhood *composition* on incumbents' outcomes. Finally, our rich set of geographical controls ensure that we are comparing outcomes for incumbent residents across origin tracts that are equidistant from the Metropolitan Division's CBD, mitigating bias from secular trends toward suburbanization among our target population.

While we ensure our controls are carefully chosen to mitigate the impact of sample selection and changes in unobserved neighborhood-level characteristics, our estimates are robust to the exclusion of any small subset of control variables. It is finally worth noting that we explored using our IVs detailed in Section 6 to estimate our reduced-form equations. We find our reduced-form estimates are sensitive to the choice and composition (i.e. years and industries selected) of the instruments, indicating substantial heterogeneity in the complier characteristics of incumbent renters' origin neighborhoods across our instruments. It is therefore difficult to interpret the reduced-form IV estimates without placing more structure on our reduced-form equations to understand how each relevant equilibrium object (e.g. rents, non-tradable job market access, and neighborhood socioeconomic composition) mediates the impact of gentrification on incumbent renters. For now we report our OLS estimates which we believe offer a more transparent depiction of the impact of neighborhood change on incumbent renters' observable outcomes.

Cox Proportional Hazards Assumption. Identification in our Cox Proportional Hazards models further require that the proportional hazards assumption is met. Namely, that the impact of neighborhood change on incumbents' hazard rates are constant across each year between 2010 and 2019. We test this assumption by plotting $-\log(-\log(\text{survival probability}))$ against $\log(\text{time})$ separately for incumbent residents originally residing in neighborhoods within each decile of our measure of neighborhood change. We observe parallel lines across all deciles of neighborhood change, consistent with the proportional hazards assumption. We also test the null hypothesis that the corresponding Schoenfeld residuals for our measure of neighborhood change are not serially correlated.[60] We report the results of Stata's `phtest` command for our measure of neighborhood change in Table A1, which indicates that we cannot reject the null at the 95 percent confidence level for our full samples of Black and non-Black incumbent renters (though we are close to doing so for non-Black renters).

TABLE A1. Schoenfeld Residuals

Shoenfeld Residuals	Black	Non-Black
$\mathbb{P} \geq \chi^2$	0.317	0.056

Notes: Table A1 reports the results of Stata's `phtest`, testing the null hypothesis that the Schoenfeld residuals are not serially correlated. The test statistic is distributed as X^2 under the null hypothesis of no serial correlation. *Sources*: Estimates were disclosed by the US Census Bureau's Disclosure Review Board. Project Number 2358. Disclosure Clearance Number CBDRB-FY24-P2358-R10936.

[60]The Schoenfeld residuals correspond to the difference between the observed covariate values and the expected covariate values under the Cox proportional hazards model for incumbent renters at each year between 2010 and 2019. If the proportional hazards assumption holds, these residuals should not be serially correlated (Kleinbaum and Klein 1996).

Table A2. Cox Model: Effect of Gentrification on Incumbent Renters' Tenure (2010-2019)

Hazard Rate	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Neighborhood Change	-0.0688	-0.107*	-0.0793	0.187*	0.195	0.204*	0.369	0.323*	-0.777**	-0.264	0.165	-0.901*
	(0.0542)	(0.0439)	(0.132)	(0.0781)	(0.147)	(0.871)	(0.204)	(0.108)	(0.286)	(0.325)	(0.555)	(0.411)
Select Household-level Controls												
Household Income ($1,000s)	-0.00381***	-0.00341***	-0.00320***	-0.00277***	-0.00273***	-0.00275***	-0.00326***	-0.00257***	-0.00365***	-0.00300***	-0.00301***	-0.00476***
	(0.000155)	(0.000110)	(0.000413)	(0.000271)	(0.000666)	(0.000343)	(0.000858)	(0.000453)	(0.000526)	(0.000438)	(0.000887)	(0.000733)
Household Head DOB	0.0157***	0.0157***	0.0167***	0.143***	0.0157***	0.0145***	0.0151***	0.0152***	0.0175***	0.0144***	0.0189***	0.0130***
	(0.000216)	(0.000154)	(0.000581)	(0.000381)	(0.000948)	(0.000503)	(0.00124)	(0.000675)	(0.000734)	(0.000593)	(0.00120)	(0.000973)
Residential Tenure	-0.102***	-0.105***	-0.116***	-0.116***	-0.112***	-0.118***	-0.120***	-0.116***	-0.116***	-0.114***	-0.105***	-0.106***
	(0.000833)	(0.000555)	(0.00375)	(0.00230)	(0.00601)	(0.00311)	(0.00782)	(0.00411)	(0.00478)	(0.00343)	(0.00804)	(0.00596)
Select Neighborhood-level Controls												
Neighborhood Churn	-0.328***	-0.358***	-0.901***	-0.698***	-0.696***	-0.693***	-0.740***	-0.662***	-0.928***	-0.612***	-1.145***	-0.712***
	(0.0171)	(0.0117)	(0.102)	(0.0639)	(0.144)	(0.0825)	(0.202)	(0.112)	(0.146)	(0.101)	(0.232)	(0.179)
Rent ($1,000s)	0.272***	0.197***	0.0661	0.122***	0.135*	0.100***	0.146	0.100*	0.0749	0.198**	-0.0528	0.134
	(0.0209)	(0.0143)	(0.0452)	(0.0313)	(0.0617)	(0.0303)	(0.0844)	(0.0399)	(0.0664)	(0.0654)	(0.121)	0.0784
College Share	.141***	0.212***	0.235***	0.285***	-0.0866	0.143***	0.116	0.138*	0.923***	0.818***	0.704*	1.039***
	(0.0275)	(0.0178)	(0.0662)	(0.0350)	(0.108)	(0.0494)	(0.156)	(0.0676)	(0.207)	(0.163)	(0.352)	(0.252)
White Share	0.0975***	-0.113***	0.221***	-0.155***	0.215***	-0.0654	0.0348	-0.0666	0.245***	-0.195***	0.140*	-0.123
	(0.0150)	(0.0133)	(0.0321)	(0.0246)	(0.0578)	(0.0367)	(0.0807)	(0.0461)	(0.0383)	(0.0347)	(0.0586)	(0.0629)
Sample Restrictions												
Race	Black	Non-Black	Black	Non-Black	Black	Non-Black	Black	Non-Black	Black	Non-Black	Black	Non-Black
Longtime Renters			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Initial College Share					High	High	High	High	Low	Low	Low	Low
Fraction Developed							High	High			Low	Low
N (1,000s)	314	688	56	156	21	89	13.5	54.5	35	67	10.5	20.5

Notes: Coefficients correspond to the percent change in the hazard rate from a one unit increase in the corresponding independent variable. A one unit increase in our measure of neighborhood change corresponds to a one hundred percentage point increase. Every specification includes the full set of controls listed and detailed in Appendix C. Standard errors in parentheses are clustered at the origin census tract level. Longtime renters are renters who have resided in their origin Census tract since at least 2005. Tracts with a "High" ("Low") initial college share are tracts whose share of college-educated adults in 2010 is above (below) the population-weighted median in our full sample. Tracts with a "High" ("Low") fraction developed are tracts whose developed land area in 2011 is above (below) the population-weighted median in our full sample. *Sources*: ACS (2005-2021), LEHD (2010), CoreLogic (2006-2017), MAF-X (2019), and MAF-ARF (2010). Estimates were disclosed by the US Census Bureau's Disclosure Review Board. Project Number 2358. Disclosure Clearance Number CBDRB-FY24-P2358-R10936.

TABLE A3. Linear Regression Model: Effect of Gentrification on Incumbent Renters (2010-2019)

Outcome Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A: Household Outcomes												
Leave Tract	-0.0176	-0.0263	-0.016	0.07*	0.104	0.0668*	0.129	0.111***	-0.305***	-0.0888	-0.0458	-0.303
	(0.0223)	(0.0174)	(0.0509)	(0.0278)	(0.0571)	(0.0308)	(0.0738)	(0.0368)	(0.106)	(0.0894)	(0.254)	(0.158)
Moved > 5 Miles	0.038	0.0167	0.0066	0.0803**	0.0438	0.0709*	0.023	0.0926**	-0.0324	-0.0049	0.386	-0.205
	(0.0217)	(0.0171)	(0.0406)	(0.0261)	(0.0505)	(0.0294)	(0.0649)	(0.0339)	(0.0783)	(0.0802)	(0.223)	(0.154)
Leave CBSA	-0.0519***	-0.0375**	-0.0158	0.0087	-0.0194	0.0191	-0.0200	0.0179	0.0012	-0.0194	-0.0278	-0.1500
	(0.013)	(0.0122)	(0.0238)	(0.0160)	(0.0300)	(0.0178)	(0.0387)	(0.0205)	(0.0451)	(0.0432)	(0.1130)	(0.089)
Income	2,872**	2,268**	-932	-86.05	-1,458	-926.6	-1,875	311.5	-9,138*	-551.8	-23,950**	-3,874
	(948.8)	(736.4)	(2,170)	(1,375)	(2,748)	(1,547)	(3,498)	(1,816)	(3,854)	(3,427)	(9,225)	(6,444)
Commute Distance	-0.2280	-1.905***	0.637	0.5900	3.106	-0.234	3.651	1.0800	-1.854	3.146	-9.659	-3.379
	(0.776)	(0.531)	(1.609)	(1.002)	(2.059)	(1.146)	(2.529)	(1.352)	(3.132)	(2.775)	(6.284)	(5.352)
Panel B: Experienced Tract Characteristics												
Rent	0.271***	0.0960***	0.389***	0.222***	0.159**	0.132***	0.171*	0.153***	0.863***	0.607***	0.759**	0.485***
	(0.0472)	(0.0288)	(0.0620)	(0.0329)	(0.0546)	(0.0335)	(0.0701)	(0.0434)	(0.145)	(0.096)	(0.241)	(0.131)
College Share	0.726***	0.299***	1.079***	0.567***	0.539***	0.392***	0.570***	0.379***	3.726***	4.108***	3.559***	3.897***
	(0.104)	(0.0598)	(0.131)	(0.0788)	(0.0714)	(0.0525)	(0.0845)	(0.0689)	(0.250)	(0.236)	(0.565)	(0.357)
White College Share	-0.298	0.635***	0.277	1.269***	0.822*	0.449***	0.806	0.358***	7.758*	7.567***	2.832	5.206***
	(0.765)	(0.156)	(1.393)	(0.226)	(0.384)	(0.0863)	(0.419)	(0.098)	(3.951)	(1.167)	(7.965)	(1.169)
Sample Restrictions												
Race	Black	Non-Black	Black	Non-Black	Black	Non-Black	Black	Non-Black	Black	Non-Black	Black	Non-Black
Longtime Renters			✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Initial College Share					High	High	High	High	Low	Low	Low	Low
Fraction Developed							High	High			Low	Low
N (1,000s)	277	599	50	138	19	78.5	12	48.5	31	59.5	9.3	18

Notes: Coefficients correspond to the impact of a 100 percentage point increase in our measure of gentrification on a change in the associated outcome variable over 2010 to 2019. The dependent variables "Leave Tract", "Moved > 5 Miles", and "Leave CBSA" are all indicator variables equal to one if the corresponding condition is satisfied. Income and rents are measured in 2010 dollars and commute distance is measured in miles. Experienced neighborhood characteristics are measured in percent changes. Every specification includes the full set of controls listed and detailed in Appendix C. Standard errors in parentheses are clustered at the origin census tract level. Longtime renters are renters who have resided in their origin Census tract since at least 2005. Tracts with a "High" ("Low") initial college share are tracts whose share of college-educated adults in 2010 is above (below) the population-weighted median in our full sample. Tracts with a "High" ("Low") fraction developed are tracts whose developed land area in 2011 is above (below) the population-weighted median in our full sample. *Sources*: ACS (2005-2021), LEHD (2010), CoreLogic (2006-2017), MAF-X (2019), and MAF-ARF (2010). Estimates were disclosed by the US Census Bureau's Disclosure Review Board. Project Number 2358. Disclosure Clearance Number CBDRB-FY24-P2358-R10936.

Appendix D. Microfoundations and Structural Estimation Details

D.1. Deriving our Estimating Equation

Consider the set of residential choices detailed in Section 5.2. Given these residential choices, we start the derivation of our moment restrictions with an application of the Hotz-Miller inversion, which amounts to differencing equation 6 across the two neighborhood choices in period t, n and n':

$$\begin{aligned} \ln\left(\frac{p^k_{xnt}}{p^k_{xn't}}\right) &= v^k_{xnt} - v^k_{xn't} \\ &= \bar{u}^k_{xnt} - \bar{u}^k_{xn't} + \delta\left(\mathbb{E}\left[\bar{V}^k_{xnt}\right] - \mathbb{E}\left[\bar{V}^k_{xn't}\right]\right), \end{aligned} \tag{A1}$$

where the expectation operator is with respect to both the observable household-level and all the city-specific state variables. By assumption (a), we can write these expectations as

$$\begin{aligned} \mathbb{E}\left[\bar{V}^k_{xnt}\right] &= \sum_{x'} \int_{\bar{\omega}'} \bar{V}^k_{xnt} dF^{\bar{\omega}}\left(\bar{\omega}'|\bar{\omega}\right) f^x\left(x'|n, x_{nt}, \bar{\omega}^{kc}_t\right) \\ &= \sum_{x'} \mathbb{E}_{\bar{\omega}'|\bar{\omega}^{kc}_t}\left[\bar{V}(x', \bar{\omega}')|\bar{\omega}^{kc}_t\right] f^x\left(x'|n, x_{nt}, \bar{\omega}^{kc}_t\right) \end{aligned}$$

where x' and $\bar{\omega}'$ denote the next period values for x and $\bar{\omega}$. We can also replace the expectation of the ex-ante continuation values with respect to the city-specific state variables with their realized counterparts and an expectational error defined in equation 8:

$$\mathbb{E}\left[\bar{V}^k_{xnt}\right] = \sum_{x'} \bar{V}(x', \bar{\omega}^{kc}_{t+1}) f^x\left(x'|n, x_{nt}, \bar{\omega}^{kc}_t\right) + e^{\bar{V}}(n, x_{nt}, \bar{\omega}^{kc}_t, \bar{\omega}^{kc}_{t+1}), \tag{A2}$$

with

$$e^{\bar{V}}\left(n, x_t, \bar{\omega}^{kc}_t, \bar{\omega}^{kc}_{t+1}\right) \equiv \sum_{x'} e^{\bar{V}}\left(x', \bar{\omega}^{kc}_t, \bar{\omega}^{kc}_{t+1}\right) f^x\left(x'|n, x_{nt}, \bar{\omega}^{kc}_t\right).$$

The use of these realized continuation values permits minimal assumptions about households' beliefs over the evolution of the city-specific state variables. Imputing our expression for households' expected continuation values conditional on their own household-level state variables in A2 to our expression for the difference in conditional choice probabilities in A1 yields

$$\begin{aligned} \ln\left(\frac{p^k_{xnt}}{p^k_{xn't}}\right) &= \bar{u}^k_{xnt} - \bar{u}^k_{xn't} \\ &+ \delta\left(\sum_{x'} \bar{V}(x', \bar{\omega}^{kc}_{t+1}) f^x\left(x'|n, x_t, \bar{\omega}^{kc}_t\right) - \sum_{x'} \bar{V}(x', \bar{\omega}^{kc}_{t+1}) f^x\left(x'|n', x_t, \bar{\omega}^{kc}_t\right)\right) \end{aligned}$$

$$+ \delta \cdot \tilde{e}(x_t, \bar{\omega}_t^{kc}, \bar{\omega}_{t+1}^{kc})$$

where $\tilde{e}(x_t, \bar{\omega}_t^{kc}, \bar{\omega}_{t+1}^{kc})$ is the difference between the expectational errors when residing in neighborhood n relative to neighborhood n' in period $t+1$,

$$\tilde{e}(x_t, \bar{\omega}_t^{kc}, \bar{\omega}_{t+1}^{kc}) \equiv e^{\bar{V}}(n, x_t, \bar{\omega}_t^{kc}, \bar{\omega}_{t+1}^{kc}) - e^{\bar{V}}(n', x_t, \bar{\omega}_t^{kc}, \bar{\omega}_{t+1}^{kc}).$$

Next, using equation 7 to substitute in for $\bar{V}(x', \bar{\omega}_{t+1}^{kc})$, we obtain,

$$\begin{aligned}
&\ln\left(\frac{p_{xnt}^k}{p_{xn't}^k}\right) - \delta\left(\sum_{x'} \ln\left(p_{x\tilde{n}t}^k\right) f^x\left(x'|n, x_t, \bar{\omega}_t^{kc}\right) - \sum_{x'} \ln\left(p_{x\tilde{n}t}^k\right) f^x\left(x'|n', x_t, \bar{\omega}_t^{kc}\right)\right) \\
&= \bar{u}_{xnt}^k - \bar{u}_{xn't}^k + \delta\left(\sum_{x'} v_{\tilde{n}}^k\left(x', \bar{\omega}_t^{kc}\right) f^x\left(x'|n, x_t, \bar{\omega}_t^{kc}\right) - \sum_{x'} v_{\tilde{n}}^k\left(x', \bar{\omega}_t^{kc}\right) f^x\left(x'|n', x_t, \bar{\omega}_t^{kc}\right)\right) \\
&+ \delta \cdot \tilde{e}(x_t, \bar{\omega}_t^{kc}, \bar{\omega}_{t+1}^{kc})
\end{aligned}$$

Recall that $\tilde{n}$ is a renewal action for both households. Therefore, the household-level state variables are set to the same values for both households regardless of their values in period t. This yields identical continuation values in period $t+1$ for both households. The above expression therefore simplifies to,

$$\begin{aligned}
&\ln\left(\frac{p_{xnt}^k}{p_{xn't}^k}\right) - \delta\left(\sum_{x'} \ln\left(p_{x\tilde{n}t}^k\right) f^x\left(x'|n, x_t, \bar{\omega}_t^{kc}\right) - \sum_{x'} \ln\left(p_{x\tilde{n}t}^k\right) f^x\left(x'|n', x_t, \bar{\omega}_t^{kc}\right)\right) \\
&= \bar{u}_{xnt}^k - \bar{u}_{xn't}^k + \delta\left(MC_t^k(\tilde{n}, n) - MC_t^k(\tilde{n}, n') + \tilde{e}(x_t, \bar{\omega}_t^{kc}, \bar{\omega}_{t+1}^{kc})\right)
\end{aligned}$$

Choosing n' as the city's outside option, applying assumption (b), and substituting in for the neighborhoods' flow utilities provides an equation linear in our model parameters,

$$\begin{aligned}
\ln\left(\frac{p_{xnt}^k}{p_{xn't}^k}\right) &- \delta\left(\sum_{x'} \ln\left(p_{x\tilde{n}t}^k\right) f^x\left(x'|n, x_t, \bar{\omega}_t^{kc}\right) - \sum_{x'} \ln\left(p_{x\tilde{n}t}^k\right) f^x\left(x'|n', x_t, \bar{\omega}_t^{kc}\right)\right) \\
&= \tilde{\alpha}_n^k + \alpha_t^k + \beta_w^k \ln(\bar{I}_{n,t}) - \beta_r^k \ln(r_{n,t}) + \beta_A^k \ln\left(\frac{Coll_{nt}}{Pop_{nt}}\right) \\
&+ \beta_\tau^k\left(\sum_{x'} \ln(\tau_{xt}(x')) f^x\left(x'|n, x_{xt}, \bar{\omega}_t^{kc}\right) - \sum_{x'} \ln(\tau_{xt}(x')) f^x\left(x'|n', x_t, \bar{\omega}_t^{kc}\right)\right) \\
&- MC_t^k(n, n_{t-1}) + MC_t^k(n', n_{t-1}) + \delta\left(MC_t^k(\tilde{n}, n) - MC_t^k(\tilde{n}, n')\right) \\
&+ \tilde{e}(x_t, \bar{\omega}_t^{kc}, \bar{\omega}_{t+1}^{kc}) + \tilde{\xi}_{nt}^k,
\end{aligned}$$

where $\tilde{\alpha}_n^k = \alpha_n^k - \alpha^{ck}$. To condense notation, we write this equation as

$$Y_{xnn'\tilde{n}t}^k = \tilde{\alpha}_n^k + \alpha_t^k + \beta_w^k \ln(\bar{I}_{n,t}) - \beta_r^k \ln(r_{n,t}) + \beta_A^k \ln\left(\frac{Coll_{nt}}{Pop_{nt}}\right) + \beta_\tau^k \tilde{\tau}_x - \widetilde{MC}_t^k + \nu_{xnn'\tilde{n}t}^k,$$

where

$$\begin{aligned}
Y_{xnn'\tilde{n}t}^k &\equiv \ln\left(\frac{p_{xnt}^k}{p_{xn't}^k}\right) - \delta\left(\sum_{x'} \ln\left(p_{x\tilde{n}t}^k\right) f^x\left(x'|n, x_t, \bar{\omega}_t^{kc}\right) - \sum_{x'} \ln\left(p_{x\tilde{n}t}^k\right) f^x\left(x'|n', x_t, \bar{\omega}_t^{kc}\right)\right) \\
\tilde{\tau}_x &\equiv \sum_{x'} \ln(\tau_{xt}(x')) f^x\left(x'|n, x_t, \bar{\omega}_t^{kc}\right) - \sum_{x'} \ln(\tau_{xt}(x')) f^x\left(x'|n', x_t, \bar{\omega}_t^{kc}\right) \\
\widetilde{MC}_t^k &\equiv MC_t^k(n, n_{xt-1}) - MC_t^k(n', n_{xt-1}) - \delta\left(MC_t^k(\tilde{n}, n) - MC_t^k(\tilde{n}, n')\right) \\
\nu_{xnn'\tilde{n}t}^k &\equiv \tilde{e}(x_t, \bar{\omega}_t^{kc}, \bar{\omega}_{t+1}^{kc}) + \tilde{\xi}_{nt}^k.
\end{aligned}$$

This is the same equation reported in 9.

D.2. First-Step Conditional Choice Probabilities

To see how equation 10 approximates the neighborhood choice problem of households in our dynamic model, start by considering our specification for a household's conditional value function defined in 5,

$$\begin{aligned}
v_n^k\left(x_{it}, \bar{\omega}_t^{kc}\right) &\equiv \bar{u}_n^k\left(x_{it}, \bar{\omega}_t^{kc}\right) + \delta\mathbb{E}_t\left[\bar{V}^k\left(x_{it+1}, \bar{\omega}_{t+1}^{kc}\right)|n, x_{it}, \bar{\omega}_t^{kc}\right] \\
&= \alpha_n^k + \alpha_t^{kc} + \beta_w^k \ln\left(\bar{I}_{n,t}\right) - \beta_r^k \ln\left(r_{n,t}\right) + \beta_A^k \ln\left(\frac{Coll_{nt}}{Pop_{nt}}\right) - MC_t^k(n_t, n_{it-1}) + \tilde{\xi}_{nt}^k \\
&\quad + \sum_{\bar{\tau}=1,2}\left(\delta\mathbb{E}_t\left[\bar{V}^k\left(x_{it+1}(n_t, \bar{\tau}), \bar{\omega}_{t+1}^{kc}\right)|n, x(n_{t-1}, \bar{\tau}_{t-1}), \bar{\omega}_t^{kc}\right]\right. \\
&\quad \left. + \beta_\tau^k \ln(\bar{\tau})\right) \cdot f^{\bar{\tau}}(\bar{\tau}|n_t, x(n_{t-1}, \bar{\tau}_{t-1}), \bar{\omega}^{kc}).
\end{aligned} \tag{A3}$$

Assume for now that $\bar{\tau}_{t-1} = 1$. Then, we can re-write A3 for some neighborhood choice n as,

$$v_n^k\left(x_{it}, \bar{\omega}_t^{kc}\right) = \gamma_{nt}^k + \mu_{\bar{\tau}} \cdot \mathbb{1}\{n = n_{t-1})\} - MC_t^k(n, n_{t-1}) \tag{A4}$$

where,

$$\begin{aligned}
\gamma_{nt}^k = \alpha_n^k + \alpha_t^{kc} + \beta_w^k \ln\left(\bar{I}_{n,t}\right) - \beta_r^k \ln\left(r_{n,t}\right) + \beta_A^k \ln\left(\frac{Coll_{nt}}{Pop_{nt}}\right) + \tilde{\xi}_{nt}^k \\
+ \delta\,\mathbb{E}_t\left[\bar{V}^k\left(x_{it+1}(n_t, \bar{\tau} = 1), \bar{\omega}_{t+1}^{kc}\right)|n, x(n_{t-1}, 1), \bar{\omega}_t^{kc}\right]
\end{aligned}$$

$$\mu_{\bar\tau} = f^{\bar\tau}(2|n_t, x(n_{t-1},1), \bar\omega^{kc})\Bigg[\beta^k_\tau \ln(2) \\ - \delta \cdot \underbrace{\Bigg[\mathbb{E}_t\left[\bar V^k\left(x_{it+1}(n_t, \bar\tau=1), \bar\omega^{kc}_{t+1}\right)|n_t = n_{t-1}, x(n_{t-1},1), \bar\omega^{kc}_t\right] - \mathbb{E}_t\left[\bar V^k\left(x_{it+1}(n_t, \bar\tau=2), \bar\omega^{kc}_{t+1}\right)|n_t = n_{t-1}, x(n_{t-1},1), \bar\omega^{kc}_t\right]\Bigg]}_{\text{Difference in continuation values w.r.t. } \bar\tau_{t-1}}$$

and where $MC^k_t(n, n_{t-1})$ is defined as usual. Equation A4 shows that if the difference in conditional values across tenure status (τ_{t-1}) from staying in one's origin neighborhood is independent of one's origin neighborhood, n_{t-1}, then we could simply estimate our first-step multinomial choice model using A4.[61] It is, however, unlikely that that the difference in continuation values with respect to neighborhood tenure is independent of one's neighborhood origin. To understand why, consider two neighborhoods n and n' in the same CBSA. If neighborhood n provides lower utility than neighborhood n', then *all* residents of neighborhood n in period t are less likely to remain there in period $t+1$ relative to residents in neighborhood n'. When all residents are unlikely to stay in a given neighborhood, the difference in continuation values for incumbent residents with different residential tenures will be small. The converse is true for neighborhoods offering higher utilities to residents.

Equation 10 must therefore capture how the difference in continuation values across neighborhood tenures vary with neighborhoods' mean utilities. We therefore augment equation A4 by incorporating an interaction term between the value of residential tenure and time-varying neighborhood utilities:

$$v^k_n\left(x_{it}, \bar\omega^{kc}_t\right) = \gamma^k_{nt} + \mu_{\bar\tau} \cdot \mathbb{1}\{n = n_{t-1})\} + \gamma^k_{nt} + \lambda_{\bar\tau} \cdot \mathbb{1}\{n = n_{t-1})\} - MC^k_t(n, n_{t-1}).$$

This is the expression that appears in equation 10 and we believe approximates the data generating process implied by our dynamic model well.

D.3. Microfounding Job Market Access and Expected Income

Setup. The workplace choice problem for college graduates is identical to the problem for each type-k household, as reported in Section 4.2. In this section we therefore simply extend the index k to also include college graduates. For now, we assume that the labor market for each k-type is segmented.[62] Recall that in Section 4.2, each household's workplace choice problem is

[61] When $\bar\tau_{t-1} = 2$, the form of γ^k_{nt} remains the same, but $\mu_{\bar\tau}$ is no longer scaled by $f^{\bar\tau}(2|n_t, x(n_{t-1},1), \bar\omega^{kc})$. The current argument therefore remains unchanged when considering $\bar\tau_{t-1} = 2$.

[62] Baum-Snow, Hartley, and Lee (2019) show how to extend the model to an integrated labor market with multiple types. The resulting expressions are identical to those we derive here when assuming segmented markets. The

to choose which neighborhood to work in to maximize their income net of commute costs:

$$\bar{I}^k_{n,t} \equiv b^k_t \cdot \max_m \frac{z^i_{m,t}}{d_{n,m}} w_{mt},$$
$$= \max_m \frac{z^i_{m,t}}{d_{n,m}} w^k_{m,t},$$

To construct our JMA instrument, we simply amend this workplace choice problem by differentiating wages across industries, I, so that the household's conditional workplace choice problem becomes,

$$\bar{I}^k_{n,t} \equiv \max_{m,I} \frac{\zeta^k_{t,I} z^i_{m,t}}{d_{h,m}} w^k_{m,t},$$
$$= \max_{m,I} \frac{z^i_{m,t}}{d_{h,m}} w^k_{m,t,I},$$

where $z^i_{m,t,I}$ is distributed iid Frechet, $F(z^i_{m,t,I}) = \exp\left(-\left(z^i_{m,t,I}\right)^{-\epsilon^k_c}\right)$ with $\epsilon^k_c > 1$, for each workplace-tract-industry in the city and for each worker i of education group k. $\zeta^k_{t,I}$ is an industry- and type-specific productivity shock that captures the comparative advantage of workers of each type across industries. The probability of worker i of education group k living in tract n taking a job in tract m is then given by,

$$\pi^k_{m|n,t} = \frac{\sum_I (w^k_{m,t,I}/d^k_{n,m})^{\epsilon^k_c}}{\sum_I \sum_{m'} (w^k_{m',t,I}/d^k_{n,m'})^{\epsilon^k_c}},$$
$$= \frac{\sum_I (w^k_{m,t,I}/d^k_{n,m})^{\epsilon^k_c}}{RMA^k_{n,t}},$$

where $RMA^k_{n,t} \equiv \sum_I RMA^k_{n,t,I} \equiv \sum_I \sum_{m'} (w^k_{m',t,I}/d^k_{h,m'})^{\epsilon^k_c}$.

Define labor supply to tract m in time t by $\ell^k_{m,t} = \sum_I \left[\left(w^k_{m,I}\right)^{\epsilon^k_c}\right] FMA^k_{m,t}$, where $FMA^k_{m,t}$, represents the access firms in tract m have to k-type workers. Equating labor supply of k-type workers to tract m in period t to workers' choice probabilities yields an expression for FMA in terms of

derivations in this section follow an established literature microfounding measures of "market access" in economic fundamentals (Donaldson and Hornbeck 2016; Tsivanidis 2022; Baum-Snow and Han 2023).

RMA:

$$
\begin{aligned}
\ell^k_{m,t} &= \sum_n \pi^k_{m|n,t} \cdot \pi^k_{n,t} \cdot N^k \\
&= \sum_n \frac{\sum_I (w^k_{m,t,I}/d^k_{n,m})^{\epsilon^k_c}}{RMA^k_{n,t}} \cdot \pi^k_{n,t} \cdot N^k \\
&= \sum_I \left[\left(w^k_{m,t,I} \right)^{\epsilon^k_c} \right] N^k \sum_n \frac{\left(\left(\pi^k_{n,t} \right)^{1/\epsilon^k_c} / d^k_{h,m} \right)^{\epsilon^k_c}}{RMA^k_{n,t}} \\
&\equiv \sum_I \left[\left(w^k_{m,t,I} \right)^{\epsilon^k_c} \right] FMA^k_{m,t} \qquad \text{(A5)}
\end{aligned}
$$

The penultimate equality obtains,

$$
FMA^k_{m,t} = N^k \sum_n \frac{\left(\left(\pi^k_{n,t} \right)^{1/\epsilon^k_c} / d^k_{h,m} \right)^{\epsilon^k_c}}{RMA^k_{n,t}}.
$$

Furthermore, dividing both sides of A5 by $\left(d^k_{h,m} \right)^{\epsilon^k_c}$ and summing over m yields an expression for $RMA^k_{n,t}$ in terms of $FMA^k_{m,t}$. We subsequently obtain the following system of equations for *RMA* and *FMA*:

$$
\begin{aligned}
FMA^k_{m,t} &= N^k \sum_n \frac{e^{-\kappa \epsilon^k_c \tau^k_{n,m}} \pi^k_{n,t}}{RMA^k_{n,t}} \\
RMA^k_{n,t} &= \sum_m \frac{e^{-\kappa \epsilon^k_c \tau^k_{n,m}} \ell^k_{m,t}}{FMA^k_{m,t}}, \qquad \text{(A6)}
\end{aligned}
$$

where we have defined $d^k_{n,m} \equiv e^{\kappa \tau^k_{n,m}}$.[63]

Job Market Access Instrument. To transparently relate our job market access instruments to the recent advances in the quasi-experimental shift-share literature, we take linear approximations of equation A6 to obtain our measure of job market access that appears in expression

[63] $1 - e^{-\kappa \tau^k_{n,m}}$ represents the portion of time that type-k workers in tract n spend commuting to tract m.

14:

$$JMA_{ndt} = \sum_{m \in \mathcal{N}^c \backslash n} e^{-\eta^c \tau_{nm}} l_{mdt},$$

This linear approximation of equation A6 relates to earlier notions of market potential in international and regional trade theory, which conceptualize the demand for goods in a given region as the sum of demands in surrounding regions, weighted by bilateral transportation costs (Harris 1954; Hanson 2005). It is common in the regional economics literature to take linear approximations of structural measures of market access (e.g. Donaldson and Hornbeck (2016) and Herzog (2021)).[64]

Expected Income. Our workplace choice model also implies that expected income discounted by commuting costs for type-k households prior to drawing the vector of neighborhood- and period-specific productivity shocks is given by,

$$\bar{I}^k_{nt} = \Gamma\left(1 - \frac{1}{\epsilon^k_c}\right)\left(RMA^k_h\right)^{1/\epsilon^k_c}, \quad \forall i \in k,$$

where we directly use the identify $RMA^k_{n,t} \equiv \sum_I \sum_{m'} (w^k_{m',t,I}/d^k_{h,m'})^{\epsilon^k_c}$ in this measure's construction. This is the expression that enters into the households' flow utilities.

We use the observed distribution of workplace wages among our sample of low-income renter households to construct $RMA^k_{n,t}$. We then regress observed workplace wages discounted by commute costs on our measure of $RMA^k_{n,t}$ and predict $\bar{I}^k_{nt}$ solely using variation in $RMA_{n,t}$. This is to account for the fact that larger labor markets have mechanically higher levels of $RMA^k_{n,t}$.

Gravity and Forecasting Equations. Our workplace choice model yields gravity equations we can use to estimate $\kappa\epsilon^k_c$. We follow Baum-Snow, Hartley, and Lee (2019) and estimate $\kappa\epsilon^k_c$ separately for each type-k household in each city, c. Estimating $\kappa\epsilon^k_c$ separately for each type-k household in each CBSA increases the accuracy of our $RMA^k_{n,t}$ measures and thus also the power of our job market access instruments. CBSA-specific estimates allow labor demand shocks to impact job market access in neighborhoods accessible by longer commutes more in CBSAs where ϵ^k_c is lower. To obtain our estimates of $\eta^{kc} \equiv \kappa\epsilon^k_c$ we take the log of $\pi^k_{m|n,t}$ to obtain the

[64] An alternative approach to taking a linear approximation of A6 would be to solve for the fixed point of $RMA^k_{n,t}$ and $FMA^k_{m,t}$. However, it is unclear exactly how to relate the findings of Goldsmith-Pinkham, Sorkin, and Swift (2020) and Borusyak, Hull, and Jaravel (2022) on linear shift-share instrumentation to such a setting.

following gravity equation,

$$\begin{aligned} \ln(\pi^k_{m|n,t}) &= \ln(RMA^k_{n,t}) + \ln\left(\sum_I \left(w^k_{m,t,I}\right)^{\epsilon^k_c}\right) - \kappa\epsilon^k_c\tau^k_{n,m} \\ &= \alpha^k_{n,t} + \rho^k_{m,t} - \underbrace{(\kappa\epsilon^k_c)}_{\eta^{kc}}\tau^k_{n,m}, \end{aligned} \tag{A7}$$

which we estimate separately for each city and type-k households (including separately for college graduates) using 2010 commute flows constructed using the LEHD.

We obtain estimates of $\tau^k_{n,m}$ using the median tract commute time for each type-k worker in the 2005-2015 ACS surveys between tracts n and m for all tract pairs that are reported to have positive commute flows for any k-type worker. Since there are far more potential tract-to-tract commuting routes than ACS survey respondents, many commute times are not observed in our data. To estimate the remaining commute times, we follow Baum-Snow, Hartley, and Lee (2019) and construct an empirical forecasting model to predict commute times between all neighborhood pairs using the distance between neighborhood centroids and the corresponding city's CBD. We estimate the following forecasting model separately for each type-k worker,

$$\ln\tau^k_{n,m} = \alpha^k_d \ln \text{Distance}_{n,m} + \alpha^k_r \ln(\text{Home CBD Dis})_n + \alpha^k_w \ln(\text{Work CBD Dis})_m + \nu_c + u^k_{h,m,c}.$$

With these predicted commute times in hand, we use the observed tract-to-tract commute flows for each type-k household observed in the LEHD to estimate equation A7.[65] As per Census Bureau Disclosure Guidance, we are unable to release our CBSA-specific estimates, but report summary statistics of our CBSA-level estimates in Table A4. We also report our estimates using distance - as opposed to commute time - between commuting tract pairs as a comparison. Our estimates suggest that a 1 minute increase in commute time leads to a 8.8 or 8.9 percent reduction in the flow of Black and Non-Black commuters, respectively. These semi-elasticities, while on the higher side, are consistent with the magnitudes found in the existing literature (e.g. Ahlfeldt et al. (2015)).

[65]To maximize the size of the sample used to construct our commute flows, we do not restrict our sample to renter households.

TABLE A4. Commute Elasticities

	Time		Distance	
	Black	Non-Black	Black	Non-Black
Median CBSA Estimate	0.0876***	0.0885***	0.0807***	0.0782***
	(0.0000)	(0.0000)	(0.0001)	(0.0001)
Standard Deviation of CBSA Estimates	0.01353	0.01528	0.01667	0.07824
N (1,000s)	966	3,752	966	3,752

Notes: Table A4 reports results from the gravity equations of equation A7. We use Poisson Pseudo Maximum Likelihood to estimate these models given the prevalence of tract-to-tract potential commuting routes with zero observed commute flows (Silva and Tenreyro 2006). Observations are weighted by number of commuters in origin census tracts. To increase the number of non-zero commute flows we observe in the data, we estimate these gravity equations for all low-income households regardless of their rental status. Time is measured in minutes and distance is measured in miles. The first row reports the (pseudo) median CBSA-level estimate along with its accompanying standard error in parentheses. The second row reports the standard deviation of our 50 CBSA-level estimates. *N* reports the underlying number of commuters used to construct the commute flows in the LEHD. *Sources*: LEHD (2010). Estimates were disclosed by the US Census Bureau's Disclosure Review Board. Project Number 2358. Disclosure Clearance Number CBDRB-FY24-P2358-R10957.

D.4. Welfare Analysis

Recall that our baseline measure of change in expected welfare for incumbent renters is given by,

$$\Delta W^k(n, \bar{\tau}) = \frac{\bar{V}^k(x = (n, \bar{\tau}), \bar{\omega}^{kc}_{2000}) - V^k_{ss}(n, \bar{\tau})}{\beta^k_r} \quad \forall\, k, n.$$

We now unpack $\Delta W^k(n, \bar{\tau})$, starting with $V^k_{ss}(n, \bar{\tau})$. $V^k_{ss}(n, \bar{\tau})$ is the expected welfare for an incumbent renter with residential tenure $\bar{\tau}$ residing in neighborhood n during 2000 under the assumption that the economy is in steady state. We calculate $V^k_{ss}(n, \bar{\tau})$ by finding time-invariant continuation values and levels of exogenous amenities, $\{\xi^k_n\}_{n,k}$, that induce a *stationary distribution* among our sample of low-income renter households when holding all the observed state variables fixed at their 2000 levels, i.e. $\omega^{kc}_t = \omega^{kc}_{2000}\ \forall t$.

Stationary Distribution. Denote $q^k_n(\bar{\tau})$ as the model-implied share of type-k households with tenure $\bar{\tau}$ living in neighborhood n. Stacking $q^k_n(\bar{\tau})$ yields the $2(N^c + 1)$–dimensional vector $\boldsymbol{q}^k$. We define a stationary equilibrium among our sample of low-income renter households as $\boldsymbol{q}^k = \boldsymbol{\Lambda}^k \boldsymbol{q}^k$, where $\boldsymbol{\Lambda}^k$ is a transition probability matrix constructed so the distribution of type-k

households' locations evolve as,

$$q_n^k(\bar{\tau}') = \begin{cases} \sum_{\bar{\tau}'} \sum_{\bar{n} \neq n} \left[q_{\bar{n}}^k(\bar{\tau}')\, p_{n,t}^k \left(x(\bar{n}, \bar{\tau}') \right) \right] + q_{\bar{n}}^k(\bar{\tau} = 1)\, p_{n,t}^k \left(x(n, \bar{\tau} = 1) \right) (1 - g_n^k) & \text{if} \quad \bar{\tau}' = 1 \\ q_{\bar{n}}^k(\bar{\tau} = 1)\, p_{n,t}^k \left(x(n, \bar{\tau} = 1) \right) g_n^k + q_{\bar{n}}^k(\bar{\tau} = 2)\, p_{n,t}^k \left(x(n, \bar{\tau} = 2) \right) & \text{if} \quad \bar{\tau}' = 2. \end{cases}$$

Steady-State Equilibrium. We define a steady-state equilibrium in the year 2000 as a stationary distribution among our sample of low-income renter households, where continuation values are time-invariant and state variables are fixed at their 2000 level: $\bar{\omega}_t^{kc} = \bar{\omega}_{2000}^{kc}\ \forall t$.

The time-invariant continuation values in our steady-state equilibrium are precisely our steady-state measures of expected welfare, $\{V_{ss}^k(n, \bar{\tau})\}_{n,\bar{\tau}}$. We outline the procedure to compute these measures as well as the vector of exogenous amenities in Algorithm 1.[66]

We now turn to $\bar{V}^k(x = (n, \bar{\tau}), \bar{\omega}_{2000}^{kc})$, which is the expected welfare of a type-k household with tenure $\bar{\tau}$ residing in neighborhood n in the year 2000. In contrast to $V_{ss}^k(n, \bar{\tau})$, we do not assume the economy is in a steady-state equilibrium when constructing this measure. We instead compute $\bar{V}^k(x = (n, \bar{\tau}), \bar{\omega}_{2000}^{kc})$ via backward induction starting in the year 2019. We do, though, hold exogenous neighborhood amenities fixed at the values found in Algorithm 1. By holding the values of exogenous amenities fixed at their 2000 levels, we can attribute variation in $\Delta W^k(n, \bar{\tau})$ to neighborhood-level changes in rents and college shares. To initiate our backward induction solution method, we assume that the economy is again in a steady state in 2019 and calculate steady-state continuation values for the year 2019 as in steps 2-4 in Algorithm 1. Since we are solely concerned with expected welfare changes from the perspective of the year 2000, this assumption is innocuous given a discount rate of $\delta = 0.85$.

To compute $\bar{V}^k(x = (n, \bar{\tau}), \bar{\omega}_{2000}^{kc})$, we must also take a stance on how households form their expectations over the market-level state variables, $\bar{\omega}_t^{kc}$. One option is to assume that households have perfect foresight. We experimented with this assumption, but found that households quickly reoptimize their location choices given perfect knowledge of future states. Instead, we assume that households' expectations are a weighted average of i) the future true states; and ii) neighborhoods' current states multiplied by the CBSA-wide average growth rate of each state variable:

$$\mathbb{E}[\bar{\omega}_t^n | \mathcal{I}_{i,2000}] = \bar{\omega}_{2000}^n \cdot \left(1 + \frac{\sum_{n' \in \mathcal{N}^c} \left(\bar{\omega}_t^{n'} - \bar{\omega}_{2000}^{n'} \right)}{\sum_{n' \in \mathcal{N}^c} \bar{\omega}_{2000}^{n'}} \right) \cdot \mu + \bar{\omega}_t^n \cdot (1 - \mu). \tag{A9}$$

[66] In step 6 of Algorithm 1, we exploit the fact that the values of exogenous amenities are the same for households regardless of their length of tenure. While we cannot guarantee Algorithm 1 induces a stationary equilibrium among type-k households with a high level of neighborhood tenure, $\bar{\tau} = 2$, we find that in practice it closely approximates a stationary distribution for these households. We define $\hat{q}_n^k(\bar{\tau} = 1)$ as the empirical counterpart to $q_n^k(\bar{\tau} = 1)$.

Algorithm 1 Compute Steady State Expected Welfare

1: Guess value of unobserved neighborhood amenities: $\{\xi_n^k\}_n$.

Compute continuation values:

2: Guess $\{V_{ss}^k(n, \bar{\tau})\}_{n,\bar{\tau}}$.

3: Compute $V_{ss}'^k(n, \bar{\tau}) = \ln \left(\sum_{n' \in \mathcal{N}^c} \exp \left(\bar{u}_{n'}^k \left(x(n, \bar{\tau}), \bar{\omega}_{2000}^{kc} \right) + \delta V_{ss}^k(n', \bar{\tau}) \right) \right) + \gamma, \quad \forall\, n, \bar{\tau}$

4: Set $V_{ss}^k(n, \bar{\tau}) = V_{ss}'^k(n, \bar{\tau}), \quad \forall\, n, \bar{\tau}$

Repeat steps 2 - 4 until $\max_{n,\bar{\tau}} \{|V_{ss}^k(n, \bar{\tau}) - V_{ss}'^k(n, \bar{\tau})|\} < \epsilon_V$ for some $\epsilon_V > 0$.

Compute exogenous amenities:

5: Compute the probability a type-k household chooses n given $\{V_{ss}^k(n, \bar{\tau})\}_{n,\bar{\tau}}$ and $\{\xi_n^k\}_n$

$$p_n^k \left(x(\bar{n}, \bar{\tau}), \bar{\omega}_{2000}^{kc} \right) = \frac{\exp \left(v_n^k \left(x(\bar{n}, \bar{\tau}), \bar{\omega}_{2000}^{kc} \right) \right)}{\sum_{n' \in \mathcal{N}^c} \exp \left(v_{n'}^k \left(x(\bar{n}, \bar{\tau}), \bar{\omega}_{2000}^{kc} \right) \right)}, \quad \forall\, \bar{\tau}, \bar{n}$$

6: Update exogenous amenities using observed neighborhood shares of households with low tenure, $q_n^k(\bar{\tau} = 1)$:

(A8)

$$\xi_n'^k = \xi_n^k + \ln \left(\hat{q}_n^k(\bar{\tau} = 1) \right) - \ln \left(\sum_{\bar{\tau}'} \sum_{\bar{n} \neq n} \left[q_{\bar{n}}^k(\bar{\tau}')\, p_{n,t}^k \left(x(\bar{n}, \bar{\tau}') \right) \right] + q_{\bar{n}}^k(\bar{\tau} = 1)\, p_{n,t}^k \left(x(n, \bar{\tau} = 1) \right) (1 - g_n^k) \right)$$

Set $\xi_n^k = \xi_n'^k$.

Repeat steps 2 - 6 until $\max_{n,\bar{\tau}} \{|\xi_n^k - \xi_n'^k|\} < \epsilon_\xi$ for some $\epsilon_\xi > 0$.

Take $\{V_{ss}^k(n, \bar{\tau})\}_{n,\bar{\tau}}$ as our measures of steady-state expected welfare.

$\mu \in [0, 1]$ determines how accurate households' beliefs are. With $\mu = 0$, households have perfect foresight, and with $\mu = 1$, households believe all neighborhoods' state variables change at an identical rate. When $\mu = 1$, households are less mobile as they expect the relative levels of neighborhoods' flow utilities to remain constant over time. We calibrate μ to match the observed average neighborhood out-migration rates in our sample of low-income renter households (we assume μ is constant across household types).[67]

Throughout our welfare analyses, we use publicly available 2010-delineated tract-level Census Survey data from IPUMS National Historical GIS (NHGIS) to compute annual rents and the distribution of households across Census tracts (Manson et al. 2022). We linearly interpolate

[67] Note that equation A9 is consistent with our rational expectations assumption (Assumption (c)) in so far as the true data generating process renders equation A9 unbiased.

these data between survey years (2000, 2007-2019). For 2007 onward, we use the 5-year ACS aggregates to compute the tract-level data. As the publicly available tract-level data are not disaggregated enough to compute the exact shares each type-k low-income renter household across census tracts and aggregated tenure states, $\bar{\tau}$, we approximate the empirical neighborhood shares used in Algorithm 1, $q_n^k(\bar{\tau} = 1)$, with the share of Black (non-Black) non-college-educated renter households living in each tract who have a tenure of less than three years.